University of Iowa

Iowa Research Online

Theses and Dissertations

Spring 2016

Inference to the best explanation and the challenge of skepticism

Bryan C. Appley
University of Iowa

Follow this and additional works at: https://ir.uiowa.edu/etd

Part of the Philosophy Commons

Inference to the Best Explanation and the Challenge of Skepticism

Bryan C. Appley

A thesis submitted in partial fulfillment
of the requirements for the Doctor of
Philosophy degree in Philosophy
in the Graduate College of
The University of Iowa

May 2016

Thesis Supervisor: Professor Richard Fumerton

Graduate College
The University of Iowa
Iowa City, Iowa

CERTIFICATE OF APPROVAL

PH.D. THESIS

This is to certify that the Ph.D. thesis of

Bryan Christian Appley

has been approved by the Examining Committee
for the thesis requirement for the Doctor of Philosophy
degree in Philosophy at the May 2016 graduation.

Thesis Committee: ....

Richard Fumerton, Thesis Supervisor

....

David Cunning

....

Evan Fales

....

Ali Hasan

....

Gregory Landini

To Alaina, Daniel, and Cindy Appley

Acknowledgements

Abstract

In this dissertation I consider the problem of external world skepticism and attempts at providing an argument to the best explanation against it.

In chapter one I consider several different ways of formulating the crucial skeptical argument, settling on an argument that centers on the question of whether we're justified in believing propositions about the external world. I then consider and reject several options for getting around this problem which I take to be inadequate. I finally conclude that the best option available to us at the moment is to argue that the antiskeptical view is the best explanation of our ordinary experiences

In chapter two I argue that, if we hope to ground what counts as defending antiskepticism in common sense, there is an argument against the possibility of ever knowing one has succeeded in defending antiskepticism. After showing that common sense is no place to look in setting a goal for our antiskeptical project, I present the view that what will be crucial to settling on our antiskeptical goal is coming to a successful analysis of the nature of physical objects. I suggest some minimal criteria that must be met by a view in order to be antiskeptical based on our intuitions about core skeptical cases, but acknowledge that a fully successful response to external world skepticism will require the antiskeptic to engage in some much more difficult analysis.

In chapter three I consider various views of the nature of explanation and conclude, tentatively, that explanation as it interests the antiskeptic is fundamentally causal.

In chapter four I consider and reject some of the core views on which best explanation facts are so fundamental that a project of attempting to vindicate probabilistically the virtues which make explanations epistemically good. In this chapter I show that views which analyze justification in terms of best explanation factors fail.

In chapter five I attempt to vindicate the various explanatory virtues probabilistically. In doing so I attempt to express or translate the various explanatory virtues

in terms of probabilities in order to show that having those virtues makes a view at least prima facie more probable.

In chapters six and seven I explain and evaluate the various arguments to the best explanation against skepticism present in current philosophical literature. I attempt to show that extant arguments fail to appreciate the virtues possessed by classical (and some new) skeptical scenarios.

In chapter eight I briefly consider some options that may be open to the antiskeptic moving forward. All routes forward contain considerable obstacles, but there are some fruitful areas of research to pursue.

Public Abstract

This dissertation considers the problem of external world skepticism. Some say that we have good reason to believe that the ordinary world of 3D objects exists. They argue that the ordinary world is the best method available to us to explain our ordinary experiences of walking around, going to work and driving our cars. I argue that those who argue this fail to appreciate how well views like the hypothesis that we're in a Matrix-world explain our experiences. As far as human beings are concerned, we're just as likely to be in the Matrix as we are to be in the real world.

Contents

Chapter 1

Skeptical Arguments and Unhelpful Replies

1.1 Introduction

The following project is essentially skeptical. I will settle on a certain important skeptical argument regarding external world propositions, and, after provisionally dismissing approaches which I do not find promising, critically evaluate the major attempts at giving inference to the best explanation arguments against external world skepticism.

1.1.1 Skeptical Arguments

A number of skeptical arguments can be applied to our knowledge claims regarding external world propositions. The most famous comes in some variation of the following:

(K1) If it's consistent with S's evidence regarding an external world proposition p that S is merely deceived that p, then S does not know that p.

(K2) It's consistent with S's evidence regarding external world proposition p that S is merely deceived that p.

(K3) S does not know that p.

The crucial premise (K1) of this argument is usually justified by appeal to so-

called abominable conjunctions. For example, a conjunction of the form "I know p but maybe p is false," seems to involve an abominable combination of truth claims. These abominable conjunctions can only be done away with completely by using something like (K1). However, the classical skeptical argument loses much of its prima facie plausibility when applied to justification as opposed to knowledge:

(J1) If it's consistent with S's evidence regarding an external world proposition p that S is merely deceived that p, then S is not justified in believing that p.

(J2) It's consistent with S's evidence regarding external world proposition p that S is merely deceived that p.

(J3) S is not justified in believing that p.

What's the source of the plausibility of (K1)-(K3)? We've suggested that abominable conjunction considerations can motivate the crucial (K1) above. Such a motivation for (J1) is, however, not forthcoming. The conjunctions just don't seem to be abominable at all when construed in terms of justification. In fact, it seems essential to the notion of justification that one can say both that one is justified in believing p and that p might not be the case. The motivation might, however, be something like the following. Perhaps one's evidence is just everything that one knows. If this is the case, then being consistent with one's evidence that not-p entails that one does not know p. However, this motivation clearly leaves the (J1)-(J3) formulation in the dust. Further, it is intuitively very plausible that the notion of justification doesn't *exclude*, as a matter of analysis, that one may be justified in believing that p while it is consistent with one's evidence that p is false. One can be justified, problem of induction allowing, in believing that all crows are black without having seen every crow or having other evidence that entails the truth of "Every crow is black."

Now, there are other versions of the argument that center around knowledge. They generally take the form of the following brain-in-a-vat argument.

(C1) If S knows external world propositions, then S knows that she is not a brain-in-a-vat.

(C2) S does not know that she is not a brain-in-a-vat.

(C3) Therefore S does not know external world propositions.

This argument, rather than depending for its plausibility on intuitions like the abominable conjunction intuition, is generally thought to depend on a principle of closure. Knowledge is thought by many to be closed under known entailment. That is, if one knows that p and one knows that p entails q, then one knows that q.[1] However, what's important for our purposes is that there are people who have attempted to get out of this skeptical scenario by denying closure. If one isn't committed to closure, it is argued, then one isn't committed to **(C1)** being true, and thus has escaped the skeptical problem for knowledge.[2] If the force behind this argument is the closure principle, I'm not sure how big the problem of skepticism is. The epistemologist tied to the closure principle can simply give up that we have *knowledge*, but emphasize that she has a highly *justified* belief in the external world. It doesn't really seem to be *that* big a deal. I, however, am interested in a more radical sort of skepticism. I'm interested in a skeptical argument that will force us to really grapple with our justification with respect to external world propositions. The knowledge arguments, with all the epistemic baggage that the various versions carry, will fail to really get at this fundamental question. We should also avoid a modified version of the **(C1)-(C3)** style arguments that substitutes justification for knowledge here. Such an argument may, perhaps, produce an argument similar to our final one, but will raise questions that will be irrelevant to our purposes, e.g., the relevance of justification closure principles.

I propose that we settle on the following form for our skeptical argument.

---

[1]There are obvious complications here. It will turn out that one needs to believe q and believe that q on the basis of p and p's entailing q.

[2]The more explicitly closure-centric version of the argument would go something like this:
(C1*) If S knows that p and knows that $p \supset q$, then S knows that q
(C2*) S does not know that she is not a hand-less brain-in-a-vat.
(C3*) S knows that if she has hands then she is not a hand-less brain-in-a-vat.
(C4*) S does not know that she is not a hand-less brain-in-a-vat.

(P1) If S's evidence regarding p makes p no more probable than it does not-p, then S is not justified in believing that p.

(P2) S's evidence regarding p makes p no more probable than it does not-p.

(P3) S is not justified in believing that p.

The argument, put this way, will have two related aspects. It will be as easy as possible for the antiskeptic to answer. All the antiskeptic need do is show that her evidence makes external world propositions more probable than not. But this very ease of response also makes it one of the most dangerous skeptical argument forms. The antiskeptic is giving up a *lot* by failing to respond to this argument. The argument, unanswered, will be an epistemic disaster. However, a number of solutions to this argument are deeply unattractive.

1.2   Epistemic Externalism and Skepticism

On one way of thinking about the issue, it seems that externalists won't have a problem with external world skeptical arguments. It seems that if we look back at the skeptical argument we settled on, a certain very liberal understanding of the notion of evidence will allow the externalist to easily get out of the skeptical problem. One's evidence will make p more probable than it makes not-p just in case one's belief-forming processes that lead to the belief that p are reliable (or track truth, or are formed according to a truth directed design plan in the environment for which they were designed, etc.). There is no need to fret about skepticism. If one meets the externalist requirements for justification, one has avoided the thrust of the second premise.

But this advantage seems to melt away if what we're after is a first-personal *assurance* about external world propositions. The natural reaction of someone seeking such reassurance would be one of disbelief. Okay, maybe if the belief-forming process by which we get our external world beliefs are reliable, then my evidence in

some sense makes the external world belief probable. In whatever sense of probability the externalist uses there will be a conditional that the person fretting about skepticism can affirm: "If my beliefs about the external world are formed by a reliable process, then my evidence makes belief in the external world more probable than not." But this will be of little comfort to her. She can honestly agree with all of that and still be worried to death about skepticism. All *I* can get in a first-personal way is this conditionalized reassurance (i.e., *if* I meet the conditions of justification, then I'm justified). What help is this?

Michael Bergmann has his doubts about whether this is a special problem for internalism. He claims that externalism is in good company with respect to this kind of worry in that competitors to externalism are also only able to give a conditionalized response to skepticism. What gets him to his conditionalized response is primarily his note that the following are some commitments that externalists share with moderate non-externalists.[3]

> *inferentialism is false:* there are some conditions that, if satisfied by her noninferential beliefs, would be sufficient for their justification.
>
> *Strong Awareness Internalism is False:* One's noninferential beliefs can be justified in virtue of satisfying certain conditions even if one neither (a) believes that those conditions are satisfied nor (b) conceives of the fact that those conditions are satisfied as being in any way relevant to the truth or justification of one's beliefs (perhaps because one never applies any concepts to them at all).
>
> *The Low-standard View is False:* It is possible for all to seem well (for nothing to seem amiss) epistemically speaking even when the conditions necessary for noninferential justification are not satisfied.(Bergmann, 2008)

He goes on to argue that this means that moderate non-externalists will have the problem just as badly as externalists:

---

[3]He rules out strong awareness internalism as having any advantage in virtue of the very plausible claim that such internalists very quickly fall into a radical skepticism.

The serious moderate nonexternalists' view is that *if* their noninferential beliefs satisfy the relevant conditions, then they are justified. But this is merely a conditional response. The real question is whether that antecedent is satisfied or, instead, they are in a skeptical scenario, the skeptic will think that they cannot tell whether the antecedent is true. Moreover, because serious moderate nonexternalists reject strong awareness internalism, they will insist that they do not *need* to know or believe that the antecedent is true.(Bergmann, 2008)

This is meant to show that the internalist is in just as poor a position as the externalist with respect to being in a position to be sure about her justification. The point of the argument is to push those views that reasonably might be thought to have an advantage into a commitment to strong awareness internalism, yielding a doomed view. The sneaking here happens in Bergmann's rejection of the low-standards view. Asserting that it is "possible for all to seem well (for nothing to seem amiss) epistemically speaking" despite one failing to meet the qualifications for justification conceals the myriad differences that make crucial differences in this case.

In his book, Bergmann makes his point more clear, pointing out how his argument applies to the various internalist views. We shall concentrate on his argument that high-standards moderate internalists are subject to the problem of only giving a conditional response to skepticism.

Let's turn, finally, to High Standard versions of moderate nonexternalism. These views differ from others insofar as they restrict noninferential justification to beliefs about facts directly before one's mind. As a result, they don't allow for cases where one has a justified belief about some fact F1 (which is not directly before one's mind) in virtue of having directly before one's mind some distinct fact F2 (which is indicative of F1 without entailing it). Skeptical scenarios often capitalize on the gap created by having before your mind something which merely indicates without being or entailing the fact your belief is about. By avoiding this gap, the High Standard versions of moderate nonexternalism make themselves invulnerable to certain sorts of skeptical hypotheses.

But invulnerability to skeptical hypotheses relying on such a gap isn't sufficient for invulnerability to all skeptical hypotheses. For even supporters of High Standard versions of moderate nonexternalism (such as Laurence BonJour and Richard Fumerton) acknowledge that one can *think* one has some fact F directly before one's mind when one doesn't. As a result one can believe *p*, thinking one has the fact that *p* directly before one's mind, when in fact one doesn't. This suggests the following skeptical scenario: a demon arranges for all of one's introspective beliefs to be mistaken because each seems to be a belief about a fact directly before the subject's mind when in fact it isn't. That is, all the introspective beliefs of this demon victim are like the wildly inaccurate introspective beliefs that a malfunctioning person can have about the number of spots in her visual field: they seem to her to be about facts directly before her mind, but in fact they aren't.

Now how does the supporter of High Standard moderate nonexternalism know that she isn't a victim of such a demon? She can't point out that many of her introspective beliefs *seem* to be about facts directly before her mind because that is exactly how things would seem if she were in the skeptical scenario in question. And since she can't tell that she's not in such a skeptical scenario, it seems that none of her introspective beliefs are justified. To this, the High Standard moderate nonexternalist can only reply that she differs from the person in the skeptical scenario as follows: her beliefs are about things that are *in fact* directly before her mind whereas it merely misleadingly seems this way to the person in the skeptical scenario. But once again, this response is just as unsatisfying philosophically as the externalist's response to skepticism.(Bergmann, 2006, 221-222)

Bergmann's strategy, though the same as before is more clear in this passage. Take an internalist view and drive a wedge between first-level justification and what one has on the second-level. What exactly he's referring to on the second-level depends on which article one looks at. In the article for the *Oxford Handbook of Skepticism*, Bergmann talks about "all seeming well," while in *Justification without Awareness* considers the possibility of the subject *thinking* all is well (or thinking she is ac-

quainted with something in the specifically strong version of moderate internalism). "All seeming well" is, I think, intentionally vague as the argument in "Externalist Responses to Skepticism" is meant to remain at a level of abstraction from particular internalist views. The crucial point is that Bergmann has pushed discussion up a level to talk about our second level situation with respect to our first level beliefs. And perhaps his point is well-taken. If the accusation of the "merely-conditional" nature of externalism vis-a-vis skepticism is a question about whether the externalist can have second level justification that she's met the externalist justification conditions, then it seems right that a wedge can be driven, in most views, between first-level justification and second-level justification.

In this sense, it can be granted that the externalist has successfully given a good company defense. But the real force of the "merely conditional" charge against externalism comes from a different direction. Suppose, instead, that we give externalists and internalists as ideal an epistemic situation as possible. Suppose that the externalist has reliable beliefs several levels up (e.g., a reliable belief that p, a reliable belief that the belief that p is reliably formed, etc.). Now look at this case from the first-person perspective. Suppose the person here in question has such beliefs but is worried about skepticism. How does this help our worried subject? "From inside," as it were, the individual has no reason for optimism at all.

Now, let us give our individual an ideal strong moderate internalist epistemic situation. The individual is acquainted with her pain, acquainted with the correspondence between her belief and the pain state, and even acquainted with the fact that she is so acquainted. The person, here, is certainly in a much better place than the person without it if we're wondering about the first-personal *assurance* that this person has regarding p.

In the bad case, the high-standards person is in a better place assurance-wise as well. Suppose our worried individual has unreliable beliefs regarding p but also has

8

an unreliable belief that the belief that p is reliably formed.[4] This it seems, is a disastrous situation.

The high standards internalist is perhaps in a better situation. Let us consider a reasonable version of a bad case for the internalist. It takes some doing to construct a bad case for the internalist, but suppose our worrier is slightly color blind, such that certain shades of green and orange are difficult to distinguish. Now suppose she owns a device with a power button that glows a shade of orange when off and a shade of green when on. This device is off (and thus the power button glows orange), but when our worrier looks at it she forms the belief that she is having an experience of the color green. She's acquainted, actually, with some determinable property between orange and green. But when she reflects on her own mental states, she incorrectly takes herself to be acquainted with her being acquainted with the color green, and concludes at the second-level that her belief is justified. Our worrier is in a bad situation, but the kind of error that's possible for the high standards internalist is an error in which one can only be slightly off. Further, first-personally, the high standards internalist is even in a better place assurance-wise with respect to the bad epistemic situation. The internalist can often very easily move from the bad case to a good case with more careful introspection. In fact, I have done the very thing above in the past, but no longer commit such errors. But this was only after much careful introspection

Further, moving to the second-level isn't all that important for assurance anyway. Fumerton raises this in "Epistemic Internalism, Philosophical Assurance, and the Skeptical Predicament".

> The matter is, I think, quite different with what I call acquaintance. I stub my toe and I believe that I am in excruciating pain. What justification do I have for thinking that I'm in pain? How do I know that I'm in pain? My answer is that I am directly aware of the pain itself—the very truth

---

[4]It's not clear that this is worse than merely having an unreliable belief that p.

maker for my belief. The pain is "there" transparently before my mind. The thought that is about the pain and the pain that is its object are both constituents of the conscious mental state that I call acquaintance. When all this is so, we are in state that is all that it could be by way of satisfying philosophical curiosity. What more could one want as an assurance of truth than the truth- maker there before one's mind? When one is directly acquainted with pain as one entertains the proposition that one is in pain, there seems to me to be no need, no point, in moving up a level and asking about the justification one has for believing that one is in this state. It is not that one can't ask the question. The question is well-formed and there is, of course, a readily available answer. Just as acquaintance with pain was a completely satisfying way of assuring oneself that one is in pain, so acquaintance with this acquaintance with pain is a completely satisfying way of assuring oneself that one is acquainted with pain.

But again, I would emphasize that it doesn't strike us as even relevant to explore the second-level question as a way of getting a better sort of assurance that one is in pain. Why would it? If I'm right, what is relevant to getting the assurance one wants as a philosopher is getting the pain itself before one's consciousness. In the second level act of acquaintance the pain is present before consciousness again as a constituent of a more complex state of affairs, but having it before consciousness in that way is no better, so to speak, than having it there as an object of first-level awareness.(Fumerton, 2011, 189)

The conditional worry for externalism and skepticism is not a worry about whether the externalist can go up to the second level. The worry is about assurance and, in this case, a high standards moderate internalist (or, as Bergmann says, nonexternalist) will make much better company than the externalist.

After making a good company defense, Bergmann then explicitly addresses the "conditional response" objection. He says that nonexternalists are happy with a subject not knowing or even believing the proposition that she's justified in believing p while that subject is yet justified in believing p.(Bergmann, 2008) But more

fundamentally he takes on Sosa, etc., in saying that it's perfectly acceptable for the externalist to assert that the higher order proposition is also justified in an externalist way. He accuses internalists making this charge of assuming that the standards for higher order justification are different than the standards for lower order justification when the externalist is concerned.[5]

This is right to an extent, but as we have seen it's not the case that the fundamental point behind the "conditional response" objection is to demand something about higher order justification. The point is, rather, that the individual worrying about skepticism being told that all one needs is a reliably formed belief (to take the simple case) is thereby given no first personal comfort, let alone assurance, regarding the existence of the external world or any other belief for that matter. The person confronting skepticism is not asking a question for the epistemic externalist (perhaps so much the worse for the externalist). This person is not asking a question about whether others can regard her as justified in believing in the external world. The question is about what reason there is for her, first-personally, to believe in the external world and why.[6] The question is about philosophical assurance with respect to the existence of the external world. We should seek an answer to the problem of external world skepticism outside of epistemic externalism.

1.3   Content Externalist Responses to Skepticism

There is another sort of response to external world skepticism that begins from externalism. This one begins from externalism about *content* rather than *justification.* Hilary Putnam, famously formulated the response grounded in the very meaningfulness of the sentence "I'm a brain in a Vat."(Putnam, 2000) The problem for the

---

[5]Fumerton takes this hesitancy on the part of some externalists to apply the same externalist standards up a level to belie the dissatisfactory nature of externalism with respect to providing philosophical assurance.(Fumerton, 2011, 190)

[6]In a real way this objection may end up being just a different way of pushing Norman-like concerns for externalism in general, but it stands nevertheless.

skeptic is that, if Putnam is right about how words gain their meaning, then it would be impossible for a brain in a vat to meaningfully formulate the statement "I'm a brain in a vat," in such a way that it means anything like what we ordinarily take it to mean. The words "brain" and "vat" get their meaning from our interacting with external world objects, i.e., *brains* and *vats*. If we were always brains in vats and never came into contact with external world objects, then it would be impossible to meaningfully formulate the brain-in-a-vat scenario.

One major response to Putnam's objection to brain-in-a-vat scenarios comes from the fact that, even if it succeeds, it fails to rule out envatment scenarios. That is, it fails to rule out cases where one begins life in a real world and then one's brain is removed and put into a vat, thereby being subjected to a brain-in-a-vat deception. This individual's language has the right history, and so the individual will be able to formulate the brain-in-a-vat scenario without problem. Some interesting work has been done by Jochen Briesen (2008) suggesting that the envatment response suggests a hybrid approach. Briesen suggests that, in response to this argument, the defender of the real world hypothesis merely need add an inference to the best explanation on top of Putnam's externalist argument and the gap can be bridged. It certainly better explains my experience that I'm in and am experiencing a real world than does the proposal that I'm a recently envatted subject of the brain-in-vat deception. The envatment scenario requires there to be *both* a real world and a mirroring magnetic pseudo-world. This is clearly a less simple scenario than the regular real world hypothesis. If best explanation arguments work, then perhaps this is a good argument against external world skepticism.

However, Keith Butler mentions another interesting criticism of Putnam's argument that isn't open to Briesen's revision, and this problem is that externalism about content seems to be vindicated not a priori but a posteriori. He argues

It is crucial to the argument that there be a fact of the matter about

what "water" means on Earth. A fact such as this may be stipulated for the purposes of the argument, but the extent to which the argument is compelling is the extent to which it is shown that there is a fact of the matter about what a word like "water" means. If Putnam or Burge were to invent a new word, stipulate its meaning on Earth, and proceed to argue in the manner just presented, the argument would convince no one, because it would fail to show that Twin Oscar's concept is distinct from Oscar's. It is the empirical grounding of the argument, that water is H2O and not XYZ, that compels us to think that the Twin Earth environment would not contain any water, and hence that Twin Oscar's concept is, by the principle enunciated, different from Oscar's.

Moreover, the principle of concept acquisition Burge and Putnam are offering seems itself to depend for support on empirical facts about cognitive development. There is, Burge and Putnam claim, no other way to acquire a concept than by exposure to the extension of the concept or to other speakers who have had exposure. It is hard to imagine how this principle could be supported without appealing to the normal cases that instantiate the principle. These normal cases, of course, are not available to us a priori, nor is the fact that they are the normal cases. (Butler, 2000, 45-46)

The problem for the Putnam response on Butler's view is that the response begs the question. One must already know there is water out there and that water is H2O in order to produce the intuition that justifies externalism. It is not enough to construct two different worlds, one with H2O and one with XYZ, as it is our strong rational conviction that water in fact *is* H2O and not XYZ that justifies us in moving toward externalism. Further, even the story of how concepts are acquired requires that we have enough justification with respect to the external world not merely to believe that it is a world of spatial objects, but that it is a world of objects that are of natural kinds. Both of these truths that are what would motivate externalism are justified only after justification in rejecting skepticism is already present. For these reasons I am not optimistic about content externalism's ability to avoid our skeptical problem.

1.4 The Moorean Response

Moore would respond to the argument for skepticism with what has lovingly been named a "Moorean Shift." His most sophisticated thought on this point[7] comes in his response to an argument from Russell in "Four Forms of Skepticism." His response to skeptical arguments goes as follows:

> What I want, however, to emphasize is this: Russell's view that I do not know for certain that this is a pencil or that you are conscious rests, if I am right, on no less than four distinct assumptions: (1) That I don't know these things immediately; (2) That they don't follow logically from any thing or things that I do know immediately; (3) That, if (1) and (2) are true, my belief in or knowledge of them must be 'based on an analogical or inductive argument'; and (4) That what is so based cannot be *certain knowledge.* And what I can't help asking myself is this: Is it, in fact, as certain that all these four assumptions are true, as that I *do* know that this is a pencil and that you are conscious? I cannot help answering: It seems to me *more* certain that I *do* know that this is a pencil and that you are conscious, than that any single one of these four assumptions is true, let alone all four... I agree with Russell that (1), (2) and (3) *are* true; yet of no one even of these three do I feel *as* certain as that I do know for certain that this is a pencil. Nay more: I do not think it is *rational* to be as certain of any one of these four propositions, as of the proposition that this is a pencil. . . .Moore (1966)

If we apply this to the argument form we've settled on we get this sort of response. We're so certain that we're justified in believing external world propositions that we're more certain of this than either premise of the argument. Let's test this assertion. **(P1)** states that If one's evidence makes p no more probable than not-p, then one is not justified in believing p, while **(P2)** states that in the case of external world propositions, one's evidence doesn't make the external world propositions

---

[7]For an argument as to why this version of his argument is the most sophisticated, see Lycan (2007).

more probable than not. There are problems with adapting Moore's response to this argument.

The problem comes if we take this response and make it into a Moorean shifted argument. That is, problems arise if we swap *modus tollens* and *modus ponens* to get the following argument form:

(MP1) If S's evidence regarding p makes p no more probable than it does not-p, then S is not justified in believing that p.

(MP2) S is justified in believing that p.

(MP3) Therefore, S's evidence regarding p makes p more probable than it does not-p.

A main line of criticism of this kind of shifted argument is that the argument fails to transmit knowledge from premise to conclusion. Now, transmission failure amounts to an epistemic equivalent of an accusation of question-begging or of circularity, depending on how one looks at the case. The idea is simply that an argument, to be good for an epistemic agent, must work in the right way. Our evidence must support the premises, and our premises must support the conclusion, while (at least part of) our justification for believing the conclusion is the satisfaction of the previous clauses. But our justification for **(MP3)** in no way can be construed as relying on the premises **(MP1)** and **(MP2)**. In fact, **(MP2)** could only be supported by premises that independently support **(MP3)** and so **(MP2)** could not be a premise in a transmitting argument from **(MP1)** to **(MP3)**. Our justification for **(MP2)** and **(MP3)** are just the same evidence, without the need for **(MP1)** at all. Further, the argument isn't even very enlightening. The person worried about skepticism, especially of the kind discussed above, is worried about *how* they can rationally be comforted with respect to external world propositions. This, at the very least, should require showing what the justification is. The concerned individual should look outside of Mooreanism for a response to the kind of external world skepticism in **(MP1)** through **(MP3)**.

1.5 Seeming Internalism's Response

In a way the seeming internalist's response is very simple and intuitive. The simple assertion is that one's having a seeming that *p* gives one *prima facie* foundational justification for believing that *p*; seemings make epistemically probable the truth of their contents. For example, the seeming I have that my tea is lukewarm justifies me in believing the proposition "My tea is lukewarm." But for the important external world propositions, it surely *seems* to us that they are true. Therefore we're justified to some degree in believing external world propositions and the seeming internalist isn't subject to the skeptical argument considered above.

A major problem with the seeming internalist's response to skepticism is that it has gotten either the nature or the epistemology of seemings wrong.[8] Consider a trivial case of a seeming, such that it seems to me that my shirt is blue. What kind of experience is this? If we separate mere seeming from other possible kinds of experience like acquaintance or bare visual experience,[9] all that a seeming seems to be is, perhaps, a felt inclination to believe.

Andrew Cullison argues that this cannot be the correct understanding of a seeming. He gives the following counterexample:

> Imagine Sam's wife suddenly undergoes a radical religious conversion. After her conversion Sam starts to worry that she might divorce him if he doesn't eventually share her new religious beliefs. Sam's love for his wife might cause him to feel an attraction toward believing some propositions of her new found religion. There may be some psychological compulsion or pull in these cases. But it seems possible to feel such an attraction or pull without there being any kind of seeming state that the propositions are true....
>
> When I was younger I found myself agreeing with my father on political

---

[8]See Hasan (2013) for an interesting response to Huemer (2007) and his argument that one ought to adopt a seemings internalist view.

[9]A "bare visual experience" may not be a category for seeming internalists, since that becomes part of a visual seeming.

issues. I would ask my dad why he thought Mr. X should be president. He would give me a few principles of government that he thought were true, and then he would tell me that Mr. X would best promote those principles. I then came to have inclinations to believe certain propositions about what governments ought to do and what a good political candidate ought to promote. I was very much inclined (and disposed) to accept certain propositions that my father believed.

Today, I think I've thought about the issues more carefully, and I also think there is a new kind of attitudinal state I have toward some of these propositions that is different in a significant way from the attitudinal states I had toward those propositions when I was younger. When I was younger I was merely inclined to believe these propositions, but now I think that I am actually apprehending a truth in some significant way.(Cullison, 2010)

But Taylor (2015, 376-377) points out that in the first case we should be careful to distinguish between the conscious inclination to believe and other forms of attraction. Sam may be attracted in some way to the propositions his wife believes, but it is important to resist saying that this constitutes a felt inclination to *believe*. It is more readily described as a case where Sam wishes he were in a state where he were at all inclined to believe p. In fact, Sam may take steps to attempt to put himself in such a situation. However, it's not a slam dunk case that Sam has a felt inclination to believe. (Sam, in certain ways of filling out this case, may also actually be consciously inclined to believe, but these are cases where it is intuitive to describe Sam as having a seeming regarding such positions). Further, Taylor suggests that the difference in the case of Cullison's childhood and adulthood political beliefs can be accounted for by the difference in directness of the seeming. Adult Cullison is inclined to accept the political propositions based on a fuller understanding of the issues and principles involved, while child Cullison is believing based on testimony. In any case, it requires further argument to establish that seemings are a *sui generis* kind of mental state. But if seemings are reducible to something like felt inclination to believe, then their

epistemic import comes into doubt.[10]

Further, supposing seemings are a *sui generis* kind of state, it yet seems that sometimes I don't have a seeming with respect to external world propositions. Take for example the proposition that there is a painting on the wall across from me. Now, I'm not at my best today, and I am still waiting for my caffeine to kick in, so I'm a little groggy and tired. Does it somehow "strike me" that it's true that there is a painting on the wall across from me in the way the seeming theorist wants it? I might have such a seeming if my brain were running on all cylinders, but right now I just don't have such a seeming. Now, there are things I do have. I have a painting-like sense datum (or am being appeared to paintingly, etc.) and I, being the hopeful antiskeptic that I am, take myself to be justified in believing that there is a painting on the wall. But how is the seeming theorist to vindicate my having justification for the belief that there is a painting on the wall? Seemings won't do the work. I think seeming theorists, even provided their view can be rescued from above worries, will need to supplement their theory with another response to skepticism.

1.6 Steup's Heterodox Evidentialist Response

Steup takes an interesting and rather different tack with respect to skeptical arguments. We'll have to modify Steup's response to work against our argument, but before doing so it will be useful to consider his argument as it applies to a classical skeptical argument. Consider the Brain-in-a-vat argument as abbreviated by Steup:

(B1) If I know that I have hands, then I know that I am not a brain in a vat.

(B2) I do not know that I am not a brain in a vat.

(B3) Therefore, I do not know that I have hands.

Steup takes on the challenge of responding to this argument without giving up closure and without giving up evidentialism. In doing so, he starts by considering

---

[10]For a more detailed discussion of the nature of seeming and of reasons to think they are conscious inclinations to believe (or at least not *sui generis* states), see Taylor (2015) or Byerly (2012).

less popular forms of skeptical argument. Specifically, he discusses round-square skepticism and Easter bunny skepticism. The arguments go precisely parallel to the brain-in-a-vat arguments, except with different deceptions. Rather than be deceived by a mad scientist tweaking a brain-in-a-vat, one is deceived by a round square or an Easter bunny respectively. He builds his case by starting from the consideration of the round-square deception. Steup asks how we know that we are not subject to a round-square deception, easily concluding that we have *evidence* against the truth of the round-square scenario, namely that round squares are contradictory. Steup also thinks that we can dismiss Easter bunny deceptions for reasons such as that bunnies don't have dexterous enough limbs to be deliverers of eggs into children's baskets. That is, we have evidence to think that there can't be an Easter bunny deception because we have evidence against there being Easter bunnies.(Steup, 2011)

Now, finally, Steup returns to the brain in a vat scenario. He asserts that we can dismiss brain-in-a-vat scenarios for similar reasons. We have evidence against the brain-in-a-vat scenario. Steup gives a list of the evidence that we have against the existence of BIVs:

1. Textbooks of neurophysiology don't have a chapter entitled 'Envatment'.

2. Departments of neurophysiology don't offer courses entitled 'Envatment 101'.

3. If you bother to call a renowned neurophysiologist or brain surgeon and ask whether envatment is possible, the answer is going to be 'no'.

4. Essay collections for courses on applied ethics don't have a chapter entitled "The ethics of envatment'.

5. No known episode of *60 Minutes* has ever investigated let alone asserted the existence of BIVs.

6. There is no known case of someone ever having been sued for or found guilty of envatting a person.(Steup, 2011)

Steup presents an extensive list of evidence against the existence of BIVs. After this he defends the argument against objections that the argument is question-begging

and against the argument that the kind of evidence for 1-6 is different in kind from the evidence against the existence of round squares (it's argued that it requires one to already know one isn't a BIV in order to use it as evidence). It will turn out that the justification centered argument avoids completely Steup's argument, even if Steup is able to address objections satisfactorily.

Let's suppose that the evidence that Steup gives is good evidence against the BIV scenario. Let's even suppose that this evidence makes the real world scenario more probable than the brain-in-a-vat scenario. This leaves all the other scenarios. Now, a defender of the Steup line might argue that there's evidence against any one skeptical scenario (although I'm not sure this kind of line would work against demon scenarios), but this will still not yet establish that one's evidence makes the real world scenario more probable than not. It could be that the real world hypothesis is a little more probable than a number of skeptical scenarios while it is more probable than not that some skeptical scenario or other is true.

One can even have reason to believe that one option is more probable than a number of other options without *any* of them being particularly probable. For example, consider the once popular McDonald's Monopoly game. This game involved upselling customers to higher priced foods with the promise of a game with valuable prizes. In fact, if one collected both the boardwalk and the park place pieces, one would win one million dollars. Now, one could simply look at the various prizes and their relative costs and get a good idea of the relative probability of any two prizes. For example, it's fairly safe to say that one is more likely to get the $1,000$ shopping spree than the new car, or the new home entertainment system, or the vacation in Maui, or the million dollar prize. But it's still very unlikely that one wins *any* of those prizes, despite the fact that we're justified in believing that the shopping spree is more probable than any of the individuals we've considered.

But Steup overlooks a distinction in the kind of evidence in the BIV case versus the

kind of evidence regarding square-circle and easter bunny deceptions. The evidence we have against the existence of round squares is a priori and probably doesn't require that we're justified in believing external world propositions at all in order to be justified in rejecting the existence of round squares. For easter bunnies, it's at least reasonable to think that the evidence he proposes is a priori. It's reasonable to think that we can be justified in believing external world propositions. But all the pieces of evidence that Steup appeals to are pieces of evidence only if we're justified in believing external world propositions. Steup can yet make this work for him in his argument. One might allow justified (but below knowledge-level) external world propositions plus the evidence Steup cites to make the difference between the mere justified belief that I have hands and the knowledge that I have hands.

But if Steup's argument can work on the knowledge question, then it will have to fail on the level of justification. The only way to make Steup's argument work on the knowledge level is to *give* him justification for the antiskeptical propositions. But there is no analogous move available to the justification argument. Steup's heterodox evidentialist response to external world skepticism fails to deal adequately with our current skeptical argument.

1.7   What Are the Data?

Now, on any antiskeptical strategy it is by appeal to sensory experience that we can justify belief in external world propositions, so the next preliminary to be considered concerns the nature of sense experience as we will consider it here. Now, I'm not interested in getting into a big argument about adverbialism versus sense data theory or anything of the like, so the following will be fairly conciliatory and tentative.

The one crucial point to assert is that in the following work we will be assuming that the sense data (or adverbial experiences or what have you) that we have access to

are subjective mental states of the individual. The antiskeptical arguments presented below all work from such a view. Now, this doesn't necessarily mean that it will be important to dismiss such views as direct realism or disjunctivism which, arguably, will have a better time with an inference to the best explanation against skepticism. Such views, though gaining more traction these days, are still minority views and carry with them much stronger commitments. The best route for the antiskeptic is to show that one can get out of skepticism even on the prevailing (and more minimal) understanding of experience. Even if direct realism or disjunctivism will give an advantage on the particular point of addressing external world skepticism, there are problems with direct realist and disjunctivist theories of experience which, perhaps, will make the cure worse than the disease. Some discussion of this and other ways to avoid skepticism will follow in the final chapter.

The following will consist of two principal sections. In the first section I will consider the necessary preliminaries to the reasonable defense of the inference to the best explanation argument against skepticism. In that section I will spend a chapter considering problems for nailing down exactly what counts as defending antiskepticism. After this I will spend three chapters considering, in detail, the nature of explanation as it relates to the kind of argument to be presented and the nature and extent of epistemic support that inference to the best explanation can give. After considering all this, the next major section is devoted to a detailed critical examination of the major competing versions that argue to the best explanation against skepticism. After finding all these arguments wanting, I spend the final chapter taking stock of the situation for aspiring antiskeptics, considering possible alternatives where antiskeptical hope may remain. But let's not get ahead of ourselves. What exactly is it to be antiskeptical anyway?

Part I

Before the Abduction

Chapter 2

The Common Sense Trap

2.1 Introduction

Before considering specific IBE arguments against skepticism, it will be interesting to think about the very nature of the antiskeptical project. It also seems important to get straight what exactly counts as an antiskeptical view before considering the specifics of an argument that purports to be antiskeptical (these considerations will also be important if these arguments fail and it's important to try to pick up the pieces). The current defenders of inferences to the best explanation (IBE's) against skepticism often allege a tight connection between common sense and the nature of the antiskeptical views they defend.

Kevin McCain, perhaps most strongly, says that the "*explanationist response to skepticism* involves using inference to the best explanation to argue in support of our commonsense external world beliefs."(McCain, 2014, 125) Vogel says, "We each adopt a body of common-sense beliefs about the world which answers to our sensory experience. In principle, however, the beliefs we base on that experience are subject to under-determination, and we can devise radical alternatives to the common-sense account."(Vogel, 1990, 658) Bonjour says that the view he's defending is "the approx-

imately commonsensical idea that my sensory experiences are systematically caused by a realm of three-dimensional objects."(BonJour, 2003, 92) One might reasonably conclude from this that the current defenders of inference to the best explanation arguments against external world skepticism take defending 'our common sense external world beliefs' to be constitutive of defending external world antiskepticism.

In the following I will, by giving an analysis of common sense, show that pinning one's hopes for a reasonable non-question-begging construal of antiskepticism results in problems. After showing that there is a problem with relying on common sense to characterize what it takes to be antiskeptical, I will consider the less problematic view that to be an antiskeptical view is to be a view that rejects skeptical scenarios. The natural question is then what it takes to be a skeptical scenario. A skeptical scenario will be one, roughly, on which material objects are eliminated. If this is the case, then, in order to properly vindicate antiskepticism as requiring those common elements which the major IBE defenders share, the defender of antiskepticism will want to go some way toward a successful analysis of the nature of material objects.

If we take the common-sense-centric view seriously, there is a threat of a problem that parallel's the paradox of learning presented in Plato's Meno, where it is worried that education is impossible. If I had no knowledge of a subject, then I wouldn't know where to start in learning. If I already had knowledge of a subject, then I can't learn because I would already know. But then it seems impossible to ever gain knowledge. In ways that should be clear, the following problem for antiskeptics will mirror Plato's Meno paradox:

(1) If we don't antecedently know what counts as an antiskeptical view, then we will never know we've succeeded in the antiskeptical project.

(2) If we *do* know what counts as an antiskeptical view, we must have already succeeded in our antiskeptical project.

(3) Therefore, we can never succeed in the antiskeptical project.

(2) is clearly the controversial premise, but it will become clear in what follows

that it is by no means easy to escape the quasi-Meno paradox. In order to escape we shall have to find some way of defining the antiskeptical project that doesn't require us to already be justified in believing in a robust external world. One candidate for avoiding the problem will be to find a priori justification for believing that such-and-such counts as succeeding in one's antiskeptical project.

One natural place to look to find the antiskeptical project is in common sense. Perhaps defending common sense about external world objects will be enough to succeed in being an antiskeptic. This turns out, as alluded to above, to be deeply problematic. One of the most natural proposals for what counts as common sense says that it is the naive position from which human beings uncritically start prior to sophisticated science. Quine says:

> Rudimentary physical science, that is, common sense about bodies, is thus needed as a springboard for scepticism. It contributes the needed notion of a distinction between reality and illusion; and that is not all. It also discerns regularities of bodily behaviour which are indispensable to its force to our knowledge that sticks do not bend by immersion; and his examples of mirages, after-images, dreams, and the rest are parasitic upon positive science, however, primitive.... I am ... making the point that sceptical doubts are scientific doubts.(Quine, 2004, 287-288)

On this view, which is shared by Russell (1940, 14-15), common sense is just the naïve view of the world. But if we aren't to doom the antiskeptical project from the start, it ought not be the case that we must give a full-bodied defense of the naive view of the world[1] in order to count as having succeeded in defending antiskepticism. This naïve view of the world is a starting point for science, but it is also a ladder doomed to be kicked away once we develop a sophisticated scientific understanding of the world.[2] Consider the following short passage from Eddington regarding the "two

---

[1]The exact nature of "the naive view of the world" is actually pretty difficult to nail down as well, but we can leave this problem aside for now.

[2]For Quine, however, as a coherentist, the naïve view of the world is instead organically developed over time through scientific practice.

tables."

> There are duplicates of every object about me—two tables, two chairs, two pens.... One of them has been familiar to me from earliest years. It is a commonplace object of that environment which I call the world.... It has extension; it is comparatively permanent; it is coloured; above all, it is substantial.... If you are a plain common-sense man, not too much worried with scientific scruples, you will be confident that you understand the nature of an ordinary table. My scientific table is mostly emptiness. Sparsely scattered in that emptiness are numerous electrical charges rushing about with great speed; but their combined bulk amounts to less than a billionth of the bulk of the table itself... I need not tell you that modern physics has by delicate test and remorseless logic assured me that my second scientific table is the only one which is really there—wherever 'there' may be. (Eddington, 1935, 5-8)

We normally speak as though there is a solid table in front of us. But according to scientific discoveries, we now know that this table is made up of tiny subatomic particles that are floating in an expanse that is so great relative to the size of the particles that the table is mostly empty space. Any classical naive understanding of "solidity" and "color" will have to be abandoned. The world is almost certainly not the world as seen by the naive. Certainly the antiskeptical project's success cannot be thought to depend on the defense of such an untenable view. In fact, supposing we are able to know a priori that this is what it takes to be a successful antiskeptic, we will have gotten out of our little Meno paradox by dooming ourselves to failure in a slightly different way.

But perhaps the prospects of grounding the nature of antiskepticism in common sense aren't so bleak. Before proceeding with our discussion it will be important to take a longish detour on the nature of common sense.

2.2 The Social Nature of Common Sense

I'm primarily interested, here, in those *propositions* that are called common sense or commonsensical for an individual. People take certain propositions to be commonsensical and speak metaphorically of common sense telling them things (e.g., one says, "common sense tells us that there are spatial objects"). It is this kind of common sense that provides interesting fodder for discovering the nature of antiskepticism.

There are three basic approaches one might take to understanding the notion of propositional common sense. One might analyze common sense epistemically, psychologically, or socially. I will first consider epistemic and psychological analyses (with the hope of settling on a notion which doesn't give rise to the quasi-Meno problem discussed above). Finally I will settle on a social notion of common sense and consider other avenues for escaping the problem.

2.2.1 Epistemic Notions of Common Sense

Foundational Justification Common Sense (FJCS)   p is a common sense proposition for subject S if and only if p is foundationally justified for S.

Several central common sense cases seem to fit **FJCS**. For example, 2+2=4 is a proposition that we will want to come out commonsensical on any adequate view. It's also desirable that propositions such as "Red is a color," turn out commonsensical. However, apart from these central cases, there are few advantages to this view. There are cases where the extension seems to be wrong in both directions. First there seem to be common sense propositions that are not justified foundationally. For example, "objects thrown into the air almost always return from the heights to which they were thrown" seems to be an item of common sense. However, this proposition is almost certainly only justified inferentially, whether by induction or somehow by inference to the best explanation.[3]

---

[3]Huemer and other conservatives may disagree here, but on such views it's hard to find much

Further, this view will call some things common sense which are intuitively not so. Consider someone with very refined taste buds. Such a person can be foundationally justified in certain beliefs about her sense impressions that may not be commonsensical at all. For example, suppose Wendy is a wine connoisseur who's developed this capacity on her own and is foundationally justified in believing that a taste sensation (which happens to be caused by a rather strange brand of wine) contains notes of green apple with white peach undertones and a rose petal finish. This deeply refined awareness of one's sensory experiences is the sort of thing that might be foundationally justified while not being commonsensical at all.[4]

Suppose, instead, that one adopts HJCS:

High Justification Common Sense (HJCS)   p is a common sense proposition for subject S if and only if S is highly justified in believing that p.

This view is a natural move from FJCS given the need to countenance propositions that are non-foundationally justified in this discussion of common sense. However, merely defining common sense in terms of justified propositions will not do, since everyone will grant that common sense propositions are not co-extensive with justified propositions. People can be justified in believing propositions that are contrary to common sense.

So maybe if one moves the justification threshold up quite a bit one will have captured an explicitly epistemic notion of common sense. However, there are problems. First, there are propositions one may be highly justified in believing that are not common sense. The refined taste case above will suffice to establish this.

There is a further problem for any view that takes common sense propositions

---

of anything that's justified that isn't to some extent justified foundationally. Nevertheless, in the case above, it does seem that it will be some kind of argument that gets justification for this claim. While it may be justified to some extent by seemings for Huemer, it seems that the bulk of this proposition's plausibility comes from some kind of non-deductive argument.

[4]If Wolterstorff (2001) is right, one way of reading Reid's understanding of common sense is as a set of "shared first principles." If this is the case, then Reid may be holding a modified version of FJCS. This proposal will be ill-fated, however, as it falls to several of the weak common sense cases proposed later.

to be co-extensive with propositions that are in some way justified. The problem is that one can be highly justified in *rejecting* common sense propositions. Science is resplendent with examples of common sense propositions which many human beings have become highly justified in rejecting (e.g., the Earth is motionless had been common sense for the first person who became justified in believing that the Earth moved).[5]

Perhaps the problem with the above epistemic proposals are that they're looking toward the wrong sort of justificational structure to get an epistemic notion of common sense. But consider the following proposal:

Center of Web Common Sense (CoWCS)   p is a common sense proposition for subject S if and only if p is close to the center of S's web of belief.

On this view, those propositions are common sense which, were one to give them up, one would have to modify a lot of one's other beliefs in order to retain a coherent system of beliefs. This account can deal with a number of intuitive cases of common sense beliefs. Beliefs such as those about the external world and mathematical truths are allowed in while even beliefs such as Craig the Creationist's[6] belief that the Earth is around 7,000 years old can be countenanced by this view (I take it that, for Craig's belief to change, a lot of other beliefs would have to change as well, such as beliefs about science and religion and even, perhaps, the nature or existence of the divine).

Consider Silly Simon, who has a really odd belief web. Simon's belief web is such that the proposition "I'm currently having an experience as of a hexagonal green pot holder," is at the very center of his web of belief. If it goes, so does most of the rest

---

[5]Some of the examples of common sense propositions apparently overturned by science that occur here or below will assume a sort of view about the history of humanity's view of things that is (according to some) naïve. Nevertheless it needs to be accounted for that if such a (perhaps naïve) view of humanity's historical views is right, then these sorts of cases are everywhere. That is, even if all of these proposals never in fact were propositions widely believed (and questions widely thought to be settled), it still requires accounting for that, a world where this is the case is a problem for the analyses of common sense there discussed.

[6]For discussion of Craig the Creationist as he relates to issues in epistemology in its original context see Dougherty (2012).

of his belief set. Intuitively the proposition about Simon's experiential states is not a common sense proposition, and yet it should be by CoWCS.

It's not clear that moving toward prima facie justification will save any of the above epistemic proposals, as cases like Wendy the wine connoisseur will apply just as well to any weaker justification-centered account(Wendy will also have prima facie justification for her particularly specific introspective beliefs). Further, if a proposition's being common sense is just a proposition's being prima facie justified, then common sense is hardly the epistemically interesting concept that many have taken it to be. Perhaps one should prefer a psychological understanding of common sense.

2.2.2 Psychological Notions of Common Sense

Another family of notions of common sense that might be interesting are psychological analyses of common sense. These will attempt to find a middle way between social (to come) and epistemic notions of common sense by making common sense explicitly internal while not defining it in explicit epistemic terms. An advantage that such views have is that it is straightforward how we might extract epistemic import from the proposed analyses of common sense. Further, common sense remains the kind of thing we're in a position to recognize prior to solving external world skeptical problems.

Seemings Common Sense (SCS)   p is a common sense proposition for subject S if and only if p seems to S to be true.

Maximal Seemings Common Sense (MSCS)   p is a common sense proposition for subject S if and only if p maximally seems to S to be true.

Belief Inclination Common Sense (BICS)   p is a common sense proposition for subject S if and only if S is inclined to believe p.

The seemings proposals are very intuitively plausible. Oftentimes (perhaps all of the time), those propositions which are flagged by people as commonsensical are propositions that also seem to them to be true. Perhaps this feature is an essential feature (or at least a feature that will get us an intuitively plausible extension) of the

31

concept of common sense. However, there are problems. First of all it seems that refined senses cases like Wendy the wine connoisseur above should provide a perfectly good counterexample to these views. But supposing this case is unintuitive, there are a number of cases of seemings and belief inclinations that no one wants to say are common sense views. For example, in the Müller-Lyer illusion it *seems* to us that the lines are of differing lengths despite the fact that one would be hard pressed to say that it's commonsensical that the lines are of differing lengths. Further, it seems that there are a number of scientists who buck common sense (e.g., Galileo) who certainly don't have a seeming or even an inclination to believe the common sense proposition. That is, it certainly didn't *seem* to Galileo that the planets were special heavenly bodies, but it was the commonsensical view that he was bucking by publicizing his discoveries regarding Venus having phases and Jupiter having moons.

Suppose a very different sort of psychological account is proposed:

Psychological Unavoidability Common Sense (PUCS)   p is a common sense proposition for subject S if and only if it is psychologically unavoidable for S to believe p or to behave as if she believed p.

This is a somewhat plausible suggestion. But an analysis seems clearly false that suggests one cannot ever give up any common sense belief without it ceasing to be commons sense. However, if a psychologically unavoidable belief is rather something that even the skeptic *behaves* as if she believes, then perhaps this view merits more consideration. The skeptic still steps out of the way of the oncoming bus, flees from the wolf, eats breakfast, etc. Whether this is a case of actual belief or simply a case of acting as if one retained a belief is beside the point.[7] All this view needs is that the case is one of certain beliefs being "psychologically unavoidable."

But how unavoidable is psychologically unavoidable? Are psychologically unavoidable beliefs those that are, at some point in time, causally psychologically unavoidable

---

[7]Another possibility is that the psychologically unavoidable belief is an A-lief in accord with Gendler's distinction in Gendler (2008).

or must they be absolutely so (whether this be causally, metaphysically, or logically)? So, for example, at the time when the skeptic encounters the bus, it is unavoidable that she (act as if she) believe that there is a bus there. But as soon as it's safely out of her sight, this belief does not retain its psychological unavoidability. Should there be a time-relativizing portion of PUCS, or should the unavoidability requirement be absolute?

Either strategy one takes in developing this view, there is a problem. This proposal has the problem of being too strong, at least if it is relativized as described above. For, it seems that there are certain beliefs that one might be incapable of avoiding but that do not strike anyone as commonsensical. Consider Jeffrey, who was raised in a Broccoli worshiping cult. He is a recovering Broccolite and is no longer a member of the Broccolite community. He has come to be shown that there's nothing divine about broccoli and that it's merely a vegetable like any other. However, despite consciously being convinced that there's no good reason to believe that broccoli is in any way more than just another vegetable, he finds it unavoidable to see the broccoli that he encounters as a special sort of thing that is, in some way, divine. His position with respect to the broccoli is exactly analogous to the skeptic's position with respect to external world beliefs. Now, as I take this view to be proposing a strong notion of common sense, like Huemer's below, it seems that one holding this view would not want it to turn out that this belief is a common sense one for Jeffrey. Further, anything stronger than causally necessary at a time will also be intolerably strong. Supposing that the real world hypothesis is commonsensical, it being *absolutely* necessary (whether causally, metaphysically, or logically) will make it so that the skeptic and the antirealist *always* believe the real world hypothesis, even in those quiet reflective moments when they are free from oncoming buses.

2.2.3 Reid's Quasi-Psychological View of Common Sense

The final psychological notion of common sense to be considered will be inspired by Thomas Reid. Reid is probably one of the most difficult philosophers to interpret on the nature of common sense.[8] There are several passages which seem to suggest notions similar to the other proposals in this chapter, which I will not mention here. However, in his *Inquiry* Thomas Reid briefly discusses what he calls the "foundation" of all common sense. Perhaps an interesting and informative analysis of common sense can be gleaned from this passage.

> A third class of natural signs comprehends those which, though we never before had any notion or conception of the things signified, do suggest it, or conjure it up, as it were, by a natural kind of magic, and at once give us a conception, and create a belief of it. I shewed formerly, that our sensations suggest to us a sentient being or mind to which they belong: a being which hath a permanent existence, although the sensations are transient and of short duration. The conception of a mind is neither an idea of sensation nor of reflection; for it is neither like any of our sensations, nor like any thing we are conscious of. The first conception of it, as well as the belief of it, and of the common relation it bears to all that we are conscious of, or remember, is suggested to every thinking being, we do not know how. . . . [This class of natural signs] is the foundation of common sense; a part of human nature which hath never been explained. (Reid, 1997, 60-61)

Reid applies this immediately to the case of the conclusion that bodies are hard:

> I think it is evident, that we cannot, by reasoning from our sensations, collect the existence of bodies at all, far less any of their qualities. This hath been proved by unanswerable arguments by the Bishop of Cloyne, and by the author of the Treatise of Human Nature. It appears as evident, that external existences cannot be produced by habit, experience, education,

---

[8]Wolterstorff and Pippin (Wolterstorff 2001) argue that there are two main notions of common sense in Reid and that these are in tension with each other. The view here defended would be a third view of common sense in Thomas Reid if right.

or any principle of human nature that hath been admitted by philosophers. At the same time, it is a fact, that such sensations are invariably connected with the conception and belief of external existences. Hence, by all rules of just reasoning, we must conclude, that this connection is the effect of our constitution, and ought to be considered as an original principle of human nature, till we find some more general principle into which it may be resolved.[9](Reid, 1997, 61)

One way to read the passage above would be to construe Reid as offering something like one of the seemings-centered understandings of common sense. If this is the case then, though perhaps historically interesting, Reid hasn't given a notion of common sense that ought to be given separate treatment.[10] However, there is another quite interesting way to read the above. Reid may be intentionally distancing himself from the internalist insistence on explicitly arguing from sense data (loosely speaking) to external world propositions (put forward by the likes of Berkeley and Hume who then argue, Reid thinks devastatingly, that such argument is not forthcoming), and proposing that perhaps this shouldn't worry anyone that much. On this reading of the above, the notion of common sense will turn out to pick out those beliefs that are the result of a particular sort of spontaneous belief-forming process. These beliefs, it seems, are not spontaneous in the sense that they are entirely unconditioned, but rather in that they are the spontaneous (and, at least prior to reflection, unavoidable) result of having certain sensory stimuli.[11] On this reading of the above passage, Reid may be hinting at one of the following analyses of common sense:

Reid's Psychological Common Sense (ReidPCS):   a belief p is common sense if and only if it is formed as a result of immediate sensory stimuli by a natural, spontaneous belief-forming process.

---

[9]The excessive comma use is present in the original.

[10]Our reading, if not reasonable reading for anachronism reasons, is at least a reasonable way of developing Reid's decidedly unclear discussion here into a full-fledged view. But this ends up endorsing one of the previously dismissed psychological views.

[11]Interestingly enough, this notion of common sense seems to be eerily similar to Lyons' externalist conception of foundational justification. (Lyons 2009)

Reid's Externalist Common Sense (ReidECS):   a belief p is common sense if and only if it is formed as a result of immediate sensory stimuli by a natural, spontaneous, and reliable belief-forming process.[12]

ReidECS seems to be the natural development of **ReidPCS**, and Reid's positive discussion of these beliefs (as well as his unwillingness to give up common sense beliefs when they aren't borne out by philosophy) seems to indicate that this modification would be a quite natural one. But what are the prospects of **ReidECS** as a candidate for the analysis of common sense?

Certainly cases like Craig the creationist's belief that the Earth is roughly 7,000 years old, if compelling, will force us to abandon Reid's notion of common sense, but let's not hang our hopes too much on such an admittedly fringe case. Consider the belief that the Earth is flat. It seems that any view of common sense will need to capture that "the Earth is flat"[13] was at one time common sense and that it is no longer. **ReidECS** is going to have trouble making the added reliability condition work.[14] On one way of construing the belief that the earth is flat and the process by which this belief was formed, this belief is formed by the most unreliable belief forming process possible. The process of forming beliefs about the shape of the Earth by extrapolating from the apparent shape of the land in one's normal field of view is

---

[12]This turns out to mirror almost exactly Goldman's view of foundational justification, such that all that needs to be added is the externalist thesis to make this an epistemic view.(Goldman, 1979) However, adding certain epistemic theses to any of the above psychological views would make them epistemic views. But this is the very kind of view that can be epistemically interesting. Nevertheless, this second view which includes the reliability condition may in fact not fit in the analysis of common sense for Reid. T.A. Sandberg, for example, argues that Reid, like Descartes, relies on the existence and providence of God to defeat skepticism(Sandberg, 2004).

[13]Now, I am not endorsing the myth that has captured the imagination of us in the twentieth century that people before Columbus thought the Earth was flat. I am, however, endorsing the proposition that at some time in the past human beings generally considered the Earth to be flat. However, any case in which a previously "common sense" view has been overturned by scientific or philosophical investigation will be sufficient, if this particular case isn't plausible.

[14]This is in a way that is perhaps not entirely disconnected from the generality problem, which is an objection to the externalist view of justification from the fact that nearly any belief forming process can be described on a variety of different levels of generality which will then construe the process as either reliable or unreliable depending on which level of generality one construes the belief forming process.

maximally unreliable. Perhaps this belief can be formed by a reliable process if it is of some other type, but some reason would have to be given for why such a process type is more salient than the process type that has been proposed. One thus seems to be pushed toward considering the reliability condition uninformative and abandoning **ReidECS** in favor of **ReidPCS**.

But both proposals will have trouble dealing with another aspect of this case. The belief that the earth is flat doesn't seem to be formed as a result of immediate sensory stimuli in any kind of spontaneous way at all. It seems like the case would, at best, involve someone wondering about the Earth's shape and making an argument based on the apparent shape of the portion of the earth with which he is familiar. There's nothing spontaneous about a belief forming process that requires one to wonder about something and then make a geometrical argument based on other information one has in order to even come to have the belief. Beliefs this far from sense perception cannot fit the character required by Reid's proposal.[15]

2.2.4   Huemer's Social Notion of Common Sense

Huemer has his own axe to grind in this debate. Now Huemer might be the person that most explicitly endorses something like a social notion of common sense, although the story isn't quite so simple for him. He gives three criteria that distinguish a proposition as one of common sense:

Huemer's Common Sense (HuemerCS)   a belief is a common sense belief if and only if "i) it is accepted by almost everyone (except some philosophers and madmen), regardless of what culture or time period one belongs to, ii) they tend to be taken for granted in ordinary life..., and iii) if a person believes a contrary of one of these propositions, then it is a sign of insanity."[Roman numerals added](Huemer, 2001, 18)

The first thing to clear up is that this is not the bottom and most basic level for Huemer. He takes the *reason* that propositions are common sense to be that they

---

[15]Certainly, if one will grant that the belief that the earth is round is now commonsensical, one will be required to give up on Reid's proposals.

are those propositions that are the most initially plausible, with plausibility being construed epistemically (since a proposition's being common sense licenses using it in a Moorean shift of an argument for its contrary) (Huemer, 2001, 33). Now, similar problems of extension arise for this epistemic notion as have been shown to occur for the other psychological and epistemic notions.[16] It will be helpful *now* to discuss issues related to the notion spelled out in i)-iii). This is very strong and vivid as a social notion of common sense but his formulation has its problems just as some social notions to be considered do. For one, his account is tailor-made to not countenance the possibility that what is common sense can change. Consider some of the following examples where what is common sense seems to have changed.

Old CS: The Earth is motionless.
New CS: The Earth is hurtling through space.

Old CS: Color, as it is in the world, resembles our experiences.
New CS: Color, as it is in the world, does not resemble our experiences

Old CS: The Sun moves around the Earth.
New CS: The Sun does not move around the Earth.

Old CS: The stars are tiny pin-pricks.
New CS: The stars are giant balls of gas.[17]

If any of the above cases of Old CS used to be common sense and no longer are, then this view has a sizable problem. Further, this account cannot accommodate the

---

[16]I take several of the previous proposals to be reasonable ways to read the more epistemic notion. See, for example, the various phenomenal conservative notions of common sense and HJCS.

[17]Some raise doubts that the propositions labeled Old CS were ever common sense or even commonly believed. One might argue that, for example, that few ever really believed that color, as it is in the world, resembles our experiences. The argument goes, I take it, something like this: People aren't reflective enough to form these kinds of beliefs, but those individuals that did formed such beliefs as a result of significant deep thought. I'm not convinced that such beliefs require such deep thought. In fact, I think a number of this kind of belief are eminently natural. It is true that a really precise scientific or philosophical outworking of what these beliefs entail in the final analysis will require significant thought and outworking. But this doesn't besmirch the beliefs above as not being commonsensical. Further, the propositions in New CS strike me as commonsensical to me right now.

intuition that different things can be commonsensical to different communities like Craig the creationist's belief that the Earth is around 7,000 years old are commonsensical to the members of his insular fundamentalist community. For these reasons, let's move on to consider some other social notions of common sense.

2.2.5 Social Notions of Common Sense

Let's look a bit more closely at the social notions of common sense. Noah Lemos gives, in passing, an account of common sense that he takes to be roughly correct. His proposal promises to be enlightening. Consider DWHeldCS:

Deeply and Widely Held Common Sense (DWHeldCS) p is a common sense proposition if and only if p is deeply and widely held. (Lemos 2004, 4)

This view is given in Lemos (2004, 4) and introduces an interesting notion of common sense as defined socially rather than by a specifically epistemic state in which one finds oneself. There are a number of problems with this view if one is to take it as anything more than a rough first pass (the author seems to take this more modest view) at a definition. The view will clearly need refinement for our purposes because Lemos seems to have intended rather to give an understanding of what it is for a proposition to be common sense full stop rather than common sense for an individual. Further, for Lemos purposes, this proposal was enough of a proposal to allow him to engage in *his* particular project. One needs some way to relativize it (or at least relate it) to an individual. What do I care about whether a proposition is deeply and widely held? Well, perhaps I believe it.

DWHeldCS′ p is a common sense proposition for subject S if and only if p is deeply and widely held and S believes that p.

This notion has the unhappy consequence that it's logically impossible to deny a proposition that is common sense to you without immediately making it non-commonsensical. But the possibility of disbelieving seems to be a central desideratum for a notion of common sense. One wants, e.g., scientists to be able to buck com-

mon sense. But consider the following example which should help move the reader's intuitions toward a better proposal:

Consider **Desert Island Molly**. Desert Island Molly was raised by non-conformists, but she doesn't know this about her parents. She learned from her parents that the little flashing lights that come out in July are fairies. For all she knows, this is what everyone believes and believes strongly. After long hours of investigation with a net and a mason jar, she comes to believe that the flashing lights are actually insects. Is the proposition that the flashing lights that come out in the summer are insects a common sense proposition for Molly? Certainly not. What is it missing? Maybe Molly needs to *believe* that the proposition is deeply and widely held. Consider the following iteration of **DWHeldCS**.

DWHeldCS″   p is a common sense proposition for subject S if and only if p is deeply and widely held and S believes that p is deeply and widely held.

This is better, but some further thought reveals systemic issues with the **DWHeldCS** derivatives. Let's start with a couple of cases that the above can accommodate. These cases will lead to an objection case to the various iterations of **DWHeldCS**, which will require a more substantive modification of the view. First consider the cases that this iteration of **DWHeldCS** *can* accommodate. One needs to accommodate the fact that what is common sense can change over time. One needs to accommodate, for example, that while it used to be commonsensical that the Earth was absolutely at rest, it is now commonsensical to say that the Earth revolves around the Sun. This jibes quite well with the **DWHeldCS** based views, but a slightly modified case will force a deeper revision of the social notion.

Let's say there is some small community that was lost around the time that the proposition that the Earth is motionless was deeply and widely held. Being lost, this civilization survives untouched. It seems right to say that "The Earth is motionless" remains common sense for members of this community even though the proposition is

not deeply and widely held in the wider world. It also seems that even if a member of this community happens to truly believe that the contradictory is deeply and widely held, that the contradictory is not common sense for him. This brings us to the following proposal:

Question Settled Common Sense(QueSetCS) p is a common sense proposition for subject S if and only if there is a community C for which the question whether p is taken to be settled in favor of p and S is a member of C.[18]

QueSetCS seems to be the notion to settle on regarding a social notion of common sense. It can account for (and come down on the right side with respect to) all the cases discussed above. This will make the Wendy the wine connoisseur case turn out properly, while yielding the right answer on many classic examples like 2+2=4 and that Craig the Creationist's view is common sense in his highly fundamentalist community. There also don't seem to be any cases that are clear counterexamples to this view. **QueSetCS** promises to be a powerful and intuitive account.

The question arises: Should a belief condition be built in for **QueSetCS**? The belief condition in the **DWHeldCS** variants worked to relativize this notion to the subject, but in the current proposal the notion *comes* relativized to the individual. However, one might expect something to be different internally when a proposition is common sense for someone and when a proposition is not common sense to her. This proposal can accommodate this. For example, it seems membership in any community involves various complex intentional states. One doesn't need to commit on exactly what this amounts to, but it seems at least initially reasonable that there's

---

[18]I talk of communities taking the question to be settled on some proposition, but I don't want to commit too strongly on exactly what this is. It should be taken to be something like taking it to be widely known or widely thought to be highly justified or something in that region, to such an extent that one needn't wonder whether the proposition is true any more. Something around that area would be acceptable, but leaving the notion undefined is better, as I think there's an intuitive sense we have of what it is for a question to be settled that it's probably difficult to fully capture with a more tortuous locution. Now, alternatively, an interesting (and nearly identical) proposal would define CS in terms of what a community takes to be central to a shared web of belief. This other view will probably have a different extension, but I'm not convinced that there are cases that should make us prefer this more restrictive view (i.e., it's harder for a belief to be central to a web than for the question to be settled in its favor) to the broader QueSetCS.

some intentional state requirement for community membership. This may suffice in terms of a requirement for a difference in mental states. However, there is no need to be dogmatic. There's room also (though a convincing case would need to be made) to insist on adding a requirement of belief that the proposition is common sense, as long as the account doesn't run afoul of allowing disbelief in common sense propositions.

Further, we may not want to hang our hats on this insistence on a difference in internal states for community membership. One might be attracted to externalist theories of mental states such that a difference in mental states will not entail an internal difference. There's also reason to think that, for at least some communities, membership does not require any difference in mental states on the part of the member. Consider, for example, the practice of infant baptism in Catholicism and of circumcision in Judaism. These are both, among other things, ceremonies meant to indicate that these very young children are members of an in group (the Catholic community—for baptism—and the Jewish community—for circumcision). That is, these are ceremonies that are thought to induct these young children (too young to know the difference) into certain specific communities. But on this more external view of community membership, this view (and perhaps others) will need a theory to account for the first personal use of a propositions apparent common sense nature. That is, this view will need to say something about the internal such that we can make sense of human beings and their tendency to appeal to and have views about what is common sense.

2.3   The Agreement Heuristic

Human beings have a heuristic for deciding whether or not a proposition is common sense. I think, often, human beings judge something to be common sense when they have the intuition that the fellow members of that community would agree with the individual's assessment of that proposition. For example, the creationist judges

that the earth's being approximately 7,000 years old is common sense by, in one way or another, considering whether the other members of his community would take the question to be settled in favor of a young Earth. Sometimes coming to this belief involves judgments of how natural the belief seems, and sometimes this involves some sort of imagination or simulation of others' beliefs(and usually both). This heuristic (though usually good enough) can go wrong when there is a disconnect between what others believe and what one takes them to believe or be disposed to believe. Remember that **QueSetCS**, the notion that we've decided upon, doesn't require that any single individual believe something except insofar as a community taking a question to be settled on some proposition requires some set of shared beliefs, but the idea behind why this heuristic often works is that a strong intuition that one's community mates would agree or find the belief natural will often track the belief being commonsense by **QueSetCS**.

There are, however, times when the heuristic fails to track the property, in which case the heuristic provides a handy error theory (even for those who like the internalized versions of **QueSetCS**), which will come to our aid in addressing objections. For example, consider McCain's example of a belief that is common sense just for me: "I am looking at a computer screen." (McCain, 2014, 153) It is not clear that **QueSetCS** will call this belief common sense in all cases. The cases where **QueSetCS** will call this belief common sense are probably the outlier cases. Most cases of my looking at a computer screen don't go along with widespread agreement about my relation to the screen in some community of which I'm a member. Perhaps it is common sense relative to the small community of people in the coffee shop, but probably not. So then why are we inclined to call such deeply first-personal beliefs "common sense"? It is because of our heuristic. We judge that it's commonsensical by thinking that it would be natural to believe, were anyone to bother to come to such a belief in an epistemic situation that is roughly similar to mine, that I am

looking at a computer screen. It just turns out that our heuristic fails to identify common sense all of the time. First-personal propositions that are intuitively not common sense, such as that I'm having an experience as of phenomenal white or the wine connoisseur case, may trip the heuristic. It certainly does not seem as natural a belief or as easy to settle whether I'm having a certain phenomenal color experience or whether my wine has notes of green apple. The heuristic we use to judge whether a proposition is common sense provides a nice neat error theory for the potentially problematic cases of first-personal reports that we take to be common sense. It may also be helpful addressing the objections below.

2.4 QueSetCS and Strong Propositional Common Sense

It will be important, as well, for QueSetCS to be able to account for the intuition of individuals like Huemer, Reid, and Moore that there's a special kind of propositional common sense that is something stronger than merely that there is some community for which the question is settled in favor of the proposition. There is a need to countenance that there is *strong* common sense.

There's a promising proposal that one can understand strong common sense as a special case of the weak notion as understood in **QueSetCS**. On this view it needs to be asserted that there's a community that one might call the community of human beings. This is the historical community of all human beings, based on perceptual and biological similarities. Those propositions that are common sense to a person by virtue of her membership in the historical community of human beings will be strongly commonsensical. So a proposition is **strong** common sense if and only if it is (a) common sense by **QueSetCS** and (b) community C is the historical community of human beings (a slightly weaker notion would have C be the community of currently existing human beings). This allows **QueSetCS**, the view that can more widely capture our intuitions about common sense, to be the main notion of common

sense while accommodating the strong common sense theorist's intuitions.

2.5   Objections

2.5.1   Obj. 1: No Scientific Research

Some will object to the very idea of there being a weak notion of common sense, and especially to cases where the community in question is scientific. One might think there is something about fields that involve deep and difficult study and research which prevents propositions in these fields from being common sense, but **QueSetCS** would qualify certain advanced scientific truths as common sense. For example, that electrons that are quantum entangled have opposite spin values (where spin is a technical mathematical concept that doesn't reflect our ordinary way of talking about spin) would be common sense for the community of professional physicists since the community takes the question to be settled in favor of the opposite spin values proposition. Now the easy and dissatisfactory answer is that capturing the weak notion of common sense that is operating in such cases is a virtue of this view and is in fact a necessity for any satisfactory view of the various kinds of common sense, but this is far from helpful, dialectically.

Nevertheless, there is a compelling case that involves something in exactly those fields that are theory and experimentation heavy. Here's a statement that seems to be a piece of common sense for scientists: "Something bad will happen if one places one's hand in front of a running particle accelerator." Now, the theoretical grounding of this would require not only huge bodies of scientific knowledge about physics but also about the effects of near light speed particles on pieces of human anatomy. This can hardly involve more of the stuff that is theory and experiment heavy, but it is also intuitively a case of common sense for the scientist. The need for much study doesn't leave a proposition out of the discussion of common sense.

2.5.2   Obj. 2: Trumantown

Probably the most difficult objection to the **QueSetCS** view (and consequently to the understanding that **QueSetCS** can ground strong common sense) comes from borderline cases. Consider the following.

Trumantown is the only town that has ever existed and resident Truman is the unfortunate subject of a deception. Despite what the residents tell Truman, all residents believe that there are no rocks (just painted Styrofoam lookalikes), for they've never seen a real rock. However, one day when Truman comes back from his morning walk, the town gathers to own up to this deception. They tell Truman that no one has ever actually believed in rocks and convince him of this. Has Truman's belief that there are rocks never been common sense, and has he just realized that "there are rocks" isn't commonsensical? One might have an intuition that "there are rocks" is a commonsensical proposition whatever the situation is for Truman. If it ceases to be common sense for Truman, then why does Truman still take the view yet to be common sense (for at least a time after being convinced that no one ever believed it)?

Here is a case where the agreement heuristic fights against reality. Truman is in a situation where, for at least some time after having the truth revealed to him, he will have a certain kind of cognitive dissonance. It will be difficult for Truman to shake the feeling that this other belief that there are rocks is a natural belief to have, and for that time the heuristic will say that the proposition is common sense, while in fact it is not. Truman will soon recognize this. There is actually a case very close to our own lives that more obviously shows the heuristic still firing after a deception accounts for the attraction of calling the view common sense.

For example, consider a case very similar to this radical deception. Little Timmy finally is sat down and told that there is no Santa Claus. Prior to his discovering this fact, Timmy had very good indications that the belief that "there is a Santa Claus" is more commonsensical than it actually is. His parents told him that there

46

was a Santa Claus, and he had every reason to believe that his parents believed it and perhaps even believed it justifiedly. Now, in this real life case was the Santa Claus hypothesis less commonsensical than Timmy thought it was? Yes. And should he give up his judgment of the commonsensical nature of his Santa Claus belief upon discovering the systematic deception? Of course. But there's still a psychological inertia keeping Timmy from accepting the truth that the Santa Claus belief isn't in fact as commonsensical as Timmy originally thought. It remains a natural seeming belief, but there will be a radical change as Timmy resolves his cognitive dissonance.

2.5.3   Obj. 3: Communities and Contradictory Deliverances

Third, one might think that **QueSetCS** cannot be maintained without giving a detailed account of the nature of communities or at least giving conditions for being a community. Now, in ethics there's a problem with cultural relativism that is connected to the question of what constitutes a culture and membership that will be illustrative. The problem is that it seems that any individual is a member of a number of cultures. These cultures may often disagree on questions of morality. So an advocate of cultural relativism, on this objection, must come up with a notion of culture membership that will not produce contradictories (e.g., eating pork being wrong and not wrong for some individual). Now, the view being defended is not worried about contradictory deliverances because each item of common sense is fundamentally community relative. **QueSetCS** will make it so that there's no contradiction in the contradictory deliverances. There will be no cases where it turns out that it's common sense that p and it's not the case that it's common sense that p, because of the relatively liberal definition given in **QueSetCS**. It may be common sense in virtue of membership in community such-and-such that p and also not be common sense in virtue of membership in some other community, but this isn't a contradiction.[19]

---

[19]This response follows very closely the view of morality presented in Harman (1975). The idea is that it's the point of such a relativist view that different actions and beliefs can be subject to different agreements

Further, it isn't a problem for this view that some proposition p is common sense while the proposition $\neg p$ is also common sense. For whereas, intuitively, the cultural relativist about morality still needs it to be false that both I ought to p and that I ought to $\neg p$, there is no reason to suppose this for the relativist view of common sense just proposed. In fact, it seems to be a virtue of the view that **QueSetCS** can take account of the confusing and sometimes contradictory deliverances of common sense (a young earth geologist—and there are such people—seems to be a classic example of this kind of internal battle). This even makes good sense of the struggles over the nature and use of common sense over the years.

2.5.4 Obj. 4: Relativity and Communities of One

There's a different way of pushing the relativity objection that threatens **QueSetCS** in the direction of a radical subjectivism, pushing it in a psychological direction. The question is whether there is a limiting case of a community of one (if there is then maybe we should understand common sense more subjectively). Consider Martha:

Martha is in a community where everyone believes that the big thing above them is a sky with a sun in it. Suppose, further, that Martha and her community are subjects of a systematic deception. They are actually in a dome carefully climate controlled and light controlled, etc., in order to look and feel like Marthatown was underneath a sky with a sun. Now, everyone in Marthatown is convinced by this deception. But then Martha goes for a hike up the highest mountain in Marthatown and at the peak she reaches up and is able to touch the dome that is over them.

At this moment has Martha become a community of one? If so, perhaps **QueSetCS** should collapse into something more strictly subjective. If there is a community of one in this case, then it seems that community isn't the driving notion behind anything being common sense. There is no community with which she can

confer and settle the question but, *ex hypothesi*, it ought to be that the proposition that there is no sky or sun is common sense for Martha. However, it simply seems more plausible to me that Martha here just has a belief that bucks common sense. If there were more than one person on this particular odyssey of self-discovery then we would very quickly be able to call this a community. But in the current case, saying her belief is not common sense is at least as plausible as there being a limiting case of a community of one, and avoids the radical shift that the objection in question threatens. Further, the agreement heuristic can account for any lingering plausibility it might have.

But there are better cases that push us toward respecting communities of one. One might propose **Desert Island Dave**. Desert Island Dave is an individual we've placed on a desert island who has beliefs and attitudes like an ordinary western human being, but he's put on this island to purposely isolate him from the rest of humanity. It seems like the standard sorts of things that are common sense for him despite it seeming like we've taken him out of any community in which he might be a member.

It should be obvious, though, that this isn't quite enough to sever all of the ties Dave has to communities. No matter how we construct the Dave case, we will be forced to have some story to tell about how Dave came to have the current beliefs and abilities he has that will make him a member of various communities. In order produce a strong intuition regarding communities of one, more extreme cases must be considered.

Consider **Deserted Island Swamp Sally**. If we take a clue from the philosophical swampman cases, perhaps more progress may be made here. So suppose that in a swamp on a deserted island lightning strikes, creating a molecule-for-molecule mental-state-for-mental-state copy of Sally, a normal western person, on the desert island. Arguably Swamp Sally isn't a member of any of the communities Sally was in, but it seems plausible that very similar things (or at least *some* things) will be com-

mon sense for Swamp Sally that are common sense for Sally. But Swamp Sally isn't a member of any communities, except perhaps a community of one. Thus **QueSetCS** will need to respect communities of one.

Before we come down on this issue, it will be important to decide whether the swamp Pope is Catholic. Clearly he would not be the Pope, as that office is closely tied to the individual's causal history (i.e., having been elected at conclave, etc.), but one might have the intuition that he is yet Catholic (surely the swamp Pope would be Catholic if any swamp person is). If the intuition is yes, then our Swamp Sally may very well still be a member of a number of communities and thus we yet will not be pushed into requiring communities of one. If the intuition is no, then more will need to be said.[20]

If the swamp Pope isn't Catholic, then Swamp Sally's creation has succeeded in creating a fully functioning individual who is not a member of any communities. Well, this isn't exactly right. We've created an individual who is only a member of one community, the community of human beings, but not a member of the other communities that regular Sally is in. The question to consider, then, is whether there is a compelling case of common sense for Swamp Sally. Consider Swamp Sally's belief that Settlers of Catan is a fun game. This is common sense for regular Sally, who is a member of a local gaming club in which all agree strongly on this, but Swamp Sally has no such club. On our view, the proposition about the game Settlers of Catan is not common sense for Swamp Sally. But the heuristic won't completely explain Swamp Sally's going wrong on this one, since Swamp Sally also has every reason to believe she's a member of the gaming club and that this proposition is taken as settled in favor of Settlers of Catan being a great game by her club. Swamp Sally is subject to a systematic deception that will push her to wrongly take this proposition

---

[20]I think, probably, the truth is somewhere in the middle. There are communities membership in which requires a particular causal history and there are some communities that center around certain kinds of shared beliefs and desires.

as a common sense one, but it isn't.

2.6   The Problem for the Antiskeptic

It will be important to briefly discuss an antiskeptical argument Moore makes in "A Defence of Common Sense," as it will raise problems for the antiskeptic. In the article Moore seems to have some sort of social notion of common sense in mind, and utilizes the concept in an interesting way to attempt to deal with a certain kind of skepticism (Moore 1993: 116). He takes the confident assertion that propositions about an external world are "common sense" to be inconsistent with holding skeptical beliefs because such confident assertion betrays that one takes oneself to know that there are other human beings that are in the world and have beliefs. On this notion of common sense at least a certain form of skepticism might be inconsistent. Namely, one ought not assert a skeptical position while confidently (as if you knew) asserting that the proposition(s) with respect to which one is skeptical are common sense. However, if Moore's right about this incoherence, there might be more troubling skeptical consequences.

Moore's argument seems to go something like this. The skeptic often admits that external world propositions, etc., are common sense propositions but also asserts that he isn't justified in believing external world propositions. However, since, among other things, a proposition's being common sense entails that there are other people out there in the world that believe it, the confident assertion that external world realism is commonsensical constitutes an assertion that one is justified in believing at least one external world proposition (namely that there are other people out there that believe external world propositions). Now, suppose that Moore is right in his accusation of inconsistency with respect to the skeptic. This will raise our quasi-Meno problem again in spades, supporting premise **(2)** from earlier.

Showing that some proposition is common sense will require that one is already

justified in believing external world propositions. Furthermore, if one wishes to moti-vate a common sense realism *without* appeal to the common sense nature of external world propositions, there are going to be problems in fixing a target for one's ar-gument. If there's no common sense as a basis on which to fix the target of one's antiskeptical arguments, why should we not limit the antiskeptical project to sim-ply getting what we can get? Why should one fix on any one set of propositions to defend? This is a problem for certain antiskeptics who wish to ground the nature of antiskepticism in common sense. One could always stipulate what one will count as an antiskeptical view, but why should anyone worried about skepticism care to defend one particular set of propositions rather than another in our antiskeptical en-deavors?[21] It's also a problem that we, in effect, have simply chosen one view amongst the very many "common sense" views of the world. For example, propositions sup-porting strong color realism may be common sense, but we certainly don't want to be forced to defend them here. More work needs to be done in order to have a robust defensible account of the nature of antiskepticism.

We saw above that a full-bodied naïve view of the world may be disastrous if it turns out to be *the common sense view* of the world. This doesn't mean that this "primitive science," as Quine calls it, is *not* the common sense view of the world, but we'll return to this concern in a minute. It will be important to clarify exactly what the defenders of inference to the best explanation arguments against skepticism take

---

[21]Further, the view that's been developed will help to settle a number of basic questions concerning common sense's epistemic role as well as shedding light on the disagreements about which view in fact is *the* common sense view of the world. **QueSetCS**, given its close relation with testimony, will yield that common sense propositions are usually justified (assuming testimony justifies) and are so justified defeasibly. It will not be the case that all common sense propositions are at least *prima facie* justified unless one takes a view of common sense analogous to the non-reductionists' view of testimony. If one takes such a non-reductionist view, where being common sense fundamentally gives evidence in much the same way that testimony that p fundamentally gives evidence that p, one will be able to respond with a resounding yes to the question of whether a proposition is justified simply in virtue of being common sense. However, the question of whether Common Sense is defeasible seems to be definitively settled (if **QueSetCS** is true) in favor of common sense propositions being capable of being undermined by philosophical argument (as well as many other kinds of argument), whatever other views one takes.

to be the common sense view.

BonJour characterizes the view he defends as:

> the approximately commonsensical idea that my sensory experiences are
> systematically caused by a realm of three-dimensional objects, (1) having
> at least approximately the shape corresponding to those reflected in the
> "nuclear" sense-data, (2) through which I move in such a way as to change
> my point of view, (3) which are spatially related to each other in the ways
> reflected in the sequences produced by my apparent movement, and (4)
> which have causal properties and change over time in the ways correspond-
> ing to the relevant further aspects of the experiential patterns.(BonJour,
> 2003, 92)

Roughly speaking, the "approximately commonsensical idea" is the idea that there
is a world of three-dimensionally extended spatial objects. This is a running theme
amongst defenders of best explanation arguments against skepticism. Vogel also
explicitly grounds his naming the Real World Hypothesis in Common Sense(Vogel,
1990, 659), and in the end the view that he defends is very similar to the view
described by BonJour above, with his argument discussing spatial properties without
exceptionVogel (1990).

McCain, however, is the only to ground his view with an *explicit* discussion of
what counts as common sense:

> The set of beliefs that constitutes CS will vary from person to person. For
> example, the set of CS beliefs that I presently have contains beliefs such
> as "I am looking at a computer screen", and "My sister is shorter than I
> am", but another's set of CS beliefs that go into CS will vary from person
> to person, there will be several beliefs that are held in common. For in-
> stance, many individuals' set of CS beliefs will include the following: "The
> sun is bigger than earth", "Infants sometimes cry", and "There is food at
> the grocery store". More important, the beliefs that constitute CS in gen-
> eral have two significant features....their truth entails a mind-independent
> external world with which we causally interact.(McCain, 2014, 153,n.15)

McCain takes what seems to be a strictly subjective understanding of common sense here in vindicating calling the antiskeptical view the "common sense" view. Common sense and its nature are left largely mysterious, but the assertion is that the truth of common sense beliefs entails "a mind-independent external world with which we causally interact". For McCain, in order to defend common sense one must show the existence of a mind-independent external world with which we occasionally have causal interaction.

There's still a problem of setting the standards for what counts as an antiskeptical view. For example, depending on how we understand "mind-independent", a theistic world, though full of three-dimensional spatial objects, would fail to be antiskeptical. Now, on a different reading of "mind-independent" in which a world is mind-independent if it's independent of *my* mind, settling for brain-in-vat scenarios and even demon scenarios would constitute succeeding in defending this *particular* antiskeptical thesis. However, in practice, McCain seems to be just as dissatisfied with this kind of modest conclusion as Vogel and BonJour seem to be. He, along with Hasan (2015), seeks to defend a world roughly the same as that which Vogel and BonJour defend, one of three-dimensional spatial objects.

This is the target that seems to be set. But let's suppose that these arguments fail to successfully defend the existence of three-dimensional spatial objects in the way desired. Is one doomed to skepticism? Suppose we decide to approach the three-dimensional external world in the same way that one normally approaches intuitive secondary qualities. On this view there exist three-dimensional spatial objects and, depending on how we understand "how we take them to be" they can reasonably be argued to be roughly as we take them. If being spherical means, for example, that I have some systematic spherical experiences in thus-and-such circumstances, then we can, with Berkeley, claim that it *is* spherical, despite, say, the cause of the experience being some idea in the mind of a demon or a god.

This can't count. Skepticism would be too easily defeated, and the result would be deeply unsatisfactory for most antiskeptics. But why? One might ground the rejection of such views' dissatisfactory nature in the problem with cashing out the phenomenal meaning of terms like "spherical" in a defensible way, but I'm not sure that this gets to the heart of why people are dissatisfied with a phenomenalist way of rescuing propositions that are 'antiskeptical.' Let's suppose such a phenomenalist account of a term like "spherical" were more defensible. It still strikes me (and I think it would strike most) as deeply conflicting with antiskepticism to take a phenomenalist reading of the world.

In fact, the deep reaction I have to this case may in fact be the common sense heuristic going off. I judge that other human beings would largely agree with me (or that it's natural to believe) that phenomenalism is wrong, and rather something closer to a primary quality analysis is the right view of sphericality. That is, perhaps the agreement heuristic will suggest one not settle for such weaker views. This would allow common sense to do some work in grounding the dissatisfaction with certain proposals for what counts as antiskepticism.

If this is the case, the problem becomes the fact that, arguably, common sense in this sense, really, is on the side of a primary quality analysis of even such properties as color. On this view, then, it seems that common sense is doomed. An opponent might first attempt to argue that this isn't really the case, but I'm not convinced. For me at least, a primary quality analysis of color seems very natural. It took me until I read Berkeley to be at all reconciled with the idea of understanding color in a reductive way, and I think the philosophically unsophisticated subject (at least those in less scientifically advanced cultures) will tend toward the more natural view that the colors I experience are actually out in the objects that cause those experiences.

Maybe this is because human beings have a tendency to confuse reductions with eliminations. So, for example, one might not like a functionalist account of mental

states, but this may be because one is confusing the assertion that mental states are *just* functional states with the claim of the eliminativist that there *aren't any* mental states.

2.7   The Less Easy Solution

There is a natural, if trivial, way to go in grounding a goal for the antiskeptical project, which may not have the added problems of the above common-sense-centric approach. One might characterize a view as antiskeptical if and only if it entails (or requires) the falsity of skeptical scenarios. But this is entirely trivial and unhelpful

What is needed immediately after characterizing antiskepticism in terms of skeptical scenarios is a reasonable characterization of skeptical scenarios. In the following I will consider proposals for the nature of a skeptical scenario taking core skeptical cases as data with which to test particular proposals. Instinctively one thinks that a skeptical scenario is just one which eliminates physical objects. But that can't be quite right. Brain-in-a-vat scenarios, arguably, don't eliminate physical objects. There are brains, vats, and computers in a brain-in-a-vat scenario. These should all be considered physical objects. It can't be simply that a skeptical scenario is just one which eliminates physical objects.

However, there might be reasonable modifications of this view which will accommodate brains-in-vats. Perhaps a skeptical scenario is a scenario which requires the elimination of the material objects of our experience. The immediate question regards what counts as requiring the elimination of the material objects of our experience. In order to answer this, it will be necessary to briefly discuss what is meant by an elimination (and comparing it to reduction) and finally to discuss what it means to require the elimination of the objects of our experience.

2.7.1 Reduction and Elimination

It will be beneficial to discuss the distinction between reduction and elimination. The rough way to put the distinction is that reduction of a thing doesn't reject the existence of something, but rather says that it's *only* something else, while an elimination simply takes the thing out of its ontology. For example, Paul Churchland is an eliminativist, asserting that there aren't minds, but rather there are only brains. On the contrary, a functionalist agrees that there are minds, but says that minds are merely certain complex causal orientations of physical objects. The functionalist is a reductionist with respect to minds.

But there are two main kinds of reduction. There is scientific and ontological reduction. Scientific reduction, on a standard view, attempts to show that the laws of one science can be expressed with the laws of another.[22] An ontological reduction claims that some purported entity or class of entities is really *just* some other entity or class of entities.

Consider how these reductions would play out in different cases. Consider a scientific reduction of, say, the laws about water to the chemical laws that govern H_2O. Scientifically, water's being reduced to H_2O involves the entity water and the laws governing it and H_2O and the laws governing those atoms being connected by bridge laws. That is, there is an organized whole which is called water which behaves in certain ways, and we in chemistry have Hydrogen and Oxygen which behave in certain ways. A scientific reduction takes the behavioral laws of water and comes up with laws that biconditionally connect the entities and laws of the higher science with the entities and laws of the science which posits Hydrogen and Oxygen only.

Hempel gives a more technical example:

> One very simple form that the derivation of a biological law from a
> physico-chemical one might take can be schematically described as follows:

---

[22]See, for example, Nagel (1961).

Let 'P_1', 'P_2' be expressions containing only physico-chemical terms, and let 'B_1','B_2' be expressions containing one or more specifically biological terms (and possibly physico-chemical ones as well). Let the statement 'all cases of P_1 are cases of P_2' be a physico-chemical law—we will call it L_p—and let the following connecting laws be given: 'All cases of B_1 are cases of P_1' and 'All cases of P_2 are cases of B_2' (the first states that physico-chemical conditions of kind P_1 are necessary for the occurrence of the biological state or condition B_1; the second, that physico-chemical chemical conditions P_2 are sufficient for biological feature B_2). Then, as is readily seen, a purely biological law can be logically deduced from namely, 'all cases of B_1 are cases of B_2' (or: 'Whenever the biological features B_1 occur then so do the biological features B_2'). (Hempel, 1966, 105)

A scientific reduction (on the classical view defended by Hempel (1966) and Nagel (1961)) involves deriving a science from another, more basic science.

Now an *ontological* reduction is best understood in reference to the debate over dualism. An ontological reduction doesn't necessarily claim that some science can be done away with in favor of some other more basic science. An ontological reduction claims that some kind isn't a basic kind of entity, but rather is merely composed of some other more basic kind of entity. In the philosophy of mind, non-dualists are usually ontological reductionists. For example, the view called functionalism takes a mental state to be a complex functional organization of physical entities(usually the functional organization matters more than the entities here). Thus minds and mental states (for the functionalist) are not basic kinds of entity. A more basic kind of entity (usually physical stuff) is all there is and minds are somehow made up of that. An ontological reductionist will not need to question the autonomy of a science regarding the mind. The ontological reductionist will simply state that some kind of entity or posit really is some other kind of entity (or organization of entities). Normal statements regarding the existence of the reduced entity remain true, nevertheless.

Now, for semantic externalists, the kind of reduction engaged in here immediately creates what I'll call an *intentional* reduction. An intentional reduction is what I call

the sort of thing that happens, for the externalist, when we discover that water is just H_2O. When we discover that water is H_2O, for the externalist we *discover* that what we had meant all along by 'water' is just 'H_2O'.(Putnam, 1973) This is what I will call an intentional reduction. Arguably, as will become more important later, Paul Churchland rejects ontological reduction of the mind in favor of elimination because of the bleak outlook of an intentional reduction. Churchland calls the mental language connected with folk psychology "confused."

> ...The eliminative materialist is also pessimistic about the prospects for reduction, but his reason is that folk psychology is a radically inadequate account of our internal activities, too confused and too defective to win survival through inter-theoretic reduction. On his view it will simply be displaced by a better theory of those activities.(Churchland, 1981, 73)

Now, for semantic internalists, ontological and intentional reduction come apart more readily. For example, an internalist will grant that even if what minds *are* is a functional organization, it is left open that "mind" yet means something quite different, but it happens to always *refer* to the functional organization. A reductionist about Superman might take Superman to really just be Clark Kent, and thus that "Superman" refers to Clark Kent, but may also maintain that "Superman" means something very different from "Clark Kent." For the internalist, intention reduction[23] doesn't necessarily go with ontological reduction.

Now, on the semantic internalist view at least, we can understand the reason Churchland may be rather more happy with elimination than with reduction. A reasonable understanding of eliminativist reasoning has the eliminativist rejecting an intention reduction as unreasonable, but then finds that the unreduced intentional content *conflicts* somehow with the ontological reduction. For example, Superman has superpowers, but, one might reasonably think, part of what I mean when I say "Clark

---

[23]This isn't strictly a reduction in the technical sense, but merely the intentional (or extensional) change required as a result of an ontological or scientific reduction

Kent" is "that ordinary non-superpowered human being." In this case one would reasonably reject the reduction of Superman to Clark Kent. When such apparent conflict exists, one will be pushed toward elimination, since the reduction can't reduce *this* thing that *I'm* talking about to something else. In this kind of situation one should rather call one's view an elimination. In the Superman case, of course, there is an obvious solution. It is not, in fact, Superman that is reducible to Clark Kent. Clark Kent is to be eliminated. There is no Clark Kent. There is just Superman in disguise (one might also call this a reduction of Clark Kent to Superman, and one can resolve the apparent intentional conflict, but this would take us a little too far afield).

But it will be important, briefly, to consider an alternative way of understanding ontological reduction and its relation to language. On this view the relationship between ontological reduction and translation is much closer. Carnap characterizes reduction in the following way:

> We say that a proposition or propositional function is "exclusively about objects $a, b, ...$" if, in its written expression, there appear as extralogical symbols only "a", "b",...; logical constants... and general variables may also occur. IF for each propositional function which is exclusively about objects a, b, c (where $b, c, ...$ may be absent) there exists a coextensive propositional function exclusively about $b, c, ...$ then a is said to be reducible to $b, c, ...$ Thus we can say more briefly but with less precision that *an object is said to be "reducible" to others, if all statements about it can be translated into statements which speak only about those other objects.* [emphasis in original](Carnap, 1967, 60)

This view which explicitly limits talk of reduction to talk of translation has the advantage of squaring with an intuitive aspect of the view discussed above. That is it squares with the view that issues in translation are evidence of a failed reduction (such as encountering contradictions in attempting to capture the truths about, say, Superman, by only appealing to facts about Clark Kent). Further, if it's correct

that philosophical metaphysics is fundamentally meaning analysis, then the two views aren't that different anyway. But even if we disagree with Carnap's view of reduction, the failure of a Carnapian reduction will help us test whether an ontological reduction has succeeded. We can look for contradictions and use our intuitions about whether certain views of the world or scenarios can capture the truths about material objects.

2.7.2 Characterizing Antiskepticism

A primary method of deciding whether an elimination has occurred is if there is intentional conflict between our concepts and the proposed reduction entities (or, in the informal version of Carnap's view, whether translation is possible). But this means that, strictly speaking, one needs to develop a way to decide whether a given scenario's posits to take the role of the material objects of our experience conflict with our concept of the "material objects of our experience." One way to do this, supposing we have a better understanding of how to use our concepts than of their technical definition, is to use particular cases which most obviously eliminate the material objects of our experience. That is, in order to get a better idea of what conflicts with our concepts of the material objects of experience, we should consider obvious cases in which they are eliminated. Perhaps by doing so we can characterize the elimination of the material objects of our experience in a way that will help us succeed in our task of accurately setting out a goal of antiskeptical reasoning.

It will need to turn out, first of all, that standard God, demon, and computer skeptical scenarios turn out to eliminate the material objects of our experiences. God and demon scenarios, intuitively, will give rise to conflict in a different way than computer scenarios. The conflict in standard God and demon scenarios is, first of all, that, if we were to reduce the material objects of our experience to something, material objects would be ideas in a mind. But intuitively, it seems to conflict with our conception of the material objects of our experience that such objects are only ideas.

Further, it's not enough that one add to a demon scenario objects with extension and shape that *have nothing to do with* our experience to rescue the "material objects of our experience" from elimination. These objects need to be the objects *of our experience.*

However, the computer skeptical scenario posits objects with extension and shape that are crucially part of the cause of our experiences. Something else must be the source of the conflict in the case of a computer scenario. There's something wrong with the objects which have extension and shape in the computer skeptical scenario. those objects don't necessarily have the same extension and shape as the objects of our experience. For example, there's nothing preventing a computer causing an experience of a triangle that appears three feet from a circle while storing the computer data encoding the triangle in a roughly circular location on a disk or storing the triangle data mere inches from the circle without affecting our experience. Perhaps the upshot of this case is that a world eliminates the material objects of our experience providing that the shape and extension of the physical objects in the world aren't necessarily the same as the shape and extension presented in veridical experience. Perhaps the "material objects of our experience" don't exist unless there exist three dimensional objects which have roughly the same shapes, extensions, and spatial relations as they are presented as having in experience.

There are certain cases which will be borderline, though. For example, suppose there is a material world that is roughly as we take it to be, but our experiences are caused by an incompetent demon who accidentally causes my experiences to exactly match the material world. It's a difficult question whether the particular conflict we're looking for exists here. Supposing it's merely by chance that the demon causes us to have the very particular experiences which mirror the nature of the world, one might take an elimination to have happened because the objects don't have the right kind of causal relationship with our experience. This would be most plausible if some

version of a causal theory of objects is true.[24] But it also seems reasonable, to some extent, to call a case of an incompetent demon a case where the material objects of our experience are not eliminated. We might just look to causation because that is our ordinary way of getting *justification* for the belief that the material objects of our experience exist.

But the way forward when being given borderline cases will be to more clearly get a picture of the meaning of our concept of the material objects of our experience. But this will mean that puttering around with examples without engaging in an analysis of the notion of a material object and of the nature of perceptual experience will not get us all the way to the view we need. We seem to have, at least roughly, gotten at an important *part* of what is required to be antiskeptical. An antiskeptical view will need to posit material objects that have roughly the shapes, extensions, and spatial relations presented in our experience. This, as the several statements from Inference to the best explanation (IBE) defenders at the beginning of this chapter make clear, will be satisfactory enough to give *a* goal to the defenders of IBE against skepticism. The IBE's against skepticism will have to, at least, establish the existence of objects which have roughly the same shapes, extensions, and spatial relations as we are presented with in experience. However, this won't be enough. In order to give a fully satisfactory response to skepticism, the defender of IBE's against skepticism will need to give a plausible analysis of physical objects. Only then will the antiskeptic be able to decide what *finally* counts as rejecting skepticism. It is only then that a fully antiskeptical project can be realized.

---

[24]This is the view where material objects and their properties are defined in terms of their causal relation to our perceptual states.

Chapter 3

The Nature of Explanation

3.1   Introduction

We've settled uncomfortably into a set of minimal requirements for the IBE arguments which antiskeptics rely on. Before we consider the merits of the arguments to the best explanation against skepticism, we should come to a view of the nature of explanation and of inference to the best explanation and consider the epistemic consequences of each. This chapter will attempt to settle on a family of views of explanation on which inferences to the best explanation can be based. We will start by examining several candidates for the nature of explanation. Ultimately we will conclude, tentatively, that explanation involves giving causes (at least insofar as IBE's against external world skepticism are concerned).

3.2   Hempel's DN and IS models

For Hempel, explanation has the following fundamental properties:

(1) law: Explanations must appeal to at least one law in their *explanans*.
(2) entailment/making probable: On the DN model, the *explanans* must entail the explanandum. On the IS model, the *explanans* must render the explanandum probable.

(3) empirical relevance: Any *explanans* must be empirically testable, at least insofar as the combinations of laws and conditions predicts at least one piece of empirical data, namely, the *explanandum*.

(4) truth: The *explanans* must be true. For example, phlogiston can't enter into a proper explanation of heat transfer because there is no such thing as phlogiston.(Hempel, 1965)

3.2.1   Problems of Asymmetric Explanation

There are lawful regularities that have an explanatory asymmetry to which Hempel's view is insensitive. There are cases with lawful regularities that go both ways but in which the lawful regularity is only explanatory in one of the directions. Consider the following case: I wake up in the morning and see a thick layer of frost on my window. Now frost being on my window is highly correlated with there being frost on my neighbor's window. It would be ludicrous to say that the frost on my window *explains* the frost on my neighbor's window or vice versa. Instead it is the very low temperature outside that explains both my and my neighbor's windows' frost. Another case which is rather famous is the case of a shadow. We can predict the length of a light pole's shadow by knowing its height and the angle of the sun *and* we can predict the light pole's height by just knowing the length of its shadow and the angle of the sun. However, we want to say that the height of the pole explains the length of the shadow and *not* that the length of its shadow explains the light pole's height. These cases seem to show that the unmodified Hempelian models cannot be sufficient conditions for something's being an explanation.

However, there is a phenomenon called screening off which may be helpful as a modification of this view. An event C screens off some event A from some other event B under the following conditions: P(A/BC)=P(A/C).(Wronski, 2010, p. 12) Cases like the frost case will stop being a problem if we build in a not being screened off condition. My window's having frost on it is positively correlated with my neighbor's window having frost on it, but if I actually take into account what will end up being

(intuitively) the *real* explanation, namely some fact about the weather, then the correlated factor that my neighbor's window has frost makes no probability difference in the current case. So in order to rescue a Hempelian model, we might propose the following fifth condition:

(5) Screening Off: No screened off event can be part of an *explanans.*

This screening off condition will have be unable to countenance cases of legitimate over-determination. For example, consider a case where Mr. Body is shot, stabbed, poisoned, and bludgeoned in such a way that each is equally likely to result in the person's death at the same time. It is plausible to think that each of the parts of the story have an equal claim on explaining Mr. Body's death.[1] The screening-off condition forces the Hempelian to take a fairly strong position *against* the legitimacy of any sort of over-determination. An additional problem comes in deciding which gets screened off here. Something has to get screened off, but it is probably impossible to choose which one. The screening off condition will lop off all the possible *explanantes* and the death of Mr. Body will be left without explanation.

It's not clear that **(5)** will help with the light pole case. There's no fussing around finding the *real* cause and correlating the rest of the case with that. It's clear that *one* of the explanation proposals is really an explanation while the other isn't. Hempel's defender must again add a condition to try to eliminate the wrong correlation in asymmetric explanations. There are a number of ways that such a condition might be phrased. We might give an earlier-than condition to force the right answer on these questions.

(6) Earlier Only The event in an *explanans* must be earlier than the event described in the *explanandum.*

This seems to be dissatisfactory as a way to rescue a broadly Hempelian view.

---

[1]It is not unreasonable that many normal ways of filling out the case of Mr. Body's death will do away with this strong claim. Further, it's not crucial that in fact the real cause of Mr. Body's death is all of the causes. All that's needed is that the case, as specified, is indistinguishable for Hempel's view when using the screening off condition

There are cases of explanation in which an ongoing event is explaining another ongoing event. For example, the wooden structure behind the walls of a house explains a house's remaining upright. There are two ongoing events. There is the ongoing event of the house having thus-and-such a wooden structure which is the *explanans* and there is the ongoing event of the house's remaining upright which is the *explanandum*. We can't weaken the condition to "must not be later than," as this would make the explanation bidirectional again. That is, the house's remaining standing would have equal claim to explaining the ongoing event of the house having thus-and-such a wooden structure. We will return to the time issue in a moment.

Now, one might argue that the causally asymmetrical cases would be taken care of by the insistence on the adequacy of **(6)** and a rejection of cases which intuitively contradict **(6)**. This sort of response will take, for example, the case of the wood structure holding up the house as an extended causal event slightly temporally shifted *before* the house-staying-up events. Those that endorse this line of reasoning often do so because they endorse the view that *no* causation is simultaneous. A principle argument for this is found in Hume(*Treatise* I, iii, 2) and in Ehring (1985).

Dummett paraphrases this argument:

> If in general causes are contemporaneous with their effects, we are faced with the dilemma which Hume posed: for the cause of the cause will in its turn be simultaneous with the effect, and we shall be unable to trace the causal ancestry of an event back a single instant of time.(Dummett, 1954)

The first problem with this argument is that it only applies to views according to which *all* causation is simultaneous. All this response needs is that *some* cases of simultaneous causation exist which exhibit the kind of asymmetry required.

Kline argues that any alleged case of simultaneous causation will either (i) fail to be simultaneous on pain of breaking the speed of light and thus contravening the special theory of relativity or (ii) will collapse into a case that is non-causal in

nature.(Kline, 1980) The classic shadow/light-pole case falls easily to this criticism. If we are to claim that the light pole/shadow case is simultaneous we will need the light to travel from the pole to the ground at faster than the speed of light, which is impossible. But consider a more intuitive case of simultaneous causation that will illustrate this kind of response. Suppose that I have a ten foot pole. Intuitively, my moving one end of the pole instantaneously moves the other end. But this is precisely the kind of thing that relativity disallows.(Kline, 1980, 296) Now, in order to respond to this argument, we will have to draw from the work of Michael Huemer, who takes *all* causation to be simultaneous and continuous.

Huemer argues that a continuous rather than a discrete view of causation forces one to accept causality as instantaneous:

> Just as the 'particulate' structure of time envisaged by Hume requires a sequential view of causation, a continuous temporal structure requires a simultaneous view of causation, given one plausible auxiliary assumption. This auxiliary, which Hume himself accepted (I, iii, 2), is that o the impossibility of action at a distance: 'nothing can operate in a time or place which is ever so little removed from those of its existence.' The thought is that a causal factor cannot exercise a direct influence on anything at a time when that factor does not exist. We interpret this to mean that no occurrence can be directly causally relevant to an occurrence at a non-zero temporal distance from it.(Huemer, 2003, 561-562)

But what of the problem mentioned by Kline? On this view of causation the taking of time for motion to propagate from one end of a stick to another works perfectly fine. The instantaneous causation happens when we get down to the smallest level. Time being continuous will allow the effect to propagate over the stick in a finite amount of time while the direct causation is simultaneous (at a temporal distance of 0).

But if this is correct, then perhaps the criticism of the Hempelian view is not so straightforward after all. Such critique is going to ultimately depend on bigger

theoretical commitments about the nature of time and causation. But it at least seems plausible that there are cases of simultaneous causation that exhibit the asymmetry required to defeat the Hempelian view. In such a case the Hempelian view will have trouble squaring with our intuitions. Consider another attempt at fixing the Hempelian view:

(6′) Only *Causal* Laws The only laws that are admissible into the *explanans* are the true causal laws.

(6′) may help, but it's problematic for the Hempelian spirit. I take the Hempelian spirit to be broadly Humean, attempting with explanation what Hume took himself to be doing with causality. The Hempelian project has failed if it is forced to admit unreduced causality into the basic account of explanation, assuming the Hempelian project is an attempt at giving an analysis of causation; we've had causation to the account. We've not successfully reduced causal explanation to non-causal factors. The prospects for an *honestly* Hempelian solution using **(6′)** are limited. However, whether this is in the Hempelian spirit or not, it may be a solution to the problems already discussed. It's not the real causal law that associates from the shadow to the height of the light pole, but the other way around. But there are other problems with this account even if **(6′)** works.

3.2.2 Problems for Necessity

There are also problems with the necessity of (1) and (2). The classic problem with (2) is the syphilis/paresis case. A small percentage of syphilis sufferers go on to contract a further disease called paresis. Now, this is a low percentage of cases, such that having syphilis doesn't make it *probable* that one will get paresis, but the only cases of paresis are in those who previously had syphilis. Now, we want to be able to say that having had syphilis explains one's having contracted paresis, but it turns out that this is not borne out by the Hempelian theory. It fails condition (2). But this is intuitively an explanation, so we should give up condition (2) as a necessary

condition of something being an explanation.

More controversially, one might think that there are examples of explanations that do not appeal to laws at all against (1). Consider almost any normal everyday explanation. Someone asks why there is broken glass on the floor in the kitchen and one answers "because I dropped a cup on the floor." This seems to be an explanation despite the fact that no laws are a part of the explanation given. Does this get rid of condition (1)? Hempel has something to say about this sort of case.

Hempel sees that there will be many cases of explanation which don't mirror Hempel's ideal model in this way and makes room for them by calling these partial explanations. We can say that the fact that ⟨I dropped the cup on the floor⟩ is an explanation only in an extended sense. These aren't strictly speaking full explanations, but provide some key portion of what would count as a full explanation. Hempel calls these "explanation sketches" and is perfectly okay with saying that cases like the broken glass case are merely this degraded form of explanation. However, this doesn't entirely do away with Hempel's problems regarding condition (1).

There is a cost to maintaining this thesis, and that cost seems to be that almost no one ever gives a proper complete explanation in the Hempelian sense. Consider normal every day explanation. It almost always mirrors the broken glass case. In everyday life, one explains one's current Pandora station by appeal to one's preferences, one's tea steeping too quickly by appeal to the water's being too hot, one's laundry having too many wrinkles by one's having failed to remove the clothes from the dryer quickly enough, and one's being out of cereal by the fact that one has eaten it all. The examples go on indefinitely. None of these are cases where one appeals to a law in giving an explanation. This is not how the practice works in ordinary human life.

In fact, there are only rare moments even in the science classroom that one gives such an ideal Hempelian explanation. There are very specific cases, such as one's being asked why the ping pong ball and the heavier bouncy ball will hit the ground

at the same time if released simultaneously from the same height in a vacuum. In such a case the explanation given will *just* be an appeal to the law that gravity accelerates objects at the same rate independent of their mass. But many other cases are not. For example, one explains the purple color of a particular neon light by asserting that it emits light waves at thus-and-such purple wavelengths,[2] but one doesn't need to appeal to any laws of optics or perceptual psychology. Perhaps this is a case of eliding a law in the mention of "thus-and-such purple wavelengths." But in any case one has given an explanation without explicit appeal to a covering law. Hempelians have to be comfortable with the Hempelian view of explanation not having much to do with most of our normal everyday explanatory practice. It seems that such relevance is at least a desideratum for a theory of explanation.

3.3   The Statistical Relevance Model

Some subset of these problems is taken by many to be destructive of the Hempelian view. Perhaps there is some way out. Now, there's another view in the broadly Hempelian spirit that may deal with some of the drawbacks of the Hempelian model. Perhaps the statistical relevance model will be of some help. The statistical relevance (or SR) model of explanation comes ready-made to deal with some of the problems of the Hempelian model. Specifically, it is tailored well to deal with the paresis case[3] and other cases of explanation where the explanation doesn't in fact make the explanandum more probable than not.

For Salmon's SR model of explanation, something's being explained amounts to producing a full set of probability calculations regarding the probability of the event

---

[2]In most cases, the color emitted by some neon light is a certain purple color rather because it is emitting light on a distribution of frequencies that end up looking purple. This technicality is not needed in the above case.

[3]This is the case discussed in the beginning of section 3.2.2 where the explanation of a patient's having the disease paresis is often that the patient has had syphilis—very few patients with syphilis get paresis, but the only patients that ever get it have had syphilis and it is largely agreed that syphilis is the cause of paresis.

given. Salmon says:

> We may think of an explanation as an answer to a question of the form, "Why does this x which is a member of A have the property B?" The answer to such a question consists of a partition of the reference class A into a number of subclasses, all of which are homogeneous with respect to B, along with the probabilities within each of these subclasses. In addition, we must say which of the members of the partition contains our particular x.(Salmon, 1988)

In explaining why x which is in A has property B, we divide the class into subclasses (with each class and subclass being a partition), not stopping until all members of each subclass all render the same probability for B. This is what Salmon (1988) means by the subclasses being statistically "homogeneous with respect to B." Finding the subclass to which x belongs is enough to give us the statistical explanation of x's being B. Salmon characterizes this more technically as follows:

> By a *partition* of a class F we mean a set of subclasses that are mutually exclusive and exhaustive within F—that is, every member of F belongs to one and only member of the partition. Each of these subclasses is a *cell* of the partition F. A partition is *relevant* with respect to some attribute G if the probability of G in each cell is different from its probability in each of the other cells. The possibility is left open that the probability of G in each cell is different from its probability in each of the other cells. The possibility is left open that the probability of G in one of the cells is equal to the probability of G in the original reference class F. A class F is *homogenous* with respect to the attribute G if no relevant partition can be made in F.(Salmon, 2006, 63)

3.3.1 Problems of sufficiency

There are problems with the SR theories of explanation. There are two main ways that you can game the statistical relevance connection to make it non-explanatory. First there are cases where an explanatory class (i.e., where intuitively an individual's membership in a class is explanatorily relevant) can be statistically homogenous with

another class that is non-explanatory. For example, suppose having been immunized for the flu is 100 percent effective in preventing the flu. Suppose, also, that there are human beings somewhere who have never been exposed to the flu virus. Someone never exposed to the flu virus also has a 100 percent probability of not getting the flu. Now, the explanation of why any one individual didn't get the flu will be one or other of those but not both. So the SR view ought to yield that there are two different explanations here. However, on the SR view, the explanation will have to be that the person who didn't receive the flu shot is in the class of people who either got the flu shot or were never exposed to the flu virus. But, for example, for the person who got the flu shot but was never exposed to the flu, the explanation for why no flu occurred is that the person was never exposed and not that they had the perfect flu shot. But the classes are not able to be partitioned any further in terms of probability. This is a problem for the view.

Screening off doesn't seem to help this case either, as its the individual's membership in the disjunctive class of those who have either gotten the flu or were never exposed to the flu that explains the person's having not contracted the flu. Perhaps we will want to add criteria to the SR model in addition to the homogeneous class condition. But such an addition, it seems, would simply require that we split the classes so that the partitions also separate out the different *real causes* from each other. But this sort of requirement again fails to analyze out causality.

3.4 Other Non-Causal Views

There are other views of explanation which denegrate the centrality of causality to an account of explanation. There are views which take familiarity to be central to explanation. In these views to explain some phenomenon is to make it familiar, usually by incorporating the phenomenon into some theory. Another view takes unification to be fundamental to explanation. On this view one must bring together many different

phenomena and unify them in order to explain. For example, what's fundamentally explanatory of the orbits of the planets is Einstein's relativity rather than Newton' laws of physics since relativity is able to bring together more and different phenomena.

But both of these views have problems. The familiarity view will have trouble understanding how explanations like we provide in quantum physics can be considered explanations. It certainly doesn't seem that the explanation of an interference pattern in the double slit experiment is explained by something as unfamiliar as wave-particle duality and superposition states. This kind of explanation seems, intuitively, to make the phenomenon *less* familiar.

Regarding unification, consider this case. Suppose there is a world which is reasonably similar to ours. There are a number of different scientific laws which apply to different kinds of phenomena. There are a number of different scientific laws which apply to different kinds of phenomena. There are laws of optics; there are laws of chemistry; there are laws of biology. Further suppose that in this universe there are broader mathematical laws which might unite these various kinds of laws. But, finally, in this constructed universe, it turns out that the forces and entities to which this set of unifying equations appeal to do not exist. In this universe the broader laws are just convenient works of fiction. The unification theorist will be committed to saying that the broader set of equations is the real explanation in this case. But intuitively this is false. The unification theorist has failed to provide an adequate understanding of explanation.[4]

3.5   Causal Theories of Explanation

There are a number of views which one could rightly classify as causal views (in fact both the Hempelian and SR views may properly be called causal albeit attempting it

---

[4]There is much more that can be said on this particular issue, but we need not dwell on the issue here. For some discussion of the unification theory of explanation and some of the connected issues see, for example, Kitcher (1981), Barnes (1992a), Barnes (1992b), and Mulnix (2011).

in a reductionist way). There are a couple of different fundamental ways that causal explanation can be understood. We can understand causality in an event-law sort of way and in an individual powers sort of way. On either of these causal views, to explain something is to describe some part of its causal history.

The biggest challenge for causal views of explanation will be spelling out exactly what part of the causal history is the right part to appeal to (or whether, implausibly, we ought to require the whole causal history of whatever event or act is in question). Further, the choice must be principled and not merely an ad hoc way of avoiding the explosion of caused causes in the causal history of any object.

The problem is best seen in connection with what needs to be shown to be *the* salient explanation in the case of the commonly accepted explanations of our experiences.[5] Consider the case of my having an experience as of a baseball. The accepted, nonskeptical explanation of my experience is that light-waves bounce off of the surface of the ball and travel through the air, hitting my retina and activating my optical nerve, followed by activating the neurons in my brain in some pattern or other. Finally, somehow an experience is produced. This is a very abbreviated portion of the causal history of my experiencing a baseball, but this is the story that one needs if there is going to be any hope of making anything like an IBE against skepticism. The cause of my experience of the baseball needs to be the baseball, but the baseball is only a tiny part of the causal history of my baseball experience.

One question to ask is what's so special about the baseball that it is named as *the* explanation of my baseball experiences? Surely our intuitions line up with this, but what's needed is a principled way to prefer the baseball to other members of the

---

[5]It will be important to be able to settle on some one part of the full set of causes of an event for antiskeptics because, for example, it needs to be that the explanation we ask for in consider IBE's against skepticism to be something beyond the mere brain processes which cause our experiences, since this will be consistent with Brain-in-vat scenarios. It also cannot turn out that the explanation of our experience turns out to just be the entire history of the universe, since the entire history of the universe is made much less probable by our evidence than the antiskeptical hypothesis even assuming that our evidence makes the antiskeptical hypothesis rather plausible.

experience's causal chain based on our theory of explanation.[6] This is where some problems arise. How does one choose among the parts of the causal history of the baseball to pick that one (or more) special part(s) and call it the explanation.

3.6   Proposals for Limitations

Some think this isn't a problem. Nancy Cartwright takes a primary advantage of the causal views of explanation to be their ability to avoid a subjective choice of relevant factors to consider.(Cartwright, 1979, 426) In a sense she is right. It's not clear that the sort of reference class problem that statistical analyses have is applicable to explicitly causal theories of explanation. Cases like the flu case are telling against this kind of view. *This* sort of problem doesn't seem to apply to causal views of explanation. If we're appealing more explicitly to causation, our explanation of my remaining healthy will not be some disjunction of probabilistically equivalent classes.

The problem for causal theories of explanation is not of the kind discussed above. The problem is, rather, picking out the baseball from the entire history of the universe as the key link in the causal chain.[7] The problem for antiskeptics is only a symptom of the larger problem. If one were to ask what one gives in a causal explanation the natural thing to say would be "the cause of the event." But really *the cause* of my baseball experience it seems should either just be the most immediate proximate cause of the experience (in this case a certain brain state probably) or the entire history of the universe (probably an entire state description of the universe plus its history back to, say, God's hammer knocking out the big bang). In a certain sense the only thing that is a full explanation of the event of my experiencing a baseball is the one that begins with the divine hammer. The immediate problem is that this

---

[6]One might assert that all we need is that the *interesting* cause of our experiences be the baseball, with some principled way to pick it out. This is true, but I think it would be nice if we could get this out of our theory of explanation. It would be advantageous if this just fell out of our theory.

[7]This isn't a unique problem for this view and will rear its head in a slightly different way for the Hempelian as well.

kind of explanation, although intuitively right in answering the question of what *the* explanation is, is impossible to give. Not only would it be impossible for a human to describe all the elements of the explanation, it would be impossible to ever hold an explanation in one's head, let alone reason about which proposal provides a *better* one. But intuitively, we (at least occasionally) give explanations and minimally for one wishing to make an IBE against skepticism these do not constitute an infinite set of propositions or an entire history of the universe. Further, we should, in our antiskeptical moods, hope that some of these explanations will mirror what we take to be the story in our baseball case above.

3.6.1 Contrastive Explanation

Due to concerns like this, some individuals like VanFraasen (1977) think that explanation as it is relevant to us is indelibly bound up in pragmatic issues. Peter Lipton thinks that pragmatics are up to the challenge of finding those elements that are salient. He takes the solution to be making all explanation essentially contrastive.(Lipton, 2004, 71-90) Asking "Why am I experiencing a baseball" isn't enough to fix exactly the portion of the causal history one is asking for. One needs to ask a contrastive explanation question to avoid the problem and give oneself a leg up in the task of causal triangulation.(Lipton, 2004, 41-54) The different contrasts you choose will affect what the proper explanation is in our limited sense.

For example, the answer to the question "Why am I experiencing a baseball rather than a football" is that the stadium I'm at is a baseball stadium and not a football stadium and that there's a baseball in front of me and not a football. The question "Why am I experiencing a baseball rather than nothing at all?" will have a more expansive answer, concentrating on facts about the functioning of my brain, my being awake and my having functioning sense organs, along with there being plenty of light bouncing off the baseball and traveling into my eyes and traveling along my working

optic nerve and into my working brain.[8]

This is all well and good apart from the apparent subjectivity of the story (it's going to be our interests, etc., that effect which questions we ask and consequently which answers we get to our explanation seeking questions). But even if there are no issues whatsoever with this way of picking out particular parts of a causal explanation, there remains the issue of finding precisely that question that is the right question to ask for which the ordinary explanation of our experiences in the classical real world hypothesis is the salient one on an antiskeptical hypothesis.

This question actually turns out to be very difficult to track down, and what the contrast class *is* may look like it will be different depending on what sort of antiskeptical argument one is attempting to make. In any case the questions will need to be much more sophisticated than the questions we ordinarily ask when asking for explanations of our experiences. The kinds of explanatory questions we attribute to the various defenders of inference to the best explanation arguments against external world skepticism will come from the thrust of the particular arguments. For someone making something like Laurence BonJour's antiskeptical argument, which takes the key data that motivate antiskeptical IBE's to be the coherence of our experience in different modalities and through moving objects with them being 3-dimensional objects in space, the questions will need to be something like "Why are my experiences consistent with being projections of the surface of a 3-dimensional baseball shape rather than not being so consistent?"(BonJour, 1999) Someone wishing to defend an argument along the lines of Vogel's, which takes the key data in motivating IBE's to be various necessary truths about physical objects, will ask various questions about our experience like "Why are my sense experiences non-colocating experiences rather than colocating experiences,"(Vogel, 1990) or "Why are my sense experiences

---

[8]One might worry that to give the full causal explanation in either case will require a very expansive story, such that contrastive questions will not help us here. If this is so, then there is yet another, perhaps damning, objection to this attempt at limitation.

triangle inequality obeying experiences rather than triangle inequality disobeying experiences?"(Vogel, 2008, 547-549)

In either case the author takes the *best* of the options for explanatory answers to these questions to be the antiskeptical hypothesis. Does the fact that the sorts of questions we ask being formed by subjective considerations like one's interests somehow cause extra skeptical problems for the defender of an IBE against skepticism? There are two ways that this approach to limiting the causal question can cause more troubles for the antiskeptic. The first problem is that limiting oneself to discussing one particular question makes it difficult to say that one has found that the best explanation of one's *experiences* is the real world hypothesis. The best one can do (assuming the argument works) is that, with respect to the particular questions one has asked, the best explanation of that particular contrastive fact is the real world hypothesis. What reason do we have to believe that the best explanation of some other contrastive fact isn't, say, the evil demon hypothesis? Why should we believe that, in fact, there aren't a large number of other questions to which the answer that's the best explanation is the evil demon hypothesis.

We can construct such cases relatively easily. Consider a classic case in which IBE's are discussed. Suppose one is on the beach and sees what looks like bootprints going in a certain direction in the sand (see Fumerton (1992) for a brief discussion of the beach footprints case). We classically think that the best explanation of these shoe prints is that a human being trod on the beach with shoes on. But consider this modification of the case:

Suppose we have a familiar case where there are shoe prints in the sand. Suppose, further, that they are particularly deep shoe prints. Now consider two different contrastive explanation questions that one might ask. One might ask "Why are there shoe prints in the sand rather than some other kind of prints?" Upon asking this question, it should be clear that the best explanation of that contrast is that the

prints were made by a human being walking along the beach rather than some other creature. However, one might also ask "Why were there shoe prints of such and such a depth rather than of any other depth?" Here the best contrastive explanation may very well be that the shoe-prints were made by a cow wearing shoes.

But this possibility then raises the question of just how strong a reason succeeding at an IBE can be if we are limited to discussing best explanations of the very particular contrastive facts present in the particular question one has chosen. This seems intuitively to weaken the maximal effectiveness of either Vogel's or BonJour's IBE's against skepticism. At best these arguments will end up providing some *prima facie* reason for believing the common sense view or real world hypothesis. Something like the depth question above may come along at any moment and provide defeat for the IBE argument.

A strategy to increase the force of IBE's that one might use is to simply combine all those questions that might result in an IBE argument against skepticism into one big contrastive explanation question. There is at least one problem with this strategy. It's not clear how feasible it is to come up with a contrastive explanation question for such combinations. How is one to ask a question that combines the questions above? Are we to ask "Why are my experiences consistent with being the projections of 3D objects and triangle inequality obeying rather than either not so consistent or no so obeying?" Perhaps this can work, but there remains the problem of perhaps unintentionally selecting those particular questions that will result in an antiskeptical answer, just as, say, a Berkeleyan idealist might ask those particular questions which will yield a Berkeleyan answer. But this objection isn't a knock-down for antiskeptical IBE's. One should just keep this in mind when one goes forward.

There may be another problem for the antiskeptic in adopting the contrastive explanation option for limiting the scope of the explanation in question and triangulating in on the cause of some *explanandum*. This is a problem if we want ordinary

human beings to be justified in holding the antiskeptical conclusions aimed at in philosophical IBE's. Ordinary human beings (even sophisticated non-philosophers) have, presumably, seldom done such sophisticated examination of which question they ought to ask in defeating skepticism. This will make it yet harder for the ordinary human being to be justified in being an antiskeptic.

3.7   Summary

Several views of explanation have been discussed, and, tentatively, we should conclude that explanation (at least insofar as it is important for inference to the best explanation arguments against external world skepticism) is fundamentally an enterprise of finding causes. If the regularity theorists are right, then this will mean we are doing something along Hempel's lines but by limiting one's appeal to laws to those that are legitimately causal. If however, other views of causation are right like, say, a view that's a more "individuals with primitive causal powers" view one can give an explanation simply by naming a cause while talk of causal laws in discussing such cases may be a useful way to understand more precisely the nature of the causal powers in play when we name a specific cause. There's no need to come down on a specific one of the many sides in the debate on causation or on how to express legitimate causal explanation. All that is required, moving forward, is that explanation as it matters to IBE's starting from experience, is fundamentally causal. Later on in discussion we will see that appeals to simplicity will take it to greatly increase an explanation's complexity to appeal to a greater number of explanatory laws. On the view mentioned just above, it would require some maneuvering to make sense of how such explanatory law complexity could weigh so heavily against a view's simplicity. We shall assume, for the sake of the argument, that such an problem can be successfully overcome.

Chapter 4

Explanation: Fundamental?

4.1 Introduction

So granting that explanation is causal (insofar as those considering antiskeptical best explanation arguments are concerned), one may expect a move directly to consideration of the arguments. But first it must be established what the supposed virtues of explanation are and whether having these virtues makes a view more probable. It seems plausible that, in order to establish this, we will have to consider proposals for virtues and attempt to reconstruct them in a way that has direct import for the probabilities of a theory.

But there are some views according to which the very impulse here is misguided. There are views of the nature of non-deductive argument and of justification that argue that explanatory virtues are more fundamental epistemic categories than epistemic probability. To hold best explanation argument hostage to epistemic probability for this kind of view, is like waiting to apply the term "unmarried" until one has been enlightened as to its relation to the term "bachelor." One has the direction of analysis all wrong. One this kind of view, insofar as the notion of epistemic probability is useful it is parasitic on explanatory considerations.

In this chapter I will consider Harman's argument that all non-deductive inference is reducible to inference to the best explanation.[1] I will show that this argument fails. Second, I will consider Ted Poston's(Poston, 2014) argument that all argument reduces to consideration of explanatory virtues and show that it fails. After doing this I will consider problems for views which assert that what justifies belief that p is fundamentally some fact about best explanations.[2] These views and arguments share the label "explanationism." We will need to grapple with these views and arguments before asking the question about probabilistic vindication of the explanatory virtues.

4.2   Harman's Argument from No False Lemmas

The basic structure of Harman's no false lemmas argument for the fundamentality of inference to the best explanation is as follows:

(1) The no false lemmas analysis of knowledge is true.

(2) If there are cases of non-deductive inference which are not reducible to inference to the best explanation, then the no false lemmas analysis of knowledge is false.

(3) Therefore there are no cases of non-deductive inference which are not reducible to inference to the best explanation.

The main idea in this argument is to point out cases where the no false lemmas response to Gettier-like worries is intuitive and show that if these inferences are not best explanation inferences then the no false lemmas response is not open to us. Now, the developed version of the no false lemmas view of knowledge is that justified true belief is knowledge only if there are no false essential premises in one's inference. If we understand all non-deductive inference on the model of inference to the best explanation then cases like this one will make more sense.

Consider what lemmas are used in obtaining knowledge from an authority.

---

[1] For Harman's claim that even simple inductions are cases of inference to the best explanation see (Harman, 1965, 90-91). For his claim that even the relationship of entailment is an explanation relationship see (Harman, 1973, 162-163).

[2] An individual who proposes one of the most plausible versions of this view is McCain (2014).

Let us imagine that the authority in question either is a person who is an expert in his field or is an authority in this sense. When an expert tells us something about a certain subject, or when we read about the subject, we are often warranted in believing that what we are told or what we read is correct. Now one condition that must be satisfied if our belief is to count as knowledge is that our belief must be true. A second condition is this: what we are told or what we read cannot be there by mistake. That is, the speaker must not have made a slip of the tongue which affects the sense. Our belief must not be based on reading a misprint. Even if the slip of the tongue or the misprint has changed a falsehood into truth, by accident, we still cannot get knowledge from it. This indicates that the inference which we make from testimony to truth must contain as a lemma the proposition that the utterance is there because it is believed and not because of a slip of the tongue or typewriter. Thus our account of this inference must show the role played by such a lemma.(Harman, 1965, 93)

The idea is that, unless the inference here is an inference to the best explanation at the most basic level, it will be impossible to capture the intuition that the absence of misprints or the absence of slips of the tongue are *essential* premises for the purpose of fulfilling the no false lemmas criterion for knowledge. The view that any of these kinds of argument are merely complex enumerative inductions cannot account for the intuition that these lemmas are important for making one's non-deductive inferences knowledge.

What's the non-Harmanian to do? As far as direct responses to this argument go, the easiest response is to argue that premise (1) is false. The no false lemmas response to the Gettier problem is far from universally accepted. Among interesting alternatives in the literature today are infallibilism (the only view that can claim a full unqualified solution to the Gettier problem, but perhaps at too high a cost)[3] and

---

[3]Infallibilism, it is argued, has a fairly straightforward skeptical result. All those propositions for which your justification falls short of certainty will be excluded from the category of knowledge (which will probably include all external world beliefs and beliefs about other persons). Such skepticism may be too high a cost to pay for an unqualified response to Gettier.

recent developments of the virtue theorist's response to the Gettier problem by John Turri.[4] As Harman's view is a sophisticated development of the no false lemmas view, Turri develops a view like Sosa's "apt belief" view.[5]

Turri argues that Knowledge is *adept* belief. He says, "you know Q just in case your truly believing Q manifests your cognitive competence." This is meant to address problems with the Sosa-style view where we can make Gettier style cases where one's belief is true in some way because of one's intellectual virtue without the belief's truth manifesting it. These sorts of case center around cases where someone, for example, will reward someone's epistemic skills by *making* the belief true.[6]

But both the infallibilist and the Turri/Sosa view are varying precise developments of the generally agreed upon fact that it is *luck* that is to be extricated if we are to fully Gettier-proof our knowledge. But all this is to say, the argument has no force if we take one of the myriad other responses to the Gettier problem that try to extricate luck in some other way from our knowledge analysis.

Further, I'm not sure it's reasonable for the defender of the Harman line to require *every* person basing her belief on precisely the set of premises that are convenient for the purposes of mounting a no-false-lemmas defense. It certainly seems *possible* that an individual believes the testimony of those in authority due to an enumerative induction rather than an inference to the best explanation. And there doesn't seem to be any reason why such a person *couldn't* retain knowledge in such a case without having to worry about the false lemma problem that the inferer to the best explanation has. But the challenge immediately becomes how to explain how the kinds of cases which Harman can explain as Gettier cases can be explained by someone who denies

---

[4]Among the other interesting ways of understanding knowledge that attempt to solve or diagnose the Gettier problem are the view that knowledge is undefeated justified true belief (see, for example Lehrer and Paxson Jr (1978)), the view that knowledge is somehow closely tied to truth tracking (see Nozick (1981), and the generally accepted view that luck is somehow important in accounting for the lack of knowledge in Gettier cases (see, for example, Unger (1968).

[5](Sosa, 2007)

[6]see Turri, 136

Harman's view.

4.3   Gettier and Harman's rejoinder

Harman would respond by pushing the critic to take an actual view on the Gettier problem. We're doing away with an answer to the Gettier problem so we owe at least a sketch of the kind of account that will move toward a solution to the problem without giving up too much ground to Harman on the nature of argumentation and justification.

An interesting family of responses to the Gettier problem are what I will call broadly infallibilist. Broad infallibilism I will define as that family of views according to which the conditions for warrant[7] make a truth condition on knowledge superfluous. Included in such views will arguably be Turri's view (to be discussed below) among others. Also included in broad infallibilism is narrow infallibilism which is that family of views according to which knowledge level justification is justification to absolute certainty. I'm attracted to a solution somewhere in the broad infallibilist family of views. In contrast to narrow infallibilism, I will call those views which are broadly but not narrowly infallibilist "friendly infallibilism."[8]

4.3.1   Narrowly Infallibilist Views

Narrowly infallibilist responses to the Gettier problem have it easy in terms of actually solving Gettier-style cases. They easily do away with the luckiness of Gettier cases by taking luck entirely out of the equation. One can't be lucky in an epistemic sense if one has infallible justification. One's justification literally couldn't go wrong.[9] But the classic formula for creating a Gettier case is to take an individual case of

---

[7]I use warrant here to refer to whatever is added to a true belief to make it a case of knowledge.

[8]I call such views "friendly infallibilism" because it seems that such views will have less drastic skeptical consequences than narrow infallibilism.

[9]This view require justification to probability 1.0, such that it's both epistemically and logically impossible to have knowledge and have a false belief or misleading evidence while having justification sufficient for knowledge.

a well-justified belief, change the world in a way that would normally make this justification misleading and the belief false. Finally, one makes the belief turn out, by chance, to be true.(Zagzebski, 1994) One can't make an infallible justification misleading or make it that the belief would normally turn out false. One also cannot make the belief true by chance.[10] Further, narrow infallibilism has such a stringent account of knowledge that it will be impossible to produce any cases that it will call knowledge but we will have the intuition that the case isn't a case of knowledge.

The narrowly infallibilist response to Gettier problems also has advantages on other epistemic issues. It will be able to avoid the issue of endorsing abominable conjunctions discussed in chapter 1, since one is always absolutely certain of those things that one knows. Further, there are other puzzles about knowledge that narrow infallibilism solves. Infallibilism also solves a problem with knowledge closure. By a reasonable closure principle, Knowledge that $\langle A \rangle$ and knowledge that $\langle B \rangle$, along with application of basic logical rules, gives one knowledge of $\langle A \wedge B \rangle$. But if one is not an infallibilist, then all one has to do is set the justification one has for $\langle A \rangle$ and for $\langle B \rangle$ right at the knowledge-level justification threshold and $\langle A \wedge B \rangle$ will fall far below the threshold for knowledge. A narrow infallibilist can, without reservation, endorse such a closure principle without getting into such trouble.[11]

---

[10]There is a tricky case for this kind of view involving the justification of necessary truths. It turns out that *any* premise, strictly speaking, entails *every* necessary truth. So it seems that we have infallible evidence for every necessary truth. In response to this problem one might require that one's evidence *relevantly* entails the truth of one's belief for knowledge to occur. This is the response that Fumerton (1995, 72-73) suggests in *Metaepistemology and Skepticism*. This kind of response, however, pushes us down the path of making heads or tails of relevance logics, which is troubling. I have some initial suspicions that if one helps oneself to a form of inferential internalism there is some hope of avoiding relevance logics. The key will be the requirement that an individual is either acquainted with the fact and justified that way or *acquainted with* the fact that one's justification entails the proposition in question. Whatever the relations that actually obtain between ordinary propositions and necessary ones, you will be unable to be acquainted with an entailment relation between your justification and the putative necessary proposition unless you already know the proposition is necessarily true.

[11]This is not to mention the lottery problem and the problem of the abominable conjunctions.

4.3.2   The Problem with Narrow Infallibilism

The principle problem with narrow infallibilism is the rather radical knowledge skepticism that seems to directly result. We have infallible justification about so little that the victory over the Gettier problem by narrow infallibilism seems empty. We do away with Gettier cases by doing away with nearly every intuitive case of knowledge. This is a problem but perhaps not insurmountable. Defenders of narrow infallibilism will happily (or at least begrudgingly) accept a more or less skeptical conclusion. But perhaps friendly infallibilism will have a more friendly result for the antiskeptic while dealing in a satisfactory enough way with the Gettier problem.

4.3.3   Friendly Infallibilism

Linda Zagzebski is a principal figure in the rise of friendly infallibilism. In "The Inescapability of the Gettier Problem" she argues that unless the conditions on warrant entail the truth of one's belief then one will be subject to Gettier problems.(Zagzebski, 1994, 72) In it she argues that defeasibility accounts that avoid Gettier problems make the truth condition superfluous and are broadly infallibilist but there are other friendly infallibilist understandings of knowledge. For example, the view alluded to above defended by Turri and developing Sosa is another promising friendly infallibilist view. I'll briefly discuss this view and some problems with it.

Turri's view develops a theme from Sosa's *A Virtue Theory of Knowledge*(Sosa, 2007) in which he argues that beliefs, along with other performances have a AAA structure. "We can distinguish a belief's accuracy, i.e., its truth; its adroitness, i.e., its manifesting epistemic virtue or competence, and its aptness, i.e., its being true *because* competent."Sosa (2007) Turri wishes to distinguish a fourth A in the structure of performances and of belief in order to give a fuller solution to the Gettier problem. See, the problem with Sosa's analysis is that it has Gettier cases.

For example, consider the case Turri calls Hobbled:

(Hobbled) A competent, though not masterful, inspection of the crime scene would yield the conclusion that a man with a limp murdered Miss Waterbury. Holmes saw through it and had already deduced that Dr. Hubble poisoned the victim under pretense of treating her. Homes also recognized that the scene would fool Watson, whose own inspection of the scene was proceeding admirably competently, though not masterfully. Watson had, after years of tutelage, achieved competence in applying Holmes's methods, and while Holmes was no sentimentalist, he didn't want Watson to be discouraged. "Look at him," Holmes thought, "measuring the distance between footprints, noting their comparative depth, and a half dozen other things, just as he ought to. There's no doubt where this will lead him—think how discouraged he will be." Holmes then resolved, "Because he's proceeding so competently, I'll see to it that he gets it right!" Holmes....strode across the street to where Hubble was, and kicked him so hard that Hubble was thereafter permanently hobbled with a limp.(Turri, 2012)

Watson then concludes that the murderer has a limp and is correct, but in a way that seems to meet Sosa's criteria for knowledge. The belief is accurate, adroit, and apt (accurate *because* adroit), but it doesn't seem to be knowledge.[12]

Turri makes an intuitive distinction between a state of affairs' merely being caused by a virtuous capacity and a state of affairs' manifesting a virtuous capacity. For example, my catching a vase may be caused by its fragility, but it is a vase's falling and breaking that *manifests* its fragility.(Turri, 2012, 136) There might be some account of directness or something else that analyzes away "manifesting" in the final analysis. But this distinction is intuitive and can do work to help out Sosa's analysis.

Turri proposes that we add adeptness to the list of A's in Sosa's understanding. A performance or belief is adept if its accuracy *manifests* its adroitness. This will help Sosa deal with the Sherlock Holmes case. Dr. Watson goes wrong by the truth of his

---

[12]Sosa can, in response, claim that this is the sort of deviant causal chain that one should disallow (as people who give causal theories of anything tend to need to require non-deviance as well). But I take it to be an advantage of Turri's view that he can use talk of "manifesting" a skill which will be able to effectively do the same thing while avoiding the intuitively unprincipled move of adding an "it has to be the right *kind* of cause" move.

belief merely being caused(and in its funny way) by his competent investigation. This account will then do away with a wide swath of Gettier cases making it a plausible response to Gettier's problem.

The only problem is that there is a seemingly core Gettier case that this view can't address. The fake barn case[13] clearly seems to be a case in which the belief is adept, that is, the truth of the belief manifests its competence. But, for many, it seems like we shouldn't call it knowledge.

Turri argues that the fake barn case *should* be called knowledge, but he also thinks he has the resources to allow for the case on his view. He argues that one can equally well say that the individual in the fake barn case *doesn't* have the competence to detect barns in fake barn county. I'm not sure what to think about this assertion, but it *does* seem to me that there is some promise to Turri's assertion that the fake barn case is not really a Gettier case. I've never been particularly impressed with the fake barn case. Turri gives a case which simply twists the fake barn case a bit. He suggests we imagine a world in which the individual traveling through fake barn country is an evil individual who goes around shooting missiles and destroying barns. Our individual, by chance, goes up to the one real barn in the county, justifiedly believes it's a barn, and shoots a missile at it. Turri (2012) takes it to be intuitively plausible that this individual knowingly destroyed the barn. I'm not so moved by this case. It seems just as plausible to say, in response to this case, that the evil barn-destroyer *intentionally* destroyed a barn. This is reasonably all we mean when we say that someone *knowingly* did some evil act. Nevertheless, I think the rejection of the fake barn case an acceptable cost to pay as I never experienced the fake barn case as a particularly intuitive Gettier case.

But even granting that the view can either account for or reject as secondary

---

[13]Suppose I'm unknowingly driving through fake barn county, which is full of barn facades. I look off the road and see what looks like a barn and conclude there is a barn there. It accidentally turns out that the one real barn in fake barn county is what I am looking at right now. My belief is highly justified and true, but it seems that it is less than knowledge.

the fake barn case, it turns out that new Gettier cases can be proposed for Turri's analysis. Ian Church proposes the following Gettier case for Turri's proposal for the nature of knowledge:

> **Expert Botanist:** David is an expert botanist, able to competently distinguish between over 20,000 different species of orchid. David is presented with an orchid and asked to identify its species. Using his amazing skill, he can clearly tell that this particular orchid is either going to be *Platanthera tescamnis* or a *Platanthera sparsiflora* (which look quite similar), and upon even further expert analysis he comes to the conclusion that it is a *Platanthera tescamnis* orchid, which it is. However, Kevin, David's nemesis and an expert botanist in his own right, decided the night before to disguise the *Platanthera tescamnis* orchid to look like a *Platanthera sparsiflora* orchid. Thankfully, however, Alvin, David's other expert botanist nemesis (who is conveniently not on speaking terms with Kevin), decided to try to trick David in the same way—arriving shortly after Kevin left, perceiving that the orchid was a *Platanthera sparsiflora*, and disguising it to look like a *Platanthera tescamnis*, which it happens to actually be.(Church, 2013)

Plausibly, the true belief here manifests David's cognitive competency, but it is also not intuitively a case of knowledge.[14] Therefore Turri's *Manifest Failure* has failed to provide a solution to the Gettier problem.

Church asserts that the best solution to this objection is to assert that this belief doesn't manifest the skill to a great enough degree, but then suggests that this will get one into further trouble. I will not take the degrees bait. Instead, I think that the key to responding properly here is to call attention to a key distinction in the warrant condition that might be being proposed by Turri. One (incorrect) way of understanding Turri's view is to think of it as saying that a true belief gains a warrant when the specific content of that belief manifests one's epistemic competence. To that

---

[14]One might wonder why one would reject the fake barn case but not this one, and that is an interesting question. I think the difference in intuitions has something to do with the fact that the flower is *disguised* in this case. However, it's not crucial what the difference is, since Turri's view has the resources to deal with this case.

extent the botanist above's belief does manifest his competence, but this is not the nature of the view. The view Turri recommends instead is that warrant comes when the *truth* of the belief manifests the virtue. It's a distinct state of affairs from the belief's mere having the content it does, as the warrant condition is pushed out of the realm of broad infallibilism by this false understanding of Turri's views.

Consider a case of a vase being crushed by a 2-ton weight. *That* the vase is crushed doesn't *manifest* the vase's fragility, while, perhaps, the vase's breaking easily does manifest the vase's fragility. Various states of affairs can manifest an ability while others (which are part of the same set of events) don't. In the flower case something very similar is happening. It is intuitively true both that the content of the belief manifests the scientist's competence and that the belief's content matches the last deceiver's intentions manifests the scientist's competence. But it's not clear to me that the belief's content's matching reality manifests the scientist's competence. The case fails to produce an intuitively plausible case of *the truth of a belief* manifesting epistemic virtue rather than it merely being a true belief manifesting one's epistemic competence.

The primary advantage of friendly infallibilism is the ability to avoid the fairly radical skepticism that the narrow infallibilist will be pushed into. However, it's not the case that friendly infallibilism is able to solve on its own several of the other puzzles about knowledge which narrow infallibilism can. But, at the very least, it seems that there are other intuitive responses to the Gettier problem than Harman's no false lemmas response. I find Turri's response intriguing, but wouldn't be entirely unfriendly to a number of views that fall into the broadly infallibilist views of knowledge. This will provide a cornucopia of attractive alternatives to the no false lemmas view and thus ways out for Harman's initial argument.

4.4 Harman's Argument from Examples

Harman's other argument presents various examples of arguments that (purportedly) can't be captured as mere cases of enumerative induction. But all that is just his argument that there are *some* cases of inference to the best explanation that are not reducible to enumerative induction. The point here is to address his argument that *all* non-deductive inference is reducible to inference to the best explanation. This is the important argument as reducing all non-deductive inference to IBE will go a long way toward making the case that inference to the best explanation is epistemically basic. If IBE can be shown to be epistemically basic, then there will be no need to make a further case that something's being the best explanation[15] of our evidence will justify us in believing it. But life will not be so easy.

He gives one main case to first show that there are *some* cases of inference to the best explanation that are irreducible to enumerative induction. The argument goes something like this:

1. There are some good non-deductive arguments that cannot be accounted for without appeal to inference to the best explanation.

2. If there are cases that fundamentally appeal to inference to the best explanation, then inference to the best explanation is a fundamental principle of inferential justification.

3. Therefore, inference to the best explanation is a fundamental principle of inferential justification.

After giving several classical examples of inference to the best explanation such as concluding that the butler must have committed the murder or a scientist concluding there are atoms, Harman then spells out his assertion with a different case. He says regarding the conclusion that a witness is telling the truth:

(i) we infer that he says what he does because he believes it; (ii) we infer that he believes what he does because he actually did witness the

---
[15]or our awareness thereof

situation which he describes. That is, our confidence in his testimony is based on our conclusion about the most plausible explanation for that testimony. Or confidence fails if we come to think there is some other possible explanation for his testimony (if, for example, he stands to gain a great deal from our believing him).(Harman, 1965, 89)

The proposal here is that, unless one were to consider this argument as fundamentally and irreducibly an inference to the best explanation it will be impossible to understand this inference at all. Of course, there are familiar responses. There are fairly easy ways to construct this argument as a complex instance of enumerative induction along the lines that Fumerton (1980) constructs the familiar footprints on the beach case. In the above case, one can simply run an induction on the huge amount of cases one has of one's own forming of beliefs on the basis of experience (as well as all the experience of others' belief forming that are consistent with this) for the inference in (ii) and one can run an induction for (i) based on the fact of one's own truth-telling in such grave situations along with a number of other facts, like one's own judgment that if one were under the threat of perjury one would say what one believes to be true and lack of countervailing experience. This is not as unreasonable as Harman seems to think. Further, the case of inferring to the existence of things like atoms or of things like microbes I'm not sure is such a slam dunk case as Harman makes it out to be. There are things that will give rise to a fairly natural complex induction, coming from the fairly strong inductive evidence one has for the truth of a principle of causality and of the success of positing entities in past scientific endeavors. But Harman has an argument that there are *no* cases in which inference to the best explanation is not fundamental.

Harman's argument doesn't end there. Harman argues that, in fact, *all* ampliative inference must be a concealed instance of IBE in order to bridge the epistemic gap between premises and conclusion. This is then meant to license a more radical claim.

1. *No* good non-deductive arguments can be accounted for without appeal to inference to the best explanation.

2. If all non-deductive arguments fundamentally appeal to inference to the best explanation, then inference to the best explanation is the fundamental principle of inferential justification.

3. Therefore, inference to the best explanation is the fundamental principle of inferential justification.

Harman argues:

> In practice we always know more about a situation than that all observed A's are B's, and before we make the inference, it is good inductive practice for us to consider the total evidence. Sometimes, in light of the total evidence, we are warranted in making the induction, at other times not. So we must ask ourselves the following question: under what conditions is one permitted to make an inductive inference?....If...we think of the inference as an inference to the best explanation, we can explain when a person is and when he is not warranted in making the inference from "All observed A's are B's" to "All A's are B's." The answer is that one is warranted in making this inference whenever the hypothesis that all A's are B's is (in the light of all the evidence) a better, simpler, more plausible (and so forth) hypothesis than the hypothesis, say, that someone is biasing the observed sample in order to make us think that all A's are B's. On the other hand, as soon as the total evidence makes some other competing hypothesis plausible, one may not infer from the past correlation in the observed sample to a complete correlation in the total population.(Harman, 1965)

Harman seems to say that for any induction we must look at not only the "All observed A's are B's" piece of evidence. We must look at the total evidence had by the subject. Now, when we do so, the best way to judge between the thesis "All A's are B's" and the thesis that this is not the case is by considering the explanatory virtues had by each hypothesis with respect to one's evidence. There are a couple of initial problems with this argument. First, Harman fails to consider what an opponent might be doing by discussing a simple case of enumerative induction. The idea is not to claim, yet, that some large subset of justified beliefs have this kind of justification. The idea is, rather, to propose a simple case to establish the epistemic principle.

The idea that a defender of enumerative induction might be getting at is that, at minimum, this simple case is a justified case of enumerative induction. The idea will then be to build the rest (or at least some larger set) of one's non-deductive reasoning capacities from enumerative induction along with other quasi-logical making-probable relations. Perhaps this fails, but Harman, at best, simply asserts that complex cases cannot reasonably be accounted for by the "induction is fundamental" camp. For good measure, let's attempt to reconstruct one of Harman's core cases.

The question isn't whether we almost always have a lot more evidence in normal cases, the question is whether, if "All observed A's are B's" is our evidence, we are justified in believing "All A's are B's." Harman seems to be asking the wrong question. He's just looking for defeaters or for evidence other than we are giving ourselves in the case. Of course it's our total evidence that matters, but this says nothing about inference to the best explanation being most basic. That is, of course in ordinary cases we are almost never making the pristine constructed case moving straight from All observed A's being B's to All of them being B's. This is immaterial to the case. Harman needs to establish that, even in the case where "All observed A's are B's" is effectively all of our evidence,[16] one still needs awareness that the concluded fact best explains this other fact in order to be justified in making the inference. This isn't what the case above warrants.

Harman's argument regarding the simplest cases brings up *another* interesting point. Harman is right to point out that there is something missing for an individual who merely has access to "All observed A's are B's" and concludes to "All A's are B's". The mere access to the enumerative inductive premise isn't enough to justify us in concluding that all A's are B's.

Harman suggests that we need to also have access to explanatory factors to be able to conclude that All A's are B's. We need to be aware that "All A's are B's"

---

[16]Inferential internalists will require that one be aware that "All observed A's are B's" makes probable the arguments conclusion, but I take this to be a non-evidential justifier.

is part of a better explanation than the competitors on which it's not the case that "All A's are B's".

But let's say that we have this fact as well. Are we now justified in believing the conclusion that "All A's are B's"? Maybe it's a stronger argument than the mere enumerative induction, but adding the best explanation fact doesn't solve our epistemic problem. There are two ways that we might think of how its being the best explanation of our evidence contributes to our justification. Either the best explanation fact is an added premise in the argument or it's not.[17] If it's an added premise, then I'm not sure we've bridged the gap between premise and conclusion. Something more is yet needed to get the agent aware that her evidence licenses some evidence. We can add even more premises if we like, and will be no closer. We'll be off on a Carroll regress.[18]

Before considering the case where the best explanation fact *isn't* a premise, there's one other option to consider. Suppose that the best explanation fact *is* a premise, but one leaves open what principle of inferential justification licenses a move from premises to conclusion in general. If one makes this move (and *doesn't* intend to make the move in the following paragraph), then one has given up enough for our purposes here. The kind of view where all non-deductive argument reduces to inference to the best explanation, but that we leave open what licenses non-deductive inference will move this explanationist view out of the camp where understanding inference to the best explanation in terms of probabilities is somehow inherently misguided.

But suppose the best explanation facts are not premises, but rather some kind of non-propositional awareness that "All A's are B's" best explains one's evidence. It's still not clear that the problem is solved on this way of thinking about our awareness

---

[17]Harman, based on his no false lemmas argument, will want to accept that it *is* a premise.

[18]I'm not saying this potential Carroll regress is a special problem for Harman's view. But all views that attempt to bridge the gap in induction between our premises and the beliefs they justify will have to avoid their principle of inferential justification resulting in a Carroll regress. Ultimately, I think, considerations in this area push one toward a fairly strong version of inferential internalism, but we will refrain from extensive defense of this view in this work.

that "All A's are B's" would best explain one's evidence. On this way of thinking the awareness of the best explanation fact is meant to be what gets us from mere awareness of the premises of a non-deductive inference to its conclusion. But why should we think that "All A's are B's" and it's best explaining of our evidence make "All A's are B's" probable unless we are already convinced of Harman's view of non-deductive reasoning? Even if we *are* convinced, how is it to get us to justified belief that "All A's are B's"? It's not clear that awareness of a best explanation fact makes a conclusion probable, on its own. There's a technical way to push this kind of problem, which I will discuss just a little later when arguing against explanationism as an analysis of justification. All I'm pointing to here is the strong intuition that we haven't finished the job of justification if the premise and the best explanation fact is all we have. It seems we need some reason to think that the best explanation of some phenomena is probably true.

By way of contrast, consider the position according to which induction is fundamental and, more importantly, that epistemic probability is fundamental. On such a view, no such intuitive problem or gap remains. One has "All observed A's are B's" as one's evidence and is aware that one's evidence makes probable "All A's are B's". There's no intuitive justificatory gap in this account. Now, one might attempt to rescue Harman's argument and view from this kind of worry, but this will require that one argue that the best explanation of one's evidence is always *probable*, which will require (at the very least) the kind of argumentation in the following chapter that the explanatory 'virtues' are probability increasing. But then, again, one has moved away from the stronger view that best explanation facts are fundamental. One has at least given up the view that explanatory virtue is so fundamental that a probabilistic vindication of these virtues is misguided. But let's set aside these issues for a moment and consider a different argument, proposed by Poston (2014), that all ampliative inference is fundamentally a weighing of explanatory virtues.

4.5 Poston's Arguments

Poston has two main arguments for his Harman-style view of ampliative inference.

(1) "If all arguments are comparisons of plausibility, then any ampliative inference is either explicitly or implicitly weighing explanatory virtues."(Poston, 2014, 53)

(2) All arguments are comparisons of plausibility.

(3) Therefore any ampliative inference is either explicitly or implicitly weighing explanatory virtue.

Premise (2) is justified by two different arguments. The two main arguments that Poston makes for this key premise are centered around the Sellarsian dilemma and the problem of induction. His first argument centers around the Sellarsian dilemma:

(1) In ampliative inference, one's premises are either unjustified, are justified foundationally, or are only able to be justified by resting on plausibility considerations.

(2) Sometimes the premises of ampliative inference are justified.

(3) Foundational justification has been shown to be impossible by the Sellarsian Dilemma.

(4) Therefore the premises of one's ampliative inferences are only able to be justified by resting on plausibility considerations.

(5) Therefore all arguments are comparisons of plausibility.

The key to evaluating Poston's argument is to respond to his defense of the Sellarsian dilemma. He worries about BonJour's most sophisticated response to the dilemma. BonJour's most recent view asserts that one is foundationally justified in believing ϕ if and only if:

One is aware of the content of ϕ, is constitutively aware of ψ and is aware of the fit between ψ and ϕ.

Poston addresses this view as follows: "How then does the original experiential awareness provide justification for the more complex cognitive act of "seeing the fit"?

Perhaps the most charitable understanding of BonJour's remark is that the cognitive act of "seeing the fit" is an infallible propositional awareness.

Poston's error is in thinking that a relationship of "seeing the fit" is a conceptual awareness in a way that runs afoul of the Sellarsian dilemma. As far as I understand "seeing the fit," the limiting case would be having an awareness of the correspondence between a proposition and one's state of experiential awareness. This is not the sort of 'awareness' state that is propositional in any justification requiring sense. We may call it conceptual in the sense that it's an awareness of a relation between something with a conceptual structure and a non-conceptual awareness, but not in the sense that it's the sort of thing that requires justification. It's not an application of a concept. Awareness of a "fit" relation or of a "correspondence" relation is exactly the kind of middle way to find in order to solve the Sellarsian dilemma. That is, the awareness is not conceptual, but it is an awareness of the fit (or correspondence) between something non-conceptual (one's awareness ψ) and something conceptual (the content ϕ).[19]

Poston, in his second argument, appeals to the problem of induction. It's hard to pin down exactly what the argument is, but there are a few things that he makes relatively clear. First, it's clear that Poston's second argument, rather than concentrating on how one's premises are justified, concentrates on the problem of connecting one's premises to some belief justified on their basis. The main contention is that there is no way to bridge the gap Hume showed to exist in all our non-deductive reasoning. Poston takes inference to the best explanation to be the only way to bridge this gap. The way it is meant to solve the problem is a little perplexing. He gives the standard circularity objection to defeat the inductive defense of induction and then goes on to

---

[19]One might think that even if the foundationalist view is indefensible against the Sellarsian dilemma the argument is established. For example, one might reject the antiskeptical premise (2). In fact, it may even be the case that certain broadly coherentist versions of subjective Bayesianism might accept two but take themselves to get justified beliefs out of the process by following the Bayesian updating mechanism. In any case, it seems that the antiskeptical premise (2) is a little two hastily accepted.

say that it's because foundationalism is bankrupt[20] that the problem exists and if we give it up in favor of Poston's explanatory systems view all will be well. One way of beginning to gesture at this problem is (since we've seen that Poston's argument against foundationalism fails) to assert that we can be foundationally justified in believing that some set of premises makes probable a conclusion.[21] This will ultimately require a controversial Keynesian view of epistemic probability where "making probable" is to be understood on the model of entailment (as a necessary relation between propositions with which we may become acquainted). Now, there's a lot more to be said regarding the problem of induction, but granting foundationalism at least leaves room for disagreement with the IBE is fundamental crowd.

But even if we grant that an explanationist view of reasoning can vindicate that something's being the best explanation makes it more likely than competing explanations(thus moving toward a solution to the problem of induction), there remains a difficult problem for those that wish to analyze inferential justification in terms of best explanation factors. There are three probability problems for the person who provides an explanationist analysis of either inferential justification or of justification in general. The first objection is that explanationist analyses fail to provide a sufficient condition for justification. The second two objections will show that sophisticated versions of explanationism will inevitably get into trouble with nested justifications and fail *again* to provide a sufficient condition for justification.

Explanationism most often posits justifiers (evidence for a common strand in explanationism) and best explanation features as what bridge the Sellarsian gap between these justifiers and justified beliefs. But there is a problem for explanationism that is perhaps well known, but I think not well enough respected as a significant difficulty that no explanationist view has solved. No currently existing proposal has

---

[20]If this is right, then dealing with the problems for foundationalism should be enough to deal with the problem of induction that is insoluble on any view that doesn't capitulate to explanationism.

[21]See, for example, Fumerton (1995).

given a sufficient condition for justification because of the problem of getting best explanation considerations to translate to a view that guarantees that propositions that explanationists call "justified" are epistemically more probable than not. In the end, there will be a way to cheat on behalf of the explanationist and get a sufficient condition for justification, but not without ceasing to be an explanationist *analysis* of justification. At the very least, the cheats available to explanationists will give up the view that a project of vindicating explanatory virtues probabilistically is somehow fundamentally misguided.

There's a problem with any explanationist view and even for anyone interested in using best explanation criteria to justify any proposition. This is a (solvable) problem for those who use inference to the best explanation (IBE) to justify but is especially problematic for explanationists due to its centrality to the explanationist's view, among other things. For the purposes of the following objection, let us formally restate the minimal explanationist thesis for the scrutiny that will follow.[22] Consider Minimal Explanationism (ME):

Minimal Explanationism (ME):   A proposition p is justified for subject S if p is part of the best explanation available to S for S's total evidence e.

For every explanandum there are more than two possible explanations, but believing that p on the basis of its being the best explanation of one's evidence tacitly assumes that there are only two: the p-explanation and the not-p-explanation. It assumes that there is the p-explanation and the not-p-explanation so that if the p-explanation is the best explanation then one ought to believe p. If this were the case every time one engaged in IBE then there would be no worries, but it's never the case. Instead there are almost always several choices for explanations of any set of events.

But this immediately raises the threshold problem for explanationism. Something

---

[22]I formulate what I take to be an *analysis* of justification in terms of sufficient conditions because my arguments in the following will primarily target sufficiency.

being the best explanation doesn't make it probable. It simply makes it more probable than any individual competitor. For example, a skeptic may easily grant that the real world hypothesis is the best explanation of our sense experiences while reasonably taking it to be the case that we ought not endorse the real world hypothesis. This is because one can be convinced that the best explanation of one's sense experience is the real world hypothesis while believing that a disjunction of its competitors is more probable. For example, suppose I bet on a long shot at a horse race. I miss the race, but afterward my friend tells me that I lost the bet. The best explanation of my losing may be that the favorite won, but more probably one of the other horses won. Explanationists struggle to avoid getting cases like this wrong.

To further illustrate this point, consider the case of evidence e which has three possible explanations A, B, and C. B and C are roughly equally good explanations of e while A is the best explanation, being marginally better than each of B and C individually. Ignoring the problem of figuring out exactly how being a better explanation translates to epistemic probability, let's say that(since an exhaustive set of competing explanations will add up to a probability 1.0) this all gets A being .4 probable, and B and C being each .3 probable on one's evidence. Now suppose that, on A, proposition p is true and on B and C, proposition p is not true. Granting all of this, one ought to believe that not-p. But ME will hold that one is justified, rather, in believing p since A is the best explanation of one's evidence. Poston attempts to address a worry adjacent to this one by modifying his criterion to require that the winning explanatory system have "sufficient explanatory virtues." His analysis of justification goes: Ex-J": S has justification for believing p if and only if p is a member of a sufficiently virtuous explanatory system, E, and E is more virtuous than any other p-relevant competing system E'. (Poston 2014)

As an attempt to block a similar worry concerning several bad explanations competing with each other Poston requires that the explanatory system that justifies be

a good one at the very least. But as vague as "sufficiently virtuous" is, there is no way for this to get Poston out of the current worry. All that is needed are more than two "sufficiently virtuous" explanations competing with each other. Since competing explanations are more or less probable by being virtuous or vicious in comparison to their competitors, again on Poston's modified view a belief may be justified while being rather less than .5 probable. No modification is needed to the above case to defeat Poston's account except perhaps that we stipulate that each of A, B, and C have sufficient explanatory virtues.

The other route one might take to address this worry is to allow disjunctive explanations. Suppose, for example, that we want to get out of the A B C case above. One might attempt to say that one should treat all those explanations on which p is false as one big disjunctive explanation. This would then yield that the explanation (B or C) is really the best explanation of one's evidence and the view would not have a problem. However, this is implausible and doesn't solve the problem. It's implausible to say, for example, that demon-world skepticism and brain-in-a-vat skepticism are part of one disjunctive explanation of our sensory experiences. At best it's a disjunction of candidates for explanation.

I take the distinction to be as follows. A disjunctive explanation can be treated as *a* proposal with its own *explanatory virtues* to be compared to other explanations. A disjunction of candidates for explanation is merely a set of *different* explanations which cannot be considered a unit which can be compared with another in terms of the disjunction's explanatory virtues. It's not *an explanation* at all. It's a bunch of fundamentally different things which we can combine epistemically and (perhaps) decide in non-virtue-comparative way whether the disjunction we're considering is more probable than not. If disjunctions are to solve the problem for the explanationist who grants the minimal explanationist thesis above, we must be able to consider disjunctions of explanatory candidates *as* disjunctive explanations.

Now, I'm asserting above that it is *initially* implausible to say that demon-world skepticism and brain-in-a-vat skepticism are part of one disjunctive explanation of our experiences. Some more reasons to doubt this will be considered below. Now there are cases where we propose different candidates for explanation as a disjunction any one of which we're happy to defend. For example, when asked to explain Mr. Red's death, a lawyer might assert that it was because the defendant either shot Mr. Red or stabbed him, but it was one of those. But this isn't the same as saying that in such cases one is proposing *an explanation* which is the {stabbed∨shot} explanation. One isn't considering the {stabbed∨shot} explanation as a single candidate with its own virtues. Rather one is lauding each individual explanation as pretty good and relying on the higher probability of a disjoined proposition than either disjunct.

But suppose that it were initially plausible to say that a disjunction of possible explanations is *an* explanation. On such a view, given that selection of explanations to disjoin will be based on the proposition in question,[23] radically disjoint views of the world will be collected together as one disjunctive explanation depending on what proposition we're considering for justification. Consider the ABC case above. Now suppose that there is also a proposition q that is included in A and also included in B, but excluded by C. With respect to proposition p we are told to treat (B or C) as a bona fide disjunctive explanation, but with respect to justifying q we are to treat (A or B) as the bona fide disjunctive explanation. But this way of splitting the competing explanations in just that way which allows us to get a justified belief is both needed by the explanationist and is problematic. This proposal **(a)** abandons one of the attractions of explanationism that it's easier to attain the justification required by an explanationist than what's required by those who take probabilities to be more fundamental, **(b)** requires very exotic and counter-intuitive collections of

---

[23]If we want to have a justified belief regarding the truth of p, we will collect all those candidates on which p is true into one explanation and we will collect all those candidates on which p is false so that a justified belief can be guaranteed by **ME**.

theses to be called a single disjunctive explanation, and (c) ends up undermining the assertion that what *fundamentally* justifies belief that p is that the explanation where p is true is the most explanatorily virtuous.

Regarding (a), one of the principle attractions of explanationism to come from a contrast it can exploit with views like Keynesianism or Bayesianism about justification and probability (at least for the internalist kinds). These views, when internalist, will require human beings to have some kind of access to probability facts, whether mere conditional probabilities or even the prior probabilities of some subset of propositions. One reasonable critique of these kinds of views is to point out that it seems implausible to think that actual human beings have access to such facts. One way to try to address this kind of worry is to take refuge in explanationism. The idea is that explanationism is more in line with what we actually do when we reason. We think about our experiences and what would explain them, and then come to a conclusion. Intuitively, we have *access* to facts about what best explains other facts.

On this view there is an individual who has her total evidence and there is the best explanation of this total evidence which is available to her. This gives a very clean picture of the justification of some proposition. One finds the best explanation of one's experiences and then reaches in and extracts a proposition from this explanation.

But the proposal that allows disjoining as many p-affirming explanations as one pleases (or at least as one needs to allow p to reach above the threshold for justification) runs afoul of this supposed advantage of explanationism. This is because for nearly any new pair of propositions q and ¬q there will be some different disjunction of explanations that is the right one to get above the threshold for either q or ¬q than one has entertained for any previous set of propositions p and ¬p (assuming p and q don't entail each other). But then this isn't the sort of thing that we more intuitively have access to in everyday life. This will require us to be able to mentally disjoin a large number of possible explanations *and* have some kind of access to the

fact that some incredibly complex disjunction has more explanatory virtues than another complex disjunction. And one does this again for every different proposition one considers for justification. This seems no easier for the individual to attain than having some kind of access to probability facts.

Regarding **(b)**, suppose (returning to our case above where I've lost a horse race and am considering what this does in getting a justified belief about who *won* the race)[24] that A is the explanation of my experience of losing the bet upon which Whitey (a white horse with brown patches) won the race, B is the explanation on which Browny (a plain brown horse) won the horse race, and C is the explanation on which Blacky (a plain black horse) won the race. On this way of cutting explanations disjunctively, it will be the best explanation of one's experience that either a white horse or a brown horse won the race, but it will also be the best explanation of one's experience that a plain colored horse won the race. But the fact that *either a white horse or a brown horse won the race* is a very odd set of facts to combine into one proposed explanation. But this is on an already artificially simplified case.

The problem is much worse. On this way of thinking one must, for every set of propositions p and ¬p, disjoin all those explanations on which p as well as all those explanations on which ¬p. But this threatens a new disjunctive explanation on every proposition one is evaluating epistemically (for a suitably sophisticated epistemic subject). There will be a different massive disjunction on which p is true and another on which q is true and so on (for every new proposition which is not logically equivalent to a previous proposition).

But then, even in ordinary cases, not only will there be very odd combinations which are to be counted as one explanation, but there will be disjuncts which are part of this explanation on which the truth of the proposition in question would have *nothing* to do with our evidence. Presumably, for example, if we were attempting to

---

[24]In this case we again will use the stipulated probabilities for A, B, and C of .4, .3, and .3 respectively for convenience.

justify a belief in the existence of cats, our set of different proposals to treat as a single explanatory candidate will need to be a very large and odd set of proposals. Among them will be, for example, the hypothesis that all apparent cats which we experience are mere holographic cat simulations, but that, by chance, there is a distant planet on which cats are the world's apex predator. Not only is this an odd proposal to consider along with our ordinary cat-experience explanation as forming one proposal, but it's also a proposal on which the fact that cats exist bears little, if any, on the experiences which are under consideration. Treating disjoined proposals as single explanations in order to avoid problems with sufficiency makes for strange bedfellows. These are such strange bedfellows that it's hard to consider them as a single proposal.

Further, again granting that disjunctions can be considered *an* explanation for our purposes here (it can even be granted that adding a disjunct to an explanation makes the disjunctive explanation more probable, since this is certainly true—for non-zero probability disjuncts that don't entail each other). There is another, perhaps deeper, problem, called (c) above. Propositions justified on disjunctive explanations will not be justified because the disjunctive explanation is a better explanation. It'll be justified, if at all, simply because, ceteris paribus, (A or B) is more probable than just A, which is a simple probability fact with no explanatory virtue that we can see to back it up. This seems like a lot to grant for the explanationist. Almost no cases are going to arise in our ordinary experiences where the explanationist will be able to avoid the candidates for justification merely being big disjunctions of explanation candidates. Many beliefs will, in the final analysis, not be justified in virtue of the explanatory virtues of a single proposal over competing explanations, but due to a simple probability fact that isn't reducible to some explanatory virtue or combination thereof.[25]

---

[25]The rejoinder to this from the stronger Harmanians will be to say that all probability relations are reducible to best explanation relations, but disjunction's ability to increase probability is one of the greatest stretches (apart from entailment relations) in this kind of account. A second strategy one might have for rescuing the view is to admit this but assert that one can keep explanationism

But suppose, abandoning the disjunction strategy and returning to the Poston-style analysis from above, that we hope to get ourselves out of the probability problem by adding a "sufficient virtues" condition. The problem here is that more than one explanation can be "sufficiently" virtuous each of which may even have very nearly the same amount of virtue. Having a virtuous explanation doesn't preclude other explanations from being virtuous (as opposed to probability facts, according to which one proposition being highly probable precludes other contrary propositions from being probable). But then, for Poston, we will have to just say that "sufficient virtues" just means "virtues sufficiently higher than any competitor such that the probability of p is guaranteed to be above .5."

The above proposal may yet succeed in proposing a sufficient condition for justification, but at a cost that is much too great. Any proposal with enough virtue compared to the other possible explanations to make it .5 more probable than any single competitor will be guaranteed to be above .5 probable. The problem is that this cuts out a huge number of cases that, for the explanationist, are intuitively justified in virtue of explanatory considerations. This will leave out a number of cases where one has reached a probability above .5 without a sufficient difference in virtue to *guarantee* being above .5. There are cases, for example, where the difference established might be .3 but normally one would want to take one to be justified in believing one of the proposals. For example, a probability distribution where the options are proposals with .58, .28, .01, and .01 probability. In such a case one, though having established a difference of only .3 between the best explanation and its nearest competitor, an adequate explanationist analysis will need to be able to call it justified.[26]

on the lines of Poston, with arguments being comparisons of plausibility essentially. But there being probability facts that are important to getting an IBE justification that are irreducible to explanation facts is enough for our purposes. If there's the possibility of probability facts and best explanation facts being orthogonal, the project of vindicating best explanation facts probabilistically will be required (or at least interesting).

[26]There will also be a way of filling out having gotten such a difference on which one is unjustified in believing the best explanation proposition, for example, if the distribution has proposals at .49, .19, .19, and .11.

I suspect, perhaps, a lot cases of explanation based justification will, in fact, be left out in virtue of failing to meet the newly added sufficiency condition. But not all hope is lost. There is another proposal that the explanationist could fall back on that would definitely propose a sufficient condition without leaving out large numbers of intuitive cases out of one's proposal. One could propose the following:

Cheater Explanationism(CE): A proposition p is justified for subject S if p is part of the best explanation available to S for S's total evidence e and p is more than .5 epistemically probable for S.[27]

This will solve, once and for all, any problems of leaving out intuitive cases or failing to provide a sufficient condition for justification, as all CE does is tack on an explicit probability requirement onto the end of ME.

But this is a cheat. It is an ad hoc addition to the explanationist proposal in order to get out of this one problem (as are any of the others above that include explicit probability requirements).[28] Further, the fact that one cannot propose an explanationist sufficient condition for justification without simply adding a condition that the proposal is above .5 epistemically probable suggests that, contra explanationist analyses of justification, a better strategy for the epistemologist is to consider epistemic probability a more basic category than explanation considerations. If, in order to propose an explanationist sufficient condition for justification, one must simply tack an *already sufficient condition* for justification on to the explanationist proposal, then perhaps we shouldn't be explanationists after all. Being forced to make an explanationist sufficient condition proposal *superfluous* in order to give a true proposal doesn't bode well for the prospects of an explanationist *analysis* of justification.

---

[27]This is different from the above proposal in that we have just added a condition that, on its own, will guarantee a belief to be justified. We do this rather than relying on a proposal according to which the *explanation* must be *sufficiently better* that the belief is justified. We've been forced to disconnect what provides a sufficient condition for justification from the part of the proposal about goodness of explanations.

[28]We might be cheeky here and suggest that we wait to decide whether being ad hoc is a vice for this view until we've given a probabilistic vindication of the vice of being ad hoc, but the explanationist views we're targeting won't like this suggestion.

But even supposing that the worries above can be assuaged, there is a second kind of threshold problem for explanationist analyses of justification, which arises when one allows fallibly justified propositions to justify other propositions in a non-entailing way. Threshold problems of this nature arise in different ways, but the core idea is that a number of formulations of explanationism fail to state a sufficient condition for justification because they avoid analyzing justification in terms of epistemic probability.

4.5.1   The Threshold Problem for Ex-Ej 2.0

Kevin McCain, in his version of explanationism, attempts to account for cases that are intuitively justified for a subject but that, it seems, cannot be part of an explanation of a subject's evidence. T. Ryan Byerly (2013) proposes a case in which one is golfing, having had a good amount of success. One putts the ball and watches, seeing it going toward the hole. One becomes justified, by this observation, in believing that the ball will roll into the hole. Intuitively, no proposition about the ball rolling into the hole in the future explains the golfer's now having the visual experience of seeing the ball go toward the hole. The problem for explanationism is that propositions about the future often seem justified, but it is implausible that their justification is a result of being part of the best explanation of one's present evidence.

In response to this objection from Byerly (2013) and related objections from Byerly and Martin (2015), McCain proposes Ex-Ej 2.0. Ex-Ej 2.0 states the following:

Ex-Ej 2.0   - A person S, with evidence e at t is justified in believing p at t iff either (i) p is part of the best explanation available to S at t for why S has e, or (ii) p is available to S as an explanatory consequence of the best explanation available to S at t for why S has e (p would be better explained by the best explanation of S's evidence available to S at t than not-p would).(McCain, 2015)

A key weakness in Ex-Ej 2.0 is that p can be part of the best explanation of one's evidence e but be relatively weakly supported by e. In fact, it can be just above whatever the threshold for justification may be. For explanationism, it should

reasonably be thought to be just over .5 as an explanationist will allow that p better explaining one's evidence than not-p will yield justification. All this will require is for one's justification to yield a probability of p above .5. That is, being the best explanation of one's evidence should make a belief more probable than not.

But if we set the probability yielded by the first clause to just above .5, then any proposition that is justified on the second clause based on the propositions this first clause justifies to just above .5 will fall below the threshold for justification provided that the first-clause propositions don't entail the second-clause propositions. But if Ex-Ej 2.0 calls 'justified' propositions that fall below the threshold for justification, then Ex-Ej 2.0 doesn't state a sufficient condition for justification.

Consider the following example:

> Bill the Bowler: Bill has been bowling with a reasonable amount of success. He releases the ball as it goes toward the bowling pins. His evidence includes many times observing his ball roll toward the pins. He has a hazy memory of getting[29] a strike slightly more than half the time. The conjunction of ⟨This ball is going toward the pins in circumstances C⟩ and ⟨Most balls going toward the pins in circumstances C will knock over all of the pins⟩ is just barely the best explanation of his experiential evidence, such that it ekes by in fulfilling Ex-Ej 2.0 clause (i) Now, ⟨The ball will knock over all of the pins⟩ is better explained by the above conjunction than is ⟨It's not the case that the ball will knock over all of the pins⟩, but barely. ("Most F's are G's" means "More than half of F's are G's.") Therefore, ⟨The ball will knock over all of the pins⟩ is explained (marginally) better than its negation.

By Ex-Ej 2.0, Bill the Bowler is justified in believing that the ball will knock over all the pins. But that is clearly false. This belief should be far below the threshold for justification. But Ex-Ej 2.0 implies it is justified. So, Ex-Ej 2.0 does not state a sufficient condition for justification.

---

[29]Set the memory to just hazy enough that the justification for ⟨Most balls going toward the pins in circumstances C will knock over all of the pins⟩ goes down to just above the threshold for justification.

Since this probability objection merely established that Ex-Ej 2.0 is not sufficient, McCain could strengthen the connection between clause (i) and clause (ii) propositions so that what is justified by clause (ii) does not fall below the degree to which a clause (i) justifying proposition is justified. He could add a condition or strengthen the connection between clause (i) propositions and clause (ii) propositions to entailment, as he originally suggested in his (2013). But these strategies, we worry, are ill-fated. The added conditions will have to explicitly or implicitly require the overall epistemic probability of the consequence proposition not to fall below the threshold for justification. If the view explicitly adds a probability condition, the view threatens to concede that explanatory relations are not the epistemically fundamental thing. As an explanationist, McCain should propose an analysis that uses explanatory facts in order to remain in the spirit of explanationism.

But suppose that, to defend Ex-Ej 2.0, one adds an implicit probability condition that bars clause (ii) propositions from being considered justified while falling below a probability of .5. The condition would have to involve best explanations in order to remain explanationist. What would such a limit look like? It seems that an implicit characterization will either have to require that, on the whole, the justified proposition is part of the best explanation of one's evidence (this proposal we take Byerly and Martin to have challenged with the golfer case) OR one will have to require that one's evidence better explains p than it explains not-p. But surely McCain has avoided this with Ex-Ej 2.0 because there is no plausible sense in which my mental states, caused by the world out there, explain the truth of any proposition about those things which cause my experience. My evidence doesn't explain facts about bowling balls hitting pins or golf-balls going into holes any more than a videotape of a tornado explains the tornado's happening.

This problem affects explanationism generally. Sometimes we are justified in believing propositions (for example, about the future) that don't explain our evi-

dence/experience, but are in some way consequences of the truth of propositions that explain our evidence/experience. These latter propositions are what McCain calls 'explanatory consequences'. Some of these consequence propositions that we are justified in believing are not entailed by the propositions that explain our experience (for example, the belief that the golf ball will go into the hole based on its trajectory and speed). Explanationists need to provide an account that allows for the justification of consequence propositions. The challenge we have presented is that any two-factor account like Ex-Ej 2.0 will run into this threshold problem. The problem is that, since the only thing the primary explanatory factor in these two-factor accounts guarantees is that the propositions which best explain one's experience are more probable than not (i.e., that they are above .5 probable) there is no way (short of entailment) to arbitrarily make the second factor require a high enough probability that fulfilling this second condition will be sufficient for justification. The explanationist is caught between a rock and a hard place. The explanationist must either deny a number of plausible cases of justified belief or give up too much ground to those that think epistemic probability is more fundamental than explanation.

4.5.2   The Threshold Problem for ME+Fallible Evidence

There is another threshold problem for explanationists who allow fallible evidence to confer justification. Suppose we take ME and we add the further thesis that we call fallible evidence (FE).

Fallible Evidence (FE):   Some fallibly justified propositions can be members of one's evidence.

Explanationists able to escape the previous problem but who accept (ME) and (FE) face a new threshold problem. Recall that according to (ME) any proposition that is above .5 probable on one's evidence is justified and that propositions that best explain one's evidence are not necessarily much more than .5 probable. But assuming that (FE) is true, so that the probability of one's evidence is not always

1.0, (ME)+(FE) will regard propositions as 'justified' that are less probable than their negation. Given that one's evidence is barely well-enough explained by the proposition in question to make the probability P(p/e) is just above .5, the final probability of p will fall below .5 while satisfying ME.

This threshold problem is less general than the previous one because the explanationist can avoid it by accepting infallibilism about evidence or by denying that propositions are evidence. A non-propositionalist about evidence like McCain can avoid this problem because the only things that count as evidence for him are facts of which one is aware and seeming states, neither of which are candidates for being justified, whether fallibly or infallibly.

We conclude that explanationism fails as an analysis of the epistemic support relation because it fails to provide a sufficient condition for justification. Also, if what's just above is correct, it fails to provide a sufficient condition for a proposition being more probable than not.

As can be seen, there are plenty of problems for the explanationist. The explanationist cannot provide a sufficient condition for justification and at the very least it does no better than other views at solving problems like the problem of induction. But given that explanationism will have such difficult problems even giving a view that presents a sufficient condition for a proposition being more probable than not, let alone providing a sufficient condition for inferential (or all) justification.[30] Explanationism is false as an analysis of justification. But then explanationism hasn't shown that the project of vindicating the probability-conduciveness of the explanatory virtues is fundamentally misled. In the following chapter we will do this, looking closely at the specific explanatory virtues and considering what, if anything, about them makes particular explanations more probable.

---

[30]It even seems to fail to show that it is the fundamental principle of inferential justification which will also nicely open up the project of the next chapter.

Chapter 5

The Explanatory Virtues

Having settled on a rough characterization of explanation as causal (at least insofar as we're interested in it) and raised doubts about views that think best explanations simply justify in virtue of being best explanations, it will be important to come to understand exactly what it is that makes an explanation good. The concept of a best explanation will then be built on top of the notion of a good explanation. There is a sense of explanation according to which there is only ever one explanation of some phenomenon, let alone a "good" one. But we are now interested in the sense of "good explanation" which is specifically *epistemic* in character. But an explanation is called epistemically good only insofar as it exemplifies some subset of the so-called "virtues" of explanation. So the discussion, going forward, should be of the "virtues" of explanation. We will start with the most highly agreed upon and move toward the more contentious as the discussion proceeds.

There are several notions of probability. A number of them are properly called frequency notions. On the most crude version of this theory one gets a conditional probability of, say, A given B by counting up all those instances of B and then counting on how many of them A is also true. An unfortunate consequence of this view is that if there is only one time a thing happens (say the first time B is true, A

is also true) it will receive of probability of 1.0.

This doesn't seem right. Other ways to play out this kind of view include moving counterfactually or talking about the preponderance of worlds in order to get defined probabilities. These are all interesting but the other views of probability are more probably the sort of thing that can have epistemic import for the individual worried about skepticism.

The subjective Bayesian view will take probabilities to mirror degrees of credence. The probability, for me, of p being true mirrors my degree of credence in p. This is no place for the antiskeptic to take refuge. When starting unjustified (with degrees of credence intuitively not being sufficient for justification), one will never *get* justified, however well one follows an updating mechanism.

On another view, named Keynesianism, probabilities are internal relations between propositions. The relation of making probable is like entailment in this way (except following the rules of probability). One might think of this kind of making probable as a relation which maxes out in entailment. A consequence of this view, as of the frequency view, will be that we should treat necessary truths as probability 1.0.[1]

But this should be slightly qualified. Necessary truths' probability is 1.0 on both a Keynesian and a frequency understanding of probability, but our access to the facts in question may vary. That is, on either way of understanding probability one will have to have access to the fact that a particular truth is a necessary truth to take advantage of this probability fact. That is, we might require that we treat any truth as contingently true, at best, unless one already has access to the fact that it is a necessary truth.[2]

---

[1]There will be worries about individuals' having access to such facts, but this is not a worry to consider here. In coming arguments, we will treat views that appeal to necessity as having an advantage without worrying about the fact that we might be imperfectly justified in believing the necessary truth.

[2]One might argue that, instead, we should ignore the necessary/contingent distinction and instead make all our arguments in terms of facts that are probable to 1.0 and facts which are probable to

5.1 Simplicity

Simplicity is a "virtue" with a very interesting history. The simplicity of a theory is almost always lauded as a virtue by its defenders, and, inevitably, the various ways in which the theory fails to be simple are cast aside as ways in which simplicity is not a virtue. As a simple case study, consider radical modal realism. Radical modal realists praise their view for its qualitative simplicity: it need not posit weird entities[3] to account for modality that other views do, and this is lauded as a virtue of the theory and a reason to think it is true. Further, when charged with failing horribly at being quantitatively simple its defenders will dismiss quantitative simplicity as among those kinds of simplicity that "don't count" for a view, or at least that the qualitative simplicity of the view just outweighs whatever other kinds of complexity the view may have. Views critical of modal realism will give precisely the same response in the opposite direction. This kind of discussion happens all over philosophy. Luckily many have taken this problem on and it is the function of the next portion of this dissertation to sort through the various views of simplicity and decide whether such virtues count for a view that has them (and, if possible, to discover why). First we shall attempt to characterize accurately each of the different kinds of simplicity.

5.1.1 Quantitative Simplicity

Quantitative simplicity is a standard form of simplicity that has been appealed to in the history of philosophy. It seems a fairly straightforward concept. A view or theory that appeals to a smaller number of entities is simpler. However, it's not

---

less than 1.0.

[3]There is one way of thinking about weird entities on which one might think that radical modal realism posits *the weirdest* kind of entities. That is, *possibilities* that are *out there* in the world that are *just as real* as entities in the world seems like a really weird kind of entity to posit. But the sense in which I'm saying it doesn't posit weird entities is that the possibilities are the same kind of entity as ordinary every day actual objects. There's no special kind of "possibility" object or property. In any case, nothing much hangs on this since it is presented as an illustration of the distinction in kinds of simplicity.

that straightforward if we want to avoid obvious objections to simplicity's status as an explanatory virtue, so let's consider this a bit more.

QuantSimp1:   A view or theory V is more quantitatively simple than another view or theory W iff V appeals to fewer *entities* than W.

According to **QuantSimp1**, any theory that appeals to fewer entities would meet the simplicity test, but this can't be what we mean by simplicity if we want it to be truth conducive. What about necessary truths? Does my view's requiring there be three necessary entities make it any less probable than someone else's view merely requiring there to be two necessary entities. Now, I'm not talking about a view which requires three entities and requires, further, that those entities be necessary. I'm talking about a view which only requires three entities which, it turns out, necessarily exist. I'm pumping the intuition that a view which appeals to the existence of three entities that are necessary isn't sticking its neck out any more than a view which appeals to two entities that necessarily exist, if that is all each view requires. Further, and even worse for **QuantSimp1**, it seems that a view which only needs to appeal to two *necessary* entities would be more probable than the view that only needs to appeal to *one* entity, if that entity is contingent. Likewise, in this case I'm not thinking of the view which appeals to two necessary truths as a view that requires two entities and requires that they're necessary entities in addition to the fact that they exist. Rather one theory requires two entities, which are both necessary. The point is that it seems, contrary to **QuantSimp1**, that in the case where one theory appeals to two entities and those entities are necessary one is in not in a worse position than a theory which appeals to one entity which is contingent. A fairly simple revision will deal with this problem:

QuantSimp2:   A view or theory V is more quantitatively simple than another view or theory W iff V appeals to fewer *contingent* entities than W.

This is a much better view of simplicity. But it seems that there are other things

which contribute to a view's complexity than simply the fact that the view appeals to more *entities*. One will want to account for the fact that, for example, a view which requires that there is at least one squirrel is simpler than a view which requires that there is at least one squirrel and that the squirrel is black. But there will be a fairly simple modification of the view which allows us to account for this intuitive requirement. We should move to talking about propositions instead of entities which will have other advantages in relating simplicity to probability. Entities don't get probability assignments. Propositions do. But it's easy to give a proposition that captures the commitment to some entity. One simply needs to make a quantified existence claim. That is, the requirement that Oranges exist can be captured by merely introducing the predicate O for "is an orange." The proposition representing the commitment to there being some oranges will then be: $(\exists x)Ox$. We can incorporate more possible contributors to complexity in our definition. Consider **QuantSimp3**:

QuantSimp3: A view or theory V is more quantitatively simple than another view or theory W iff V requires the truth of more *contingent* propositions than W.

This still isn't enough, as there is the problem of ballooning entailments that will make every view come out infinitely complex on this view. On this view of simplicity one could generate an infinite number of contingent propositions that are required to be true. Suppose that a view requires that ϕ be true. It will also require the following to be true: $\phi \vee \psi_1, \phi \vee \psi_2, ..., \phi \vee \psi_n$. All views, on **QuantSimp3** will be infinitely complex.

Perhaps **QuantSimp4** will be better:

QuantSimp4: A view or theory V is more quantitatively simple than another view or theory W iff V requires the truth of more *independent* contingent propositions than W.

This is yet another over-correction. Independence is too strict a requirement, as independence requires that the conditional probability be the same as the prior probability for that proposition. That is, **QuantSimp4**'s independence condition is

too strict. A is independent of B when $P(A/B) = P(A)$,[4] but a candidate that will more straightforwardly get out of this issue will simply require that the propositions are such that none entail each other. This is because dependent probabilities are still subject to the problem of diminishing probability. $P(A/B) * P(B)$ is still less than the probability of B as long as $P(A/B)$ is not 1. The question is how to word a candidate that will take advantage of this fact in an accurate and not stilted way.

QuantSimp5:   A view or theory V is less quantitatively simple than another view or theory W iff V requires the truth of more contingent propositions, none of which are entailed by the other propositions.

This is a much better candidate. Now all the contingent propositions that matter will be added on, while those that come for free with some other proposition will drop off. One might think that the entailment ballooning problem will continue, since V might require θ, but also require $\phi \vee \psi_1$ and $\phi \vee \psi_2$, etc. on to infinity. But none of these would entail each other. Now, this could be understood in two ways. On one way of filling out this case, the reasons for the ballooning of propositions of the form $\phi \vee \psi_n$ is that V requires ϕ. But on this way of filling out the case, all of the propositions of the form $\phi \vee \psi_n$ will be lopped off since they are entailed by ϕ. Suppose instead that the view requires an infinite set of disjunctions of the form $\phi \vee \psi_n$ but *doesn't* require ϕ. In this case we are left without a generator of the infinite disjunction, but also intuitively we are left with a theory V that is *hopelessly* complex, which will also have a hopelessly small probability.[5]

Another objection that might be considered is that there might be a (potentially) infinite number of propositions that aren't entailed by each other if we consider different ways of picking out individuals as different propositions. If we name an individual *Bob*, who has white hair, there is a proposition that ⟨Bob has white

---

[4]It follows trivially from this that $P(B/A) = P(B)$.

[5]At the least the view will reduce a little in probability for every one of the disjunctions it requires. But since the view requires an infinite number of disjunctions, the theory's probability will approach zero.

hair⟩. Now suppose Bob is called by an infinite number of names. One can call him $Bob, Bob_1, Bob_2, ..., Bob_n$ such that here will thus be n number of propositions true about Bob of the form ⟨Bob_n has white hair⟩. But arguably these propositions don't entail each other, since the proposition, e.g., ⟨Bob_{75} has white hair⟩ doesn't say anything about whether there is a Bob_{76} or if he has white hair. So here we have another case where the number of propositions in a theory will balloon infinitely (or at least arbitrarily largely ballooning). This ballooning problem is solvable in two ways, depending on what we take the various numerically indexed Bobs to be. If the various Bobs are taken to be legitimate names as we use them in ordinary language then it is not trivial to use the different names of the various Bobs. We would want to discharge these names with definite descriptions and, ultimately, it will not be a trivial theory that includes all of the 'Bob' names. If they are variables, we may simply treat co-referentials as not producing distinct propositions for our purposes.[6]

One might also object that, really, one shouldn't use propositions which are arguably truth-bearers to define a characteristic like simplicity, which is classically taken to depend rather on something in the realm of truth-makers, whether it be more facts, more properties, more individuals, or what have you. But it is important to realize that I am not taking this definition to be a fundamental account of simplicity. Rather I am attempting to come up with a proposal that will capture intuitive cases of simplicity but will also easily translate into a consideration of probability.

But perhaps, in the end, **QuantSimp5** is shades off of being correct. There is a problem of choosing which consequence propositions count as the ones to choose among the set of propositions with various entailment relations to each other. So, for

---

[6]There is also the question of which propositions to exclude when there are entailments. For example if theory V requires ϕ and it requires ψ it will also require $\phi \wedge \psi$. But should we exclude both of ϕ and ψ or should we exclude the conjunction? I think the answer is that we should always exclude the more compound proposition when there is a choice, as this will intuitively match the way we count better. Requiring both ϕ and ψ will add two units of complexity and not just one. We would need to build this into a more precise formulation of **QuantSimp5**, but for present purposes it is sufficient to make this parenthetical remark.

example, we need to pick a set of propositions required by V but that don't entail each other. Since V requires both p and q, we need to choose whether we admit $p \wedge q$ or both p and q. The more atomic propositions p and q should be the ones we choose as it will more closely correlate with our intuitive notion of simplicity. It seems odd at the very least that in this case V, in virtue of entailing itself, is one of the "required" propositions (which needs to be the case in order to get rid of the infinitary problem of disjoining a bunch of stuff with V) and will then be removed with a stroke of the pen after being found to be entailed by the rest of the requirements.[7] There is no real decision procedure here for deciding what gets included and what doesn't. We can solve this with the following proposal.

QuantSimp6: A theory V is more complex than U iff the set of the *least-complex-constituents* of V is larger than the set of the *least-complex-constituents* of U.

Set of the *least-complex-constituents* of some theory V =df. the set of contingent constituents of a theory (perhaps propositions, perhaps facts or beliefs) that is formed by first excluding all necessary truths and collecting all of the atomic propositions that are constituents of the theory,[8] then adding all of the more complex constituents, excluding any member entailed by the conjunction of previous (included or excluded) members, by slowly growing in complexity as one adds members. The set is complete when one has added or excluded all constituents of V.[9]

QuantSimp6 simply makes explicit the sort of decision procedure I had in mind

---

[7]One can very easily and perhaps endlessly do this with almost any theory. If V requires p and requires q, then trivially one will get $p \equiv q$. But then all three jointly entail and then don't contribute to simplicity (or one will have to not include p or not include q which would be an untoward consequence of the view.

[8]If $A\&B$ is required by a theory, then A and B are both treated as constituents of the theory.

[9]This criterion will clearly work for a propositional system, but there's no clear way of capturing this in a straightforward way with respect to systems with quantifiers. Supposing a theory requires merely that there exists something with the properties $P\&Q$, it will be difficult to capture this commitment with the above requirement. There is one very contentious way of thinking about systems with quantifiers that may work. One might take any case of a quantification and collect all the entities in the domain of the quantifiers and attempt a translation of existentially or universally quantified statements into complex disjunctions and conjunctions respectively. One major problem with this is that, in an infinite domain, every universal quantification will be infinitely complex. Another way one might treat the case is to treat every quantified statement as atomic for our purposes. This will, perhaps, not capture every way that a theory may be simple, but it will at least give us a way to apply this principle in ordinary cases. the drawback of this route is that it will all a view to be incredibly simple by being expressed as a long conjunction within a superfluous quantifier. In any case, the above, though not universally applicable, is at least progress toward a view that can technically characterize simplicity in a clear way.

for collecting the propositions that matter in deciding on the simplicity of a theory. One might worry about **QuantSimp6** that it is getting further and further from an everyday notion of simplicity. But it doesn't need to exactly mirror such a notion if it is a virtue that we can recognize and apply it fruitfully to theories.

One worry regarding **QuantSimp6** would be that there are propositions or entities that contribute to a view's complexity that are in danger of being left out. The challenge for **QuantSimp6** will be to make sure that unintended requirements of a theory get counted against simplicity. For example, Vogel points out that a fact about how skeptical scenarios will have to explain certain regularities in experience will contribute to skeptical scenarios' complexity despite the fact that no skeptic will have included this as part of the skeptical scenario. We'll discuss this specific case more in chapter 6, but Vogel's assertion is that the real world hypothesis employs necessary truths about geometry as part of the explanation for our experiences' obedience to things like the triangle inequality. Further, he asserts, the skeptical hypothesis can't appeal to such necessary truths in explaining these aspects of human experience. Thus the skeptical hypothesis must add some contingent truth in order to so account, and thus will be a less simple theory than the real world hypothesis. The problem with applying **QuantSimp6** is that the skeptic hasn't included this in his theory until Vogel makes his argument. However, it seems to be required by the skeptic's theory. We don't want **QuantSimp6** to exclude such consequences of a theory as contributing to the complexity of a theory.

There is a fairly straightforward response that doesn't require any fingernail biting about whether unintended consequences get included. One may simply treat the two as different theories. The one excluding the additional hypothesis will be simpler but (perhaps) less probable in some other way. The theory *with* the additional hypothesis will perhaps improve in other ways while being the less simple theory. The theory with the unintended consequences can then, by **QuantSimp6**, be a more complex view

that explains more, while the theory that doesn't include the unintended consequences will be a simpler view that, perhaps, explains less.

The Probability Question

The next question to consider is whether **QuantSimp6** is a virtue and will be vindicated probabilistically. We will take some feature of a view to be a virtue if *ceteris paribus* a view or theory with the virtue is more probable than a view or theory without it.[10] This is clearly true for **QuantSimp6**. Let's say we have evidence E, and it is roughly equally well accounted for by two views. Now suppose that one view is simpler than the other by just one proposition, and that all other things about the cases will be the same. It follows that the more simple view or theory will be more probable than the less simple one by diminishing probabilities. The probability calculus for conjunction is $P(A \wedge B) = P(A) * P(B/A)$. Thus as long as the propositions aren't necessary and as long as the probability of B given A isn't 1.0 (both ensured in **QuantSimp6**), the probability of more complex theories will be lower, assuming a principle of indifference.

Now, suppose we are presented with two theories that explain the same data and all we know about them is that one is more simple than the other. Should we prefer the simpler theory to the more complex one? This question is difficult to the extent that we find it difficult to accept a principle of indifference. If we accept the principle of indifference, then we can answer yes, but the principle of indifference is notoriously (and controversially) subject to Bertrand's paradox. Salmon gives a case to illustrate the problem as follows:

> Suppose a car has traversed a distance of 1 mile, and we know that the time take was between one and two minutes, but we know nothing further about it. Applying the principle of indifference, we conclude that there is a probability of $\frac{1}{2}$ that the time taken was in the range of 1 to $1\frac{1}{2}$

---

[10]This ceteris paribus condition is meant to allow that what are called virtues may, perhaps easily, be undermined by other considerations, perhaps easily.

minutes, and a probability of $\frac{1}{2}$ that the time taken was in the range of $1\frac{1}{2}$ and 2 minutes. A logically equivalent way of expressing our knowledge is to say that the car covered the distance at an average speed between 30 and 60 miles per hour. Applying the principle of indifference again, we conclude that there is a probability of $\frac{1}{2}$ that the average speed was between 30 and 45 miles per hour, and a probability of $\frac{1}{2}$ that the average speed was between 45 and 60 miles per hour. Unfortunately we have just been guilty of self-contradiction. A time of $1\frac{1}{2}$ minutes for a distance of one mile is an average speed of 40, not 45, miles per hour. On the basis of the same information, formulated in different but equivalent terms, we get the result that there is a probability of $\frac{1}{2}$ that the average speed is between 30 and 40 miles per hour, and also that there is a probability of $\frac{1}{2}$ that the average speed is between 30 and 45 miles per hour. Since it is not impossible that the average speed is between 40 and 45 miles per hour, the foregoing results are mutually incompatible.(Salmon, 1974, 94)

The paradox seems to show that applying the principle of indifference will result in contradiction. Now, there are arguments that attempt to rescue the principle from contradiction, but I won't take a strong position on whether one or any of the responses successfully answers Bertrand's paradox.[11]

There is another less technical problem for the indifference principle and that is the question of why one would think the indifference principle should be accepted.[12] There is another seemingly rational approach one might take to assigning probabilities. One might simply *refrain* from applying prior probability where one doesn't have some good reason to apply them in any particular case. That is, one may simply decide not to apply probabilities to the set.

An initial response to this problem comes from our own practice in probability assignment. There are all sorts of cases where we seem to apply a principle of indif-

---

[11]See, for example, Shackel (2007), Jaynes (1973), or Marinoff (1994).

[12]I'm looking for something beyond a mere pragmatic argument. That is, the argument that it would sure help us justify belief in more propositions isn't the kind of thing I'm looking for. At the very least, barring a reason to accept it as a rational principle, establishing that there are clear cases where our ordinary (presumably) probability judgments rely on a principle of indifference will at least help.

ference and we also take ourselves to be rationally acting in doing so. For example, when we roll a six-sided die we take each of the six possible outcomes as equally probable unless we get further information that the die is loaded or something else gives us reason to apply the probability weights differently. We do the same with the flipping of quarters and with the assignment of probabilities in roulette and standard decks of cards.

So how can such practice in probability assignments be considered rational if some kind of principle of indifference is *not* justified? If we consider how we pick which alternatives will fall under a principle of indifference in the most plausible cases of its application, we will see that there is probably more going on in such ordinary cases. Consider the case of the six-sided die roll. In a six-sided die roll, we automatically think of it as a case where there are six and just six possibilities. The die could come up as a 1, 2, 3, 4, 5, or 6. But why should we take this to be the case? Aren't there a near infinite number of possible orientations of a cube in three-dimensional space that don't involve one flat-side up? It seems that if we really applied a principle of indifference we would either assign a very low probability to any one number coming up or quasi-arbitrarily defining wide sets of orientations as one, two, etc. So why don't we apply a real principle of indifference? One interesting and plausible answer is that in all these apparent cases of the principle of indifference, our common sense (in the Quinean sense) physics judgments are what give us our choices in these cases as well as our decision to assign specific probabilities to each choice. The fact that we dismiss a near infinite number of other possibilities as not worth considering is further support that our ordinary physics judgments are affecting our assignment of probabilities here. We dismiss all but the smallest number of possible orientations of the die because we take the probability that any of those possibilities actually occur is very very small. But what do we apply to get a roughly equal probability for each of the numbers one through six?

Assumptions of uniformity (along with, perhaps inductive argument and naïve physics) get us from settling which possibilities are to be taken seriously to giving roughly the same probability weight to each of the six sides of the die coming up. This is at least a plausible story for how we come to the same conclusion that a principle of indifference would have us come to. This leaves us in a relatively agnostic place with respect to whether merely being simple (and given no other information) gives one reason (alone) to prefer one view over another. However, perhaps getting probability with *ceteris paribus* conditions is enough.

The Access Problem

Another concern for settling on **QuantSimp6** is that, though the probability relevance is fairly clear and simple, this kind of simplicity may not be as clear and easy to pick up on as some of the earlier candidates. One needs much more (and much more sophisticated) access to probability facts and relations among propositions as well as a very extensive knowledge of the consequences of a view that it seems difficult to say that people have in ordinary contexts often correctly and fruitfully applied the simplicity criterion to a case.

We're left with something of a dilemma for the virtue of simplicity. Either the virtue of simplicity is clearly a virtue and is not very reasonably thought to be the sort of thing less sophisticated subjects can successfully apply *or* the virtue of simplicity is not so obviously a virtue and is the sort of thing that is easier to successfully apply (e.g., **QuantSimp1**). Until such time as it proves problematic for our purposes, we will stick with the more clearly virtuous reading **QuantSimp6**.

There is also the problem of giving necessary truths for free when the epistemic subject may not be in a position to know that each necessary truth is necessary. Without having such access, the individual will not be in a position to know, with respect to each necessary truth that enters into a view, that it comes for free. But the error that this problem will cause is minimal. One may simply require that, unless one

is justified to 1.0 that some necessary truth is true (and necessary), one includes the proposition in one's collection of propositions according to **QuantSimp6**. Another strategy for response that would require modification is, rather than exclude necessary truths, to exclude all propositions of probability 1.0 on one's evidence. I'm not sure which way to go on this one, but either route is acceptable.

Two Other Objections

There are two other objections that we might consider to **QuantSimp6**. One consequence that initially seems problematic is that, it seems, the null hypothesis (that there is no explanation of the observation in question, or even that there is nothing at all) will always win on simplicity grounds. So, if we have an experience as of a teacup, on **QuantSimp6**, the hypothesis that there is no explanation for this will have a huge leg up by being so simple. But this is untoward in the final analysis, since, presumably, all other views will have other virtues that work in their favor (and that the virtue only establishes an advantage given other things are equal).

Finally, another infinite ballooning threat may raise its head. Suppose that we have the fairly simple explanation of my pen failing to write which claims that is out of ink. Now, how does one pick out all the propositions which constitute this hypothesis? Well, it seems this theory requires that it's not the case that the pen was caused to fail to write by an invisible anti-writing alien device. It also seems to require the falsity of all the other explanation candidates for the observation. But these aren't strictly required. It is metaphysically possible that my pen failing to make marks when moved along the page is explained both by the alien device and by the pen being out of ink. The possibility[13] of true over-determination allows us to not build the negation of "opposing" explanations into what's required by our account.

---

[13]or at least the fact that it shouldn't be excluded from consideration *a priori*.

5.1.2   Qualitative Simplicity

Qualitative simplicity is much harder to characterize in a way that is clearly virtuous and yet is clearly the virtue that is used in philosophical argumentation. A main line of criticism comes from Sober (2002) and Huemer (2009). Sober argues that, since the best way we have to understand the virtue of simplicity is by reference to mathematical simplicity, only in those cases that a view or theory that is qualitatively simple is also mathematically simple will qualitative simplicity be a virtue of an explanation. An answer to this argument would simply consist in another way to ground qualitative simplicity being a probability increaser.

However, there's another kind of objection to qualitative simplicity being a virtue which comes from the fact that it is also not clear that any reasons to think that simpler explanations are more probable when quantitatively simple apply to qualitatively simple explanations (or at least those cases to which we appeal in philosophy).

Huemer attempts to apply the various avenues he pursues for a probabilistic vindication of simplicity and finds that each of them fails to vindicate parsimony-based arguments for physicalism or for nominalism. Part of his argument centers around questioning whether such views really are parsimonious. Now, ultimately many of the criticisms center around showing that, in the end, having the intuitive "qualitative simplicity" really ended up saying nothing about the probability of the views in question (at least if we wish the probability of these views to be increased in virtue of the view's being simple.)(Huemer, 2009, 231-233)

I think that there is a way to think about the cases where qualitative or type simplicity is used in philosophical argumentation and understand it as a virtue. The only concession will be that perhaps it will turn out in the final analysis that this virtue ought to take another name that isn't 'simplicity.'

Consider the case of idealism and physicalism. These views both laud themselves as being qualitatively simpler than the various dualistic theories and thus more prob-

ableother things being equal. But the reasoning, I think, is not so simple.

Let's say I own a bunch of cats. Let's also say in my experience (as far as I know) there are only two kinds of animals in the world, human beings and cats. Now suppose that a pile of excrement shows up on my lawn. Knowing that cats exist along with the fact that cats could explain a pile of excrement's existing on a lawn counts for the "cat pooped on my lawn" explanation. Now, suppose I have another theory, which asserts that a dog pooped on my lawn. It seems that I should clearly prefer the cat thesis. Now this might be called simplicity, but it's more fundamentally about using what one already has in one's metaphysics to explain (if it can be done).

But why should we think this is probability conducive? The key here may be to capture this kind of intuition: "having reason to believe that at least one instance of a type exists raises the probability that other instances of that type exist, ceteris paribus." That is, if, on one's background knowledge, there's already good reason to believe that there is at least some individual of kind K, then the probability of other instances of kind K existing is raised. This applies to cases like the poop case and the physicalism or idealism case. The idealist hypothesis (for someone already justified in believing there are minds) will be more likely than the normal external physical world hypothesis, given two hypotheses that are roughly the same prior probability and provide roughly the same quality of explanation of one's experience. The idealist has reason to prefer the idealist explanation since it appeals only to entities the existence of which one already has reason to believe exist. I, in the cat/human world would be more justified in appealing to some extra cat pooping on my lawn than appealing to some extra dog.

Is it true that "having reason to believe at least one instance of a type exists raises the probability that other instances of that type exist"? This can be proven (within limits) from Bayes theorem.

The virtue of qualitative simplicity really is in appealing only to theories (and

entities) instances of the type of which one already has reason to exist.[14] This can be shown as probability-raising from Bayes theorem. Bayes' theorem states that $P(A/B) = \frac{P(B/A)P(A)}{P(B)}$. Now cases where qualitative simplicity often come up in philosophy are where one is appealing to an explanation that involves the existence of entities of certain kinds. For example, one is explaining sense experience by appealing to minds. But such an explanandum entails that there are some minds. But if the explanandum entails that there are some minds then, if we stipulate that by A we mean this mental explanandum and that by B we mean that there are some minds, the $P(B/A)$ term will drop out so that now $P(A/B) = \frac{P(A)}{P(B)}$ But then for any consistent value of the probabilities of A and of B where $0 < P(A) < 1$ and $0 < P(B) < 1$, The Probability of A on B will be greater than the prior probability of A: $P(A/B) > P(A)$. The probability of the idealist response is increased because it only appeals to entities for which we already have fulfilled some prerequisites. But since, on these simple theories we know B, we go with the updated higher probability for A. So, ceteris paribus, the theory that has this kind of qualitative simplicity will be more probable.

5.1.3   The Explanatory Simplicities

Kevin McCain, James Beebe, and William Lycan propose forms of what is called "explanatory simplicity." McCain proposes the following Explanatory Simplicities:

Explanatory Simplicity: All else being equal, an explanation that posits fewer fundamental explanatory regularities is preferable to an explanation that posits more.

Explanatory Questions: All else being equal, an explanation that raises fewer unanswerable explanatory questions is preferable to an explanation that raises more. (McCain, 2014)

Beebe gives three similar proposals:

Explanatory Simplicity I: Other things being equal, a theory whose structure is more elegant or straightforward should be preferred to a theory that is less elegant

---

[14]We will call this *having precedent* for a type of entity when discussing this kind of case below.

or straightforward.

Explanatory Simplicity II: Other things being equal, a theory that raises fewer further explanatory questions should be preferred to a theory that raises more.[15]

Explanatory Simplicity III: Other things being equal, a theory that posits fewer primitive explanatory notions should be preferred to one that posits more.[16](Beebe, 2009)

Explanatory Simplicity I will be similar to what we'll call mathematical simplicity, so we'll forget about this proposal in the next section. Due to the similarity between **Explanatory Simplicity III** and McCain's **Explanatory Simplicity**, and between **Explanatory Simplicity II** and McCain's **Explanatory Questions**, in this selection we will consider them as roughly equivalent and consider McCain's **Explanatory Simplicity** and **Explanatory Questions**.

Not much has been said regarding either of these explanatory simplicities with respect to vindicating them as virtuous, but it will be worth giving it a go.

Explanatory Simplicity seems to straightforwardly incorporate itself into **QuantSimp6**, and thus true. That is, if these are meant to be contingent regularities (as McCain seems to say later in the same passage), then cases of violating **Explanatory Simplicity** would also be cases of violating **QuantSimp6**. This proposal is a good one, though perhaps unnecessarily narrow in the presence of **QuantSimp6**.

Explanatory Questions is a little harder to vindicate as a virtue. It's also hard to see it as particularly suited to be called a version of simplicity. But let's assume that it can be called simplicity and do our best to consider how this can be vindicated as a virtue. It will be helpful to start from the beginning. One first needs to distinguish between ways that one might call a question "unanswerable." A question may be unanswerable in three ways. (i) A question may be incapable of being given *any* answer by us that meet even the most minimal of rationality requirements, (ii) a question may be incapable of being given any *true* answer by us, and (iii) a question

---

[15] Beebe (2009) attributes this to Huemer and Lycan.
[16] Beebe (2009) attributes this to Lycan.

may be incapable of being given an answer which we have good reason to believe is true. It seems clear from the context that McCain takes (iii) to be the appropriate way of understanding **Explanatory Questions**.

It is difficult to consider the value of this proposed virtue because it's difficult to come up with cases where this principle is clearly applicable outside of the debate over external world skepticism to which McCain applies the principle. Let's see if we can understand this virtue. The first problem here is that this virtue seems to either be a form of consilience[17] or, perhaps, merely a pragmatic virtue. That is, one way of thinking about **Explanatory Questions** construes it as a virtue that prefers views which explain a lot of different things (which will mirror consilience), making those views that get rid of the most questions epistemically good. But this doesn't quite seem the kind of thing that McCain and Beebe are getting at. The idea here is that views which raise fewer *further* unanswerable questions are more probable, *ceteris paribus.* The virtue of **explanatory questions**, it seems, is almost certainly a pragmatic virtue. Explanations that don't raise further questions are explanations according to which we can be satisfied with ceasing in inquiry so that we may *act* in the world. If we are to pose an explanation that raises further unanswerable questions we may end up paralyzed by such things and never stop the inquisitive process. Raising an unanswerable question (and not using it as a mark against the view) might put us in the situation of the robot in science fiction when posed unanswerable puzzles. We would cease to function and black smoke would start pouring out of our ears.

This being right doesn't disqualify **explanatory questions** from being an explanatory virtue. But there may be reason to doubt that this is a useful measure. It seems that *any* theory will raise a very large number of questions that are unanswerable for us. Take, for example, the theory that I exist. An incredibly large list

---

[17]We will discuss consilience more below, but it should suffice to say, here, that consilience privileges explanations which can bring together more truths that are apparently separate to ones which do not do so.

of questions will come up that cannot be answered. We can generate unanswerable questions quite easily from this proposal. From the proposal that I exist, the question arises why I exist rather than Pete, or why I exist rather than $Pete_1$, or why I exist rather than $Pete_3$, . . ., or why I exist rather than an extra stapler in the world, etc. We can ask contrastive explanation questions until the cows come home which are not answerable to us in sense (iii). This may end up being an infinite list for many views that we take to be justified (if we can ask contrastive questions by varying features that fall on a continuum, then we should be able to do so). For example, the question of why my arm is up in the air will give us questions like ⟨Why is my arm up in the air in this position rather than that?⟩ But if arm position is a property that varies on a continuum, then we can construct an infinite number of explanatory questions that are intuitively unanswerable by us in sense (iii) of unanswerable. Now, if everything raises so many unanswerable questions, then how can this function as an *epistemic* virtue (or vice, as in this case)?

McCain might object by distinguishing questions that are *in fact* unanswerable(in sense (iii)) and questions that are *in principle* unanswerable (i.e., in sense (i)). Perhaps, for example, the question whether there are angels may be *in fact* unanswerable (at least definitively), but the question of how many angels can fit on the head of a pin is, arguably, unanswerable in principle. It, at least, seems unanswerable in any way that meets minimal rationality requirements. The contrastive questions in the above enormous set of questions seem answerable in principle, but what we want from explanations is that they don't raise questions that are unanswerable in the first sense *in principle*. But if this is the way to read the virtue, then it merely seems to be asking that we don't accept views that are incoherent, which is acceptable, but not a virtue of simplicity. This does, however, clearly count against such a view (an inconsistent view has probability zero). Perhaps there are other questions that are unanswerable in principle that don't imply incoherence. Certain interpretations

of quantum mechanics intuitively seem to yield questions that are unanswerable in principle. Knowing the momentum of a photon makes the question of its position unanswerable in principle on the standard Copenhagen interpretation of quantum mechanics. But this kind of question is still much more rare than is useful in theory choice.

5.1.4 Mathematical Simplicity

Mathematical Simplicity is a little more technical notion having to do with curve fit. That is, it has to do with what *kind* of equation we use to attempt to get a curve that best fits a set of data. For example, we may have data that plot the age of death of individuals against their percentage of body fat. Now, the relationship between the individuals in this set of data may be linear, parabolic, hyperbolic, logarithmic, and so on. Mathematical simplicity governs which kind of curve we should prefer when two different curves predict the same data. The short way of describing mathematical simplicity is that the model that has the fewest adjustable parameters is simpler. So, for example, suppose you have two kinds of model that are consistent with your data, each positing a relationship between x and y. One model describes the relationship as linear ($y = ax + b$), while another describes it as parametric ($y = ax^2 + bx + c$). The linear model here is simpler, and more probably true. In a moment we shall show how this works, but, for now consider this example from Jefferys and Berger:

> Suppose a Friend who has a reputation as a prankster offers to flip a coin to decide who will perform a little chore: heads he wins, tails he loses. Knowing your friend's reputation, you might well be concerned that he would use trickery (perhaps a two-headed coin) to win the toss. The hypothesis H_{HH} that the coin has two heads is, under this understanding, a simpler one than the hypothesis H_{HT} that the coin is fair. In a series of many coin tosses, H_{HH} will be falsified if tails comes up even once, whereas any sequence of heads and tails could arise from H_{HT}.
>
> Before the coin is flipped, you might believe that the hypotheses H_{HH} and H_{HT} are equally likely. Then the coin is tossed, and it indeed comes up

heads. Your degree of belief in the two hypotheses will change as a result of this information, and (by Bayes' theorem) the posterior probability that you assign to H_{HH} should now be twice what you assign to H_{HT}. Still, the evidence that your friend is trying to fool you is not very strong at this point, perhaps not strong enough to challenge him for a close look at the coin. On the other hand, if the coin comes up heads on five occasions in a row, you will be rather inclined to think that your friend is playing a joke on you. Even though both hypotheses remain consistent with the data, the simpler one is now considerably more credible.(Jefferys and Berger, 1992, 68)

This doesn't quite straightforwardly yield neat equations where it's easy to see how there are more adjustable parameters in the fair coin case versus the two-headed coin case, but it will be a fairly intuitive illustration. In the above case H_{HH} has zero adjustable parameters. The prediction is just that heads will come up every time. However, the hypothesis H_{HT} will have at least one adjustable parameter, since the coin can take both the value of heads or of tails.

We can also use the coin case to illustrate how this view generates more probable views. Now, models with fewer adjustable parameters, of which H_{HH} is an extreme case, will be harder to square with any random or arbitrary set of data. The first time tails comes up, H_{HH} is completely falsified. However, the view with more adjustable parameters (for our purpose, the fair coin model) is easy to square with nearly any data set. Now, this difference between models makes it so that, with each additional piece of data that remains consistent with H_{HH} (the hypothesis with zero adjustable parameters that the coin is two-headed), the data will confirm H_{HH} much more strongly than it does H_{HT} (the hypothesis with one adjustable parameter that the coin has both a heads side and a tails side).

That is, since the higher the number of adjustable parameters the easier it is to square the model with any data set, it matters for complex theories a lot less if some prediction comes true than it would for a simpler theory. Huemer explains as follows:

If we have just three data points, [a parabolic curve] is guaranteed to fit the data to any desired degree of accuracy: for any three points in a plane, there exists a parabola passing arbitrarily close to them. [A linear curve], however, is not guaranteed to fit the data; there are infinitely many more triples of points that are non-collinear than triples that are collinear. For this reason if [a theory asserting the data are related in a strictly linear way] is false [i.e., the data don't vary linearly], the probability that [a linear curve] would accommodate the data perfectly is zero... A similar point holds more generally, for larger data sets and equations with more adjustable parameters: the more parameters an equation has, the wider the range of sets of points that can be made to fit, to a given degree of accuracy, an equation of that form.(Huemer, 2009)

It should be intuitively plausible that a model being harder to square with arbitrary data will make the model more strongly confirmed by evidence consistent with the model. The probability of the new evidence on the model will be higher. Thus the posterior probability of the simpler will go up more than the posterior probability of the complex model when adjusting one's belief by Bayes theorem: $P(M/E) = \frac{P(M)*P(E/M)}{P(E)}$. As the probability of the particular evidence on the model goes up, the posterior probability of the model on the particular evidence goes up.

Often philosophers, such as Huemer (2009) and Sober (2002), will attempt to use mathematical simplicity as a model by which to understand the use of simplicity as a virtue in both quantitative and qualitative contexts. Huemer takes this mathematical model (and a few other models) of simplicity and applies it to debates like nominalism/realism and physicalism/dualism. He argues that the assertion that nominalism (and physicalism) are qualitatively more simple fails to have mathematical import in a way that will make mathematical simplicity applicable. After showing that the mathematical model (as well as a quantitative model) fails to vindicate such qualitative simplicity arguments, he suggests "that many philosophers' 'taste for desert landscapes' is indeed an aesthetic rather than an epistemic preference. As we have shown above, there is another interesting way to think about qualitative simplicity,

so Huemer's rejection of monistic arguments can be dodged. This will simply require that we reject the attempt to subsume other kinds of simplicity under the umbrella of mathematical simplicity. Mathematical simplicity may have its own uses and applications, but I think there are cases where quantitative and qualitative simplicity cannot be captured by appeal to mathematical simplicity. The different kinds of simplicity are to be probabilistically vindicated in different ways. We have shown a strategy for this vindication in the above sections.

Consider, for example, the appeal to simplicity in physicalism (or idealism) versus dualism. I think that Huemer has successfully argued that the mathematical notion of simplicity doesn't help justify one over the other (where dualism will have adjustable parameters with respect to the mental laws, physicalism will make up for such adjustable parameters with the various adjustable parameters roughly regarding which physical or functional states the mental supervene on, etc.).(Huemer, 2009) However, I think it is yet important to understand that this is not what is going on in appeals to qualitative simplicity and especially in cases like the debate between physicalism and dualism(see above for discussion)

I'm not convinced that all cases of quantitative simplicity can be captured intuitively by the mathematical account either. Views with completely superfluous complexities should be judged lacking in simplicity as well as views where the complexity involves something substantive. Suppose I have two different explanations for my teacup having only a little tea left in it. On one proposal, the teacup is nearly empty because I drank most of it. On a second proposal, the teacup is nearly empty because I drank most of it and I'm a foreign spy. The second proposal should be rejected as being less simple despite the fact that the adjustable parameters each proposal will generate are, it seems, the same. Both have zero adjustable parameters. Consider, again, the case of the two-headed coin above. The hypothesis that the coin has two sides that read heads has zero adjustable parameters. Predictions

cannot vary on any axis. But suppose I simply tack on a useless extra proposition to the two-headed coin hypothesis above. It becomes less simple quantitatively, but it doesn't necessarily change the number of adjustable parameters in the two-headed coin case. The mathematical model cannot intuitively capture quantitative simplicity. For these reasons, I think we should reject mathematical simplicity as a proposal for a unifying account of simplicity, while admitting that mathematical simplicity is indeed a virtue.

5.2 Consilience

5.2.1 The Nature of Consilience

Consilience is the name we give to views and theories that are able to take vast disparate data and make sense of them. This is sometimes called unification as well. The idea is that if a view can take various independent pieces of data and tie them together in a whole that makes sense,[18] the view is *ceteris paribus* more probable than another view which doesn't do so. For example, the belief that there is no Santa Claus brings together the fact that my friends seem not to believe in him with the fact that my parents act suspiciously whenever questions come up and the fact that the presents people get seem to mirror their socioeconomic status. Thus the hypothesis that there is no Santa Claus has a high degree of consilience.

5.2.2 Consilience and Justification

Tim McGrew has an influential account of of consilience. He argues:

> The proof is straightforward. Assume for the sake of the argument that $P(H_1) = P(H_2)$, that for all n, $P(E_n/H_1) = P(E_n/H_2)$, and that the various E_n are positively relevant to each other conditional on H_1 but independent of each other on H_2, i.e.,

---

[18]To put this probabilistically, a more consilient view takes propositions that don't seem to be positively correlated and correlates them positively.

$$P(E_1 \& ... \& E_n / H_1) > P(E_1 / H_1) \times ... \times P(E_n / H_1)$$

but

$$P(E_1 \& ... \& E_n / H_2) = P(E_1 / H_2) \times ... \times P(E_n / H_2)$$

Then by Bayes' Theorem and a bit of trivial algebra,

$$P(H_1 / E_1 \& ... \& E_n) > P(H_2 / E_1 \& ... \& E_n)$$

Thus H_1 emerges as clearly superior to H_2, in straightforward confirmational terms, despite the fact that on a case-by-case basis it has no predictive advantage over H_2. The result is easily generalized to yield the theorem in question. This provides a convincing demonstration of the confirmational relevance of what we will call 'theoretical consilience'—the consilience that obtains when a hypothesis or theory reduces independence among the data.

Tim McGrew shows that on this way of understanding consilience, it is definitely a virtue. But there are challenges to McGrew's proposal (and proposals that give roughly the same view).[19] One challenge for McGrew is that our ordinary way of thinking about consilience only roughly moves us toward this probabilistic fact. At best the ordinary cases where we apply the virtue will approximate the technical apparatus that McGrew has given. For example, that there is no Santa brings together many disparate facts like my parents' acting suspiciously around Christmas and Christmas gifts' mirroring of children's socioeconomic class, but we don't have that these were necessarily *independent* on the opposing theory. But at the very least we can see that there is probability theory which stands a chance of vindicating this virtue.

---

[19]For further discussion of consilience and McGrew's account thereof see, for example, Lange (2004), Myrvold (2003), and Schupbach (2005)

5.3 Explanatory Power

5.3.1 The Nature of Explanatory Power

Explanatory power is another virtue of explanation. On some ways of thinking of explanatory power, it should be thought of as something like Aristotle's "eudaimonia", a virtue covering all the rest of the virtues. That is, explanatory power, on this way of understanding it, shouldn't be thought of as a separate proposal for *an* explanatory virtue, but rather as the virtue that an explanation has when it is the better explanation. However, instead I think it is interesting to consider two virtues that I think are properly categorized under the heading "explanatory power". One is called explanatory breadth and the other is called explanatory depth.

Explanatory Breadth

Breadth roughly refers to a view's explaining a larger number of phenomena. This is to be distinguished from explanatory depth, where a view may explain very little but it does so in a deep way (whatever that means). A view has more explanatory breadth than another when the explanandum includes a larger set of data than in the less broad explanation. This is supposed to make broader explanations preferable. But how is this to be cashed out in terms that will make it clear that a view which has more explanatory breadth is more probable, *ceteris paribus*? Well, other things being equal, one more proposition that is positively confirming for a hypothesis as compared to another hypothesis makes it more probable. This is to be distinguished from consilience in that consilience holds constant the number of pieces of evidence that the theory explains. Breadth is just adding another fact to the list of pieces of evidence explained by the theory.

Explanatory Depth

Explanatory Depth is a little harder to nail down. On one way of thinking about what the expression "explanatory depth" means, a view is explanatorily deeper if it explains what it explains *better*. But if this is true, then explanatory depth will just be, as described above for explanatory power, the *eudaimonia* of explanatory virtues. Is there another way to make depth a virtue? Explanatory depth could, rather, mean something more on the order of explaining the individual pieces of evidence in a way that makes them more probable. This will again yield a higher probability of H on one's evidence *ceteris paribus*, by some simple Bayesian calculation. This, however, becomes a rather simple virtue that appeals to nothing other than higher conditional probabilities. This will force us to stretch the meaning of "depth", perhaps to the breaking point. This also makes views on which a hypothesis *entails* some evidence, even in a trivial sort of way, have the very *highest* explanatory depth with respect to that evidence. This can't be right.

Thagard (2007) proposes a view of explanatory depth that sounds like it roughly mirrors (and will perhaps illuminate) McCain's **Explanatory Questions**. He says,

> a deeper explanation for an explanatory mechanism M1 is a more fundamental mechanism M2 that explains how and why M1 works. M1 consists of parts, and M2 describes parts of those parts whose properties and relations change in ways that generate the changes in the properties and relations of the parts in M1.(Thagard, 2007, 38-39)

An explanation is *deeper* if it proposes more further explanations for the aspects of that explanation. This is meant to make an explanation more probable. It seems difficult, initially, to understand how this is the case. One first might wonder whether there is some *a priori* reason to think that for any explanatory mechanism, there is some further deeper explanatory mechanism. There are a couple ways that such a case might work.

Suppose that we know that the world has 5 explanatory levels. That is, suppose that a world contains five different kinds of mechanism, each one a more basic kind of mechanism than the last.[20] Now suppose, also, that we pick an explanation from the 5 leveled pool of explanations at random (say we write them on playing cards and draw from those, if we're worried about a principle of indifference). The probability, assuming there are roughly the same number of explanations on each level, of picking an explanation from the bottom level will be $\frac{1}{5}$. The probability of picking an explanation from the bottom *two* levels is $\frac{2}{5}$. This means that for any explanation in this world we've constructed, it is more likely than not (if we've picked an explanation randomly) that any explanation will have at least two more fundamental explanations out there. This seems to indicate that an individual explanation is probably not the deepest level explanation that one can give for a phenomenon (supposing we don't have independent knowledge of which level our explanation is on).

Now suppose that we don't know that there are 5 explanatory levels. Suppose the number of levels of explanations in the world is a random number between one and infinity. The *a priori* probability that one gives the bottom level explanation is nearly infinitesimal(supposing we have no independent information about how much deeper an explanation we can get), and, the more levels of explanation one gives the more probable it is that one has gotten the true explanation. The problem is that a deeper explanation often will be probabilistically relevant to much more about the world than a less deep explanation. But this means that the prior probability of deeper explanations should be smaller than less deep explanations, since deeper explanations will need to square with more data about the world.

However, one way to rescue this virtue is to make it parasitic on consilience or unification. An explanation which is deeper will usually be more consilient, i.e., it

---

[20]For example, suppose we proposed an explanation of my arm moving in terms of the macro-level contraction of muscles pulling a bone with a tendon and give a different explanation which appealed to molecules and the forces that molecules exert on one another. In such a case the explanation appealing to molecules and their forces on one another would be a deeper explanation.

will bring together more disparate data, than a higher level explanation. This means that even if the prior probability of a deeper explanation is lower, it will nevertheless be more highly confirmed by predicted evidence. The virtue of depth is parasitic on the virtue of consilience and of breadth (a deeper explanation, presumably, will often appeal to more basic mechanisms and thus will have a greater explanatory breadth in virtue of explaining a broader swath of things).

5.4   Not Ad Hoc

Coming up with an account of ad hoc-ness that will reveal it as a vice is significantly more troublesome. The standard historical view of what it is to be *ad hoc* comes from Popper. Popper seems to distinguish two notions of *ad hoc*, one of which is merely a liability and one of which renders a theory bunk and worthy of derision. A proposal gets put under the broadly *ad hoc* umbrella if it is posited just to deal with some particular difficulty with a theory. It becomes a proposal worthy of dismissal and derision if it also cannot be tested independently.(Popper, 1974, 986) In the following, I will follow Bamford's use of subscripts to indicate which sort of ad-hockery we are discussing. He gives the broader category the subscript 'e' and gives the full Popperian category a subscripted 'p'.(Bamford, 1999) We will adopt this convention in the following as needed.

For example, suppose I propose the hypothesis that aliens cause all cases of lunch going missing from the office fridge. Now suppose we set up a camera and never see an alien enter the break room. I may modify my proposal to say that invisible undetectable aliens are responsible for lunches going missing. This proposal would be a textbook case of ad-hockery for Popper. It's a proposal whose only function is to save the alien hypothesis in the face of disconfirming evidence and it is (and in fact is constructed to be) not independently testable.

But the waters almost immediately muddy. The muddying comes in attempting

to clear up the crucial criterion that makes an *ad hoc_e* hypothesis also *ad hoc_p*. Does the hypothesis need to not be independently testable in principle or merely actually and pragmatically? Consider as an example of this difference, the hypothesis that an *undetectable* alien stole my lunch. This is untestable *in principle*. Contrarily, the hypothesis that an *invisible* alien stole my lunch is untestable for me pragmatically, but in the future I may be able to obtain a set of infrared goggles and detect whether such an alien *has* been stealing my food.

"In principle" is clearly too weak to capture some of the cases that have come up in the history of science. But it's harder to delineate what counts as *ad hoc_p* from what doesn't (as well as how bad it is to be *ad hoc_p*) when we move in the more pragmatic direction.

> In response to a criticism by Grover Maxwell . . . that the problem of *ad hoc*ness was over-rated, Popper likewise retreats to the stronger requirement for testability in practice. On the very occasion where he distinguishes *ad hoc_p* from auxiliary hypotheses, he ... suggests the distinction is 'a little vague' and re-defines the former to mean 'at the time untestable auxiliary hypotheses', a move which enables him to annexe the neutrino hypothesis as *ad hoc*. According to Popper, Pauli introduced this hypothesis 'quite consciously' as an *ad hoc* hypothesis but that, since the hypothesis did eventually 'shed its *ad hoc* character'..., 'we have a warning here not to pronounce too severe an edict against *ad hoc* hypotheses'. Just as well, for as Maxwell noted, Pauli's conjecture 'opened up an extensive and exciting new area of inquiry, neutrino astronomy'! Within a few years of its invention, the neutrino hypothesis was prominent in theoretical physics and possible tests had been devised for detecting neutrinos, even if detection did then seem a remote possibility and took a further twenty years.

> We did not have to wait for the problem of beta decay in the twentieth century, however, to recognise that the gap between testability in principle and testability in practice is unremarkable. Contemplating the difficulty the apparent absence of stellar parallax posed for Copernicus's heliocentric universe, Wesley Salmon describes the suggestion that the stars might be

'almost unimaginably distant', such that no parallax would be detectable, as a 'highly ad hoc, and consequently implausible, auxiliary hypothesis to adopt just to save the Copernican system'. Salmon concludes that Tycho Brahe's cosmology, in which the sun and moon orbit the earth and the other planets orbit the sun, was therefore 'clearly the best available theory' before Newtonian dynamics. In cobbling together bits of the Ptolemaic and Copernican systems to avoid particular anomalies, however, is Brahe's proposal any the less *ad hoc* than a bold conjecture about the distance of the fixed stars, and how is Salmon's 'clear' preference for Brahe's system warranted given the absence of reliable measurements of stellar distance?(Bamford, 1999)

The problem, for Bamford, with the move toward the pragmatic view is, first of all, that it makes the hypotheses able to move from *ad hoc* to not *ad hoc* in a very short time without anything changing about the internal nature of the theory. Further, as illustrated by the case of the various views of the solar system and the planetary orbits, there are cases where the Popperian view on which what matters is the pragmatic testability of the theory will yield counter-intuitive results. Testability alone fails to get the classifications of theories right. Further, whereas the classical Popperian analysis might call the Copernican system ad hoc but not Tycho Brahe's, Brahe's system seems, on an intuitive level, also to properly be called an *ad hoc* hypothesis.

Jonathan Vogel proposes a broader view of a proposal or theory's being *ad hoc* in "Cartesian Skepticism and the Inference to the Best Explanation."(Vogel, 1990)[21] He says that an hypothesis is *ad hoc* to the extent that it is 'isolated from other explanations and data (it ought to be independently testable, it must figure in the explanation of something other than B, etc.).' It seems to be decidedly in the Popperian camp. But perhaps Vogel's proposal is broad enough to avoid some of the worries with Popper's proposal while making it possible to recognize *ad hoc*ness as a vice. If

---

[21]This broader view may end up being subsumable under some other category, but at the very least it's a proposal that doesn't depend as closely on the independent testability condition.

the heart of being *ad hoc* is a proposal's not providing an integral interlocking set of explanations, there is at least an opening to address Bamford's objections (although we'll have to take some liberties with the parenthetical explanation of what constitutes failing the ad hoc test). Further, there may be reason to think that this broad way of characterizing the vice of being ad hoc will make it easier to probabilistically vindicate it's vicehood.

There may be a simple account (on Vogel's account) of the vice of being ad hoc that makes it a probabilistic liability. A proposal (part of a bigger theory) that's explanatorily isolated will be a liability since, if it explains just one thing, it is likely to be more probabilistically independent from the rest of a body of theory. This makes the conjunction of propositions including the ad hoc proposal less probable than an otherwise probabilistically identical one where the proposal isn't *ad hoc*. Now this can be quite a liability. If we render one conjunct more nearly independent of the rest of the theory we will be significantly reducing the probability of the whole by adding such *ad hoc* hypotheses. However, this doesn't render an ad hoc hypothesis less probable, necessarily, than roughly equal non-ad hoc hypotheses. Vogel's account only gets us that being *ad hoc* is *usually* a liability for a theory.

The key question to ask in deciding whether being *ad hoc* is a liability in all cases is whether there are ways for a fact to make another fact more probable without one explaining the other. In order for explanation factors to justify, there must be probability relations. But if there are such relations then it is open that there are parts of an explanation which would be classified by Vogel as *ad hoc* which are *explanatorily* isolated but which are not *probabilistically* isolated, then there can be cases where being *ad hoc* in Vogel's sense is not an epistemic liability.[22] The key to being able to use a theory's being *ad hoc* against it will be whether one can determine whether

---

[22]I will leave aside the problem with the *ad hoc* vice that it's not it's own vice. The way Vogel reasonably characterizes the vice of being *ad hoc* makes it sound a lot like it's just an extreme case of a view that lacks consilience.

such non-explanatory probability relations exist in the individual cases in which the category *ad hoc* applies.

But now that we have a list of explanatory virtues that are probabilistically vindicated, we will be able to consider the merits of arguments against skepticism which appeal to such virtues. We will also be able to discount parts of argument which appeal to spurious virtues.

Part II

The Arguments and Their

Problems

Chapter 6

Vogel's IBE Against Skepticism

The pioneers of sophisticated defenses of inferences to the best explanation against external world skepticism are Jonathan Vogel, Laurence BonJour, and Tim McGrew, with Ali Hasan and Kevin McCain giving useful development and clarification of this style of argumentation in subsequent work. In this chapter I will discuss Vogel's argument against external world skepticism, and in the next I will discuss BonJour's. Vogel has written a number of articles defending his version of the IBE. In the following chapter we will begin with Vogel (1990) and treat his subsequent essays (as well as McCain's work) as dialectical developments of this original argument.

6.1   Acceptable Skeptical Hypotheses

Vogel begins by arguing that a skeptical hypothesis that is going to be acceptable enough to be a competitor must be isomorphic with the Real World Hypothesis (the real world hypothesis being the hypothesis, roughly, that there are three dimensional physical objects and that these objects have roughly the size, shape, and spatial relations we take them to). Vogel says:

(i) it should invoke items corresponding to the elements of the RWH, and

(ii) it should also posit, as holding of these items, a pattern of prop-

erties, relations, and explanatory generalizations mirroring those of the RWH.(Vogel, 1990, 660)

His argument proceeds primarily from two cases. First he proposes what I will call the Naïve Double Demon Hypothesis created to explain an individual's having an individual experience that is, for example, snow-like. The Naïve Double Demon Hypothesis is a development of the Naïve Demon Hypothesis, which is the label we give to a bare and unrefined demon hypothesis. That is, the naïve demon hypothesis is the demon hypothesis (i.e., the hypothesis that there is some evil demon causing my current experience) that is *bare*, with no auxiliary hypotheses added to fill out the theory. The Naïve Double Demon Hypothesis is a development of the Naïve Demon Hypothesis that remains mostly unrefined, but contains one additional hypothesis. On the Naïve Double Demon Hypothesis we're positing that our experience of the snow-like sense datum is caused by a demon. There is one additional detail of the view which, for example, in attempting to answer why the demon decided to cause me to have an experience as of snow, asserts that the demon is being *told* to cause this experience by another demon. The immediate problem that Vogel raises for this view is that it is "explanatorily idle" and "ad hoc" (Vogel, 1990, 660). It's hard to explain exactly what this accusation amounts to, but I think the point is that this is a "just so" story to explain this particular sensory experience and thus we wouldn't expect this view to explain any of the regularities of our experiences that we take the world outside our heads to explain. Further, even if we were to posit this kind of "explanation" to explain broader experiential regularities, we will at the very least be greatly sinning against simplicity and at worst be failing to provide an explanation at all. Surely this will make the *naïve* skeptical hypothesis far too implausible to be taken seriously.

Suppose that the skeptic wishes to avoid the issues of the Naïve Double Demon Hypothesis, and suppose that she does so by forgoing giving *any* higher level expla-

nation at all. Vogel says that this is worse because of the very fact that it fails to give higher level explanations at all. This is presumably because it will fail to meet the demands for explanatory depth. As a result, Vogel suggests that the skeptic must go to great lengths to match the "rich" and "well-integrated" Real World Hypothesis. Vogel takes requirements (i) and (ii) to encapsulate the basic requirement of the Real World Hypothesis in order to make the explanatory cut.[1]

6.1.1 Do Two Crazies Make a Probable?

The first challenge that this initial argument must meet is the challenge that arises from collecting a number of these ludicrous hypotheses together. Presumably, Vogel is taking the cases of naïve skeptical hypotheses as a case study to show that there are hypotheses that are bad enough explanations to be so improbable that the ordinary reasoner need not consider them seriously. There is a problem with this, which will take its inspiration from Huemer's attempts to come to terms with the intuition that simpler hypotheses are more probable. He argues that the case can be made that an individual simple hypothesis is more probable than any individual complex hypothesis despite the fact of it, perhaps, turning out to be true that it is more probable than not that the true hypothesis is some individual chosen from the pool of complex hypotheses Even if the nonskeptical hypothesis is significantly more probable than any of the crazy skeptical proposals, it may turn out that the skeptical proposals, when disjoined with each other, outweigh the probability of the disjunction of antiskeptical hypotheses. The real world hypothesis being *much* more probable than any individual naïve skeptical competitor doesn't force the skeptic to take refuge in isomorphic skeptical hypotheses in order to have a robust skeptical worry regarding the external world.

---

[1] I still have trouble seeing how richness makes an explanation more probable. An initial guess is that richer hypotheses will have stricter limits on the data that can be squared with them and thus will be confirmed more strongly by friendly data.

6.1.2 Cheats for Explanatory Depth Problem

Vogel's assertion with respect to the Naïve Demon Hypotheses is that they commit the sin of being *ad hoc* and of failing to give adequately deep explanations of our sensory experience. As we have seen above, any *a priori* increase in the probability of theories that give deeper explanations will be parasitic on the probability increasing nature of consilience or explanatory breadth. But I'm not sure that the case can be made that the RWH brings together more disparate hypotheses than Naïve Double Demon Hypotheses. I'm not sure what such a hypothesis will do to the probabilities of various experiences on each other. However, there's a way that depth and lack of depth can still be relevant to probability. Certain things seem to "call out" for explanation, and it's tough to nail down what it is about a thing that makes it call out so. It's not just being a contingent fact as there are some necessary truths that seem to call out for explanation (for example, on some versions of theism the facts about what God wills are necessary truths but they would still seem to be the sort of fact that calls out for explanation, despite being necessary)[2] and some contingent facts may have little to no explanation (for example, on certain interpretations of quantum mechanics, quantum tunneling may result in the existence of a particle which has no explanation for its coming into existence).[3] So we can see how, sometimes, one can fail to provide enough explanatory depth and thus be giving a much less probable explanation. If there is *anything* that "calls out" for explanation, then clearly such a case would include the regularities that our experiences exhibit.

But Vogel is right to accuse the Naïve Double Demon Hypothesis of failing to

---

[2]This, though a fringe case (and perhaps difficult to call a legitimate case of a necessary truth calling out for explanation) is probably one of the easiest cases to come up with which is a case where there is a necessary truth which yet yields some intuition that the truth requires causal explanation). It may not be a particularly intuitive case, but there are few such cases. It will be important later for Vogel that such kinds of explanations and callings out for explanations can happen in the case of necessary truths.

[3]This would have an explanation on the SR model of explanation, but it's not clear that the phenomenon would have a *cause* in the sense in which we normally use the term.

provide adequate explanatory depth here. The Naïve Double Demon Hypothesis (NDDH) fails to explain the very regularities that most call out for explanation in our experiences. The Naïve Demon Hypothesis without positing another demon or other higher level explanation does no better on this account.

However, we can simply construct some *slightly less* naïve demon hypotheses and avoid the problems here while not going so far as giving Vogel what he wants in providing an isomorphic skeptical hypothesis and still retain the required explanatory depth. There are a couple of modules that can be added onto any of the naïve demon hypotheses which will work toward explaining experiential regularities that will not commit the explanatory sins of the original naïve proposals. Consider the following.

An important question for an additional level of depth to explain is why the demon's victims are subject to experiences that exhibit such-and-such regularities as suggested above. One initially attractive option would explain *away* rather than explain such regularities, but this is hardly a mark against a view on its own. The explanation, on this view, is just that my brain is one that is hard-wired to see regularities even when such regularities aren't there. Examples of this tendency are legion. Consider our penchant for seeing faces and animals in cloud arrangements, seeing faces in pieces of toast and door frames and seeing headlights when presented with just the right arrangement of fireflies when driving on a dark road at night. Or consider our tendency to commit the gambler's fallacy. The gambler's fallacy is reasonably construed as our brain insisting on seeing randomly selected events as a kind of regularity, e.g., "The roulette wheel has selected so many reds that I'm *due* for a black soon." And consider confirmation bias, in which, having reached some conclusion (say that there is such and such a regularity in experience) we unconsciously ignore evidence to the contrary. Further, this tendency does more than that. The human need for regularities sometimes even changes experiences such that the human mind will *create* the experiential regularity. Consider, e.g., the case of the blind spot

in our retina. Our eyes have a blind spot. The optic nerve travels through our retina to travel back to our brain. But we don't experience an empty spot. There's no black spot in our experience. There are visual experiments that you can do yourself in which there is an x on a sheet of paper which, when focused on in one's blind spot, will disappear. This will leave an experience as of a blank page (it even works on lined notebook paper and paper of various colors). This is a case where a tendency to see regularities where they don't honestly exist can actually *generate* them in experience. Our tendency to see regular shading patterns on a page produces the actual state of affairs of the experiential regularity. The person who endorses the naïve demon hypotheses (and even the NDDH where the first demon produces a deception based on a second demon's whims) can appeal to psychology to deepen her explanation without moving much distance at all toward Vogel's insistence on isomorphic skeptical hypotheses. Perhaps there will be other reasons why this explanation is implausible,[4] but the grounds of its implausibility do not seem to be the explanation's failure to be isomorphic with the real world hypothesis.

But this proposal is not a great one. It strains credulity to propose that *none* of the ordinary experiential regularities are actually regularities. But a slightly modified case which acknowledges the regularities may be more plausible. One might go closer to Kant and say that the experiential regularities are there because the mind, in some way, can't help but experience things in that way. This will, rather than explaining *away* the regularities of experiences, *explain* them from by appeal to the human being having experiences than by way of the demon. A theory like this may only gain plausibility when limited to explaining certain necessities of experience (see the discussion of isomorphic skeptical hypotheses for discussion of this possibility).

---

[4]The degree of success this proposal enjoys will depend almost entirely on the extent to which the view intends to explain away regularities. It may work as a module to explain some subset of them, but if *all* the regularities are explained by human beings' tendency to see them whether they're there or not, very radical skeptical problems ensue. One will have to be prepared to give up much more than the external world in order to maintain this as a *full* explanation of the regularities present in our experience.

Secondly, rather than explaining *away* the regularities, we can explain them without significantly thickening the naïve demon hypotheses. When given Vogel's question of how to explain his current experience, one appeals to the demon's causing his experience. To the further questions of why a snowy experience, the answer will be that the demon causing his experience recognizes that causing such an experience would be part of a good deception. This has the advantage of giving a deeper explanation while not requiring one to add much "thickness" to a naïve demon hypothesis. This is because the main further questions will be questions about why such and such experiences constitute a good deception, which will mean that the naïve demon hypothesis' explanatory depth comes largely "free" in terms of simplicity. All the facts that one will be appealing to in deepening this explanation (in this direction) further will either be facts about human psychology or necessary truths.

Vogel may not be satisfied with this. He may assert that the other places where demon hypotheses lack depth are also problematic. For example the question of the demon's motivation for deceiving us may come up. But this isn't the sort of question that is germane or harms explanatory depth, unless one has some antecedent reason to think that the individual behind such a deception wouldn't do such a thing. Consider two different cases:

Friendly Fiona: Suppose Friendly Fiona is one of my best friends. Now, Fiona invites me to come over and play cards and when I arrive she consistently wins, and wins big. Suppose also that a good explanation of this fact is that she is cheating at the card game. In such a case we would expect that an explanation, in order to be acceptable, illuminate why Fiona might cheat me in this way as a condition of this explanation being good enough to accept.

This is not true in the case of **Sketchy Steve**

Sketchy Steve: Suppose Sketchy Steve is completely unknown to me. Further, I end up playing cards with Sketchy Steve, whom I have no antecedent reason to believe is either very honest or very dishonest. Now, Steve consistently wins and wins big. Suppose that a good explanation of this fact is that he is cheating at the card game. In such a case we would *not* expect an explanation for *why* Steve is cheating me in order to make the theory acceptable.

The demon hypothesis above is analogous to **Sketchy Steve** in a close enough way that his motivations need not be provided.[5] Further, no particular motivation needs to be given to make the theory better, despite the fact that there are a few motives one might attribute to this demon (for example, he might be treating the deception as some kind of demonic science experiment, he may want to wait for the moment of death to cause the human the suffering of realizing none of it was real all along, or he may simply think humans are stupid and enjoy the laugh of trolling them so hard with his demon deception). The theory doesn't need to appeal to motives to explain, and is better remaining agnostic on such questions.

There *are* however, questions which may be more troubling for the naïve demon theorist. For example, there is the question of how the demon *knows* that such-and-such experiences would be part of a good deception. This is a more interesting question that threatens the naïve and non-isomorphic quality of the current skeptical hypothesis. However, one can simply give the demon knowledge of the truths about what would be a good deception in the same way that he knows whatever other truths he presumably knows (by some sort of direct insight perhaps). Further, one would assume that the demon has intimate knowledge of the human mind as well. These two kinds of knowledge will come together to complete this request for explanation. It's still not clear that one is required to go so far as to propose an isomorphic skeptical hypothesis in order to generate a good enough skeptical competitor.

Yet it remains intuitively true that the real world hypothesis is a more *rich* explanation than the modified naïve skeptical hypotheses. However, it's hard to make the case that we should just prefer such a "rich" explanation over other explanations unless we are already convinced that we are in a complex physical world with basic and derived natural laws. If we're convinced of this, then explanations should be able to go to numerous levels, but the mistake is that in the skeptical hypothesis we aren't

---

[5]It may be that demon hypotheses are in an even better condition (the perpetrator of the deception is called a demon, after all).

appealing to such a world. It's a world with just the mind of the demon and my mind. The intuition that we need more and more basic mechanisms misses the mark in discussing such worlds.

But it will be helpful to move on to consider what Vogel says to make the case that we should prefer (keeping in mind our modified naïve skeptical hypotheses— MNSH from here on) RWH to the richer and more well-integrated isomorphic skeptical hypotheses.

6.2   Skeptical Hypotheses and Non-Colocation

As an example of an isomorphic skeptical hypothesis Vogel introduces CSH (the Computer Skeptical Hypothesis). On the CSH, rather than there being physical objects with locations in 3-dimensional space interacting according to physical laws, there are pseudo-(physical objects) which have pseudo(location) and interact according to pseudo-(physical laws). That is, there is a computer that simulates a real world, with certain arrangements of data on a disk representing objects and their locations with the computer behaving in ways that emulate physical laws' operation on real world objects. "In general, what the RWH explains by reference to genuine location, the CSH will explain in terms of pseudo locations."(Vogel, 1990, 664)

This leads to a problem, for as in the Real World Hypothesis different objects all have different locations, the CSH will need to incorporate into its program a rule that assigns each of the different pseudo objects to a different pseudo location. But this, for Vogel, is a huge advantage in terms of quantitative simplicity for the RWH. The Real World Hypothesis has a different location for every object just because it's a metaphysically necessary truth that physical objects don't colocate. Whereas the Real World Hypothesis can appeal to a necessary truth to account for the non-colocation of objects (which allows the real world hypothesis to avoid adding to its complexity in order to so account), the Computer Skeptical Hypothesis must appeal

to contingent truths about the programming of the deception computer in order to account for the same widespread regularity in experience. Vogel thus seems to have established that the real world hypothesis is more virtuous than the CSH. It is natural to suppose that other skeptical scenarios will have similar such weaknesses.

6.2.1 Response 1: Saving the Computer Skeptical Hypothesis

Vogel's main contention here is that, while the reason that real world objects don't colocate in our experience is the necessary truth that physical objects don't colocate, the Computer Skeptical Hypothesis cannot appeal to that necessary truth in order to explain the fact that pseudo objects don't pseudo-colocate. I'm not convinced this is true. Vogel's case has just failed to consider a certain possibility for how the computer in the CSH is programmed.

Aside on Data Structures: Arrays

An array is a particular kind of data structure used widely in computer programming. One way to think of this kind of data structure is as a multi-dimensional table for which one can set the number of elements in each dimension. When one creates an array, one creates a set of locations in which to put data that are addressed by a location in the table (the locations are also actual physical locations on a disk). Think of the data structure like an ice cube tray that one has created into which one can put different things. Picture an ice cube tray two cubes wide and 8 cubes long. Now, suppose one puts different flavors of juice into each spot on the tray. Let's say the third spot on the left column of one's ice cube tray has frozen grape juice in it. In this case the frozen grape juice would be the "data" in location 1,3 of the array. This is the way that arrays store data, and one can have as many dimensions in an array as one wants. I propose that the computer program to which the BIV is subject program the physical world as a set of arrays, with each pseudo-object's location in the array treated by the program as the physical location of the object in the world.

A brief metaphor will help to clarify the difference between programming a world using arrays and without using them.

Basic Bob is going to be put in charge of pushing the buttons that directly cause people's experiences. The data for this analog program are stored in boxes that are strewn around the room. For each object, Bob is given a piece of paper with a name and an numerical value on it, with the name being the name of the object and the numerical value representing for Bob the location of the data box in the room. When Bob is performing the instructions for his program called "systematic deception," his program asks him to go to the various boxes and retrieve the piece of paper with the data about the particular object in it, remove it, and use the data on the paper to decide what experience-causing buttons to push. For example, say the program is causing an experience as of a chair. Bob finds the piece of paper labeled chair, looks at the location on the piece of paper, and goes to retrieve the box in the correct location. He then opens the box and looks at the data sheet, which has stated on it the location, color, etc. of the chair. Finally he follows his instructions and produces the experience as of the chair as being in such and such a location, of such and such a color, and so on.

Array Alice is, in an alternate scenario, going to be the one who's put in charge of pushing the buttons that directly cause people's experiences. The data for this analog program are stored in a set of boxes arranged in a cube in the center of the room. Alice is given a piece of paper with a set of coordinates on it, representing to her the location in the room of each box stacked in the cube of boxes in the center. When Alice performs the instructions for her program called "systematic deception using arrays," her program asks her to go to the various boxes in the cube of boxes and use the data on the paper in each box and the data's location to decide what experience-causing buttons to push. For example, suppose the instructions will want Alice to display the experience that would be caused by the object in some particular

location in the array. Alice is instructed to go to the box in location $\{1, 3, 5\}$. She opens the box and the piece of paper inside says the object's name is "chair", its color is tan, etc. Alex then presses the buttons which cause an experience as of a chair to whatever individual is looking at location $\{1, 3, 5\}$.

The case of arrays in computer programming is similar to the case of Alice as contrasted with Bob above, although the particular reality of the issue is not so simple. Arrays feature data stored in locations within a multi-dimensional matrix and are addressed in terms of this location, and my suggestion is that a computer skeptical scenario can treat the location in an array as the physical location of objects.

Computer Skeptical Hypothesis Using Arrays

Suppose that, on the Computer Skeptical Hypothesis, the pseudo world is stored as a three dimensional array, such that, for example, at the memory location $\{132, 500, 200\}$ is a pseudo-(table-in-a-coffee-shop). And suppose that memory location $\{132, 500, 202\}$ is a pseudo-(chair-in-a-coffee-shop). What explains the fact that these two objects cannot have the same pseudo-location is just the necessary truth that objects cannot colocate (this is because the pseudo-location of the objects in the world isn't some value assigned to a variable but rather is the location on the disk addressed by the three-dimensional array). But the two pseudo objects can't both be stored at the same location on the disk, so objects don't colocate in the experience of the sufferer of the Computer Skeptical Hypothesis.

The first natural objection is to accuse the advocate of the CSH of misunderstanding the point of RWH's advantage over CSH. For the RWH, the necessary truth that objects don't colocate *explains* the absence of apparent colocation of objects in our experience. But the modified CSH has merely *partially* explained the absence of apparent colocation of objects in our experience by this necessary truth. The CSH must also use other contingent facts, like facts about how the computer is programmed to explain the phenomenal non-colocation. The claim is that this makes the RWH

simpler than the CSH. But RWH is no better on this front.

Hold on there. The RWH isn't *that* good.

Remember that it is a *phenomenal regularity* that is being explained by the RWH and the CSH differentially here. It is thus important to get straight what the phenomena actually are that are meant to be explained. In order to do this the reader should engage in the following exercise:

IKEA Bookshelf Exercise: Imagine two very similar IKEA bookshelves, each black with square cubby-hole style shelves. Imagine both standing on the same wall. Now imagine them slowly pushing into one another such that they partially phenomenally colocate.

Has one succeeded in imagining the kind of thing the absence of which needs to be explained? If one has, then what was said above is true. RWH is no better than the modified CSH above. The necessary truth that objects don't colocate (granting, still, that this is a necessary truth) is only a *partial* explanation of the fact that objects don't phenomenally colocate, just like in modified CSH. For if there were odd physical laws, this very phenomenon could be produced by a world in which objects don't colocate. If we were in a universe with laws where objects, when they touch, combined into one in a certain way, we could have a universe in which IKEA bookshelf style cases are abundant. So, at best, in the Real World Hypothesis the necessary truth that objects don't colocate merely explains why objects don't phenomenally colocate for *one* reason. They don't explain why the many other ways that objects could seem to colocate don't obtain. The RWH theorist must appeal to other contingent facts (or at least weak necessary facts, which I take to not be explanatorily "free" like other necessary facts), like facts about the specific world's laws (or how the world is programmed, as it were).

6.2.2   A Dilemma Arises

One might argue against the applicability of the above exercise that it is actually impossible to have a phenomenal experience of objects as colocated, veridical or not. This is obvious with respect to total colocation, since there would be no phenomenal difference between this case and our ordinary perception (except, perhaps, in a very highly specified case to be discussed below). In fact, if we're talking about total colocation, we seem to have experience that is indeterminate between being an experience as of non-colocation and as of an infinite number of objects colocated. This cannot count as experience of objects as colocated. It wouldn't be a regularity of experience to explain if it happened all the time. We're stuck trying to make sense of imagining two objects as colocated. There are four main ways to attempt to produce a mental image according to which objects are partially phenomenally colocated, which take their inspiration from the ways that movies attempt to represent these kinds of phenomena (the following will harken back to the IKEA Bookshelf Exercise):

The Ghostly Strategy: One way of imagining two objects as colocated is to adopt the Ghostly Strategy. One way to make objects move visually through each other is to make the objects ghostly or gaseous at the points where they touch. This is the way that ghosts are often portrayed as passing through each other in movies. Using the Ghostly Strategy, the bookshelves can be imagined to be partially overlapping. The bookshelves are imagined as gases, able to unproblematically pass through each other.

The One on Top Strategy: Another strategy is to imagine one as always on top. That is, one may push the bookshelves together phenomenally by imagining that, wherever the objects would overlap, one is imagined as on top, so that one only sees one object at any one time on top, seeing the other objects as "filling the gaps" in the other object. In the bookshelf exercise one might imagine the left bookshelf as "on top" such that all one sees of the other bookshelf is where it's emerged into the spaces of the left bookshelf. With this strategy it's also possible to pick and choose which object one will imagine "on top" in any particular situation. For example, one might imagine that at some points, the right bookshelf is in front, say the top shelf.

The Amoeba Strategy: One might also adopt the Amoeba Strategy. In the Amoeba Strategy one imagines the two objects as becoming *one* object sharing some of the properties of each and losing some of the others. On this strategy one imagines the

objects as one amoebic thing. For example, the bookshelves being pushed together would just look like one really oddly arranged bookshelf.

The Magic Eye Strategy Another strategy one might adopt in order to imagine objects as colocated is to adopt the kind of strategy that one uses with one's eyes in order to see the quasi three-dimensional objects in magic eye paintings. In this strategy the individual crosses her eyes in particular ways and attempts to get her brain to combine the double image that is present into a single image. One can attempt to implement this sort of thing with ordinary objects wherever one is. The key to this strategy is to produce such a double image in one's head and attempt to combine it into one.

These are either ways of imagining objects as colocated or they aren't. If they are, then the RWH's case above is further weakened against the modified CSH. But there is a reasonable case to be made that they aren't, and further that it is impossible to *imagine* objects as colocated.

Consider each strategy in kind. Consider an image produced according to the Ghost Strategy. There are a couple ways of thinking about this kind of case, neither of which seem to be imagining objects as colocated. On one way of thinking about the case, the change that happens when the objects overlap is one where the two objects become one gas with various distinguishable colors and shapes within the gas (loosely speaking, where we might call steam or clouds gaseous or a kind of gas). We also might think of them as two gases passing through each other according to odd physical rules. When gases normally run into each other, the particles in each gas act on the particles in the other and there is a big swirling mess. In the Ghost Strategy, one could be thought of as imagining two gases passing through each other in a less interactive sort of way. Where the real objects are just the parts of the gas and they happen to act in a way that retains their shape and some colors as they pass through each other. Neither of these are imaginings of objects as colocated, and certainly they are not cases of imagining *solid* objects as colocated.[6]

---

[6]Now, one might point out that ordinary objects aren't really solid in this way, so that there is some non-zero probability that my two identical pens, on running into one another, simply move such that they entirely colocate. There are several ways one might picture the experience of this phenomenon. One might picture it by using one of the strategies already mentioned above, or one might picture the experience as one that looks like one object now. But this isn't experiencing

Consider, second, the One on Top Strategy. It should be clear that the One on Top Strategy is not successful in producing objects phenomenally colocated. You have just imagined, in the pristine case, one bookshelf with various bookshelf parts in the spaces of the bookshelf. One has again not imagined colocated objects. The same goes for the Amoeba Strategy. Taking the Amoeba Strategy one has just imagined two objects becoming *one* object. This is yet again not imagining phenomenally colocated objects.

The Magic Eye Strategy is a little better than One on Top and Amoeba, because it has the advantage of preserving two objects. Suppose I look at the two stools in the coffee shop and cross my eyes to push them nearer each other, and then attempt to make the two images into one image. The problem is that one of two eventualities happens when attempting to combine this into an image. One either succeeds, in which case one is, inevitably, simply imagining one object, or one fails, in which case one doesn't have *an* image that one is imagining, and further, whatever success one has achieved in this strategy simply mirrors either the One on Top or the Amoeba Strategy.

Having attempted to make sense of a kind of experience that would reasonably be called "as of objects colocating," we've reached our inevitable failure. But having exhausted the methods for making colocation experiences possible for humans, we are inclined to call colocation experiences impossible for human beings. Being

objects as colocated. One might, instead construe the case as one where I have two pens which are identical in every way except that one is blue and one is red. One might think that the experience then would be an experience as of pushing together two pens and then having an experience as of *one* pen which is purple. Further, it may be that the pen(s) feel twice as heavy as a normal pen. This kind of case is one which attempts to produce the experience as of objects co-locating over time (i.e., diachronically) rather than in an instant of visual experience (i.e., synchronically). In this case I'm not sure why we would call this an *experience* as of co-location. Rather, it seems to me to be various non-co-location experiences which, when put together, might lead us to conclude that objects have co-located. The experiences before the complete merging will be dealt with by our various strategies discussed above and below and the experiences once there has been a merger are completely ordinary experiences. Further, it's not even clear that these are experiences which should lead us to conclude that objects have co-located, unless we already antecedently have a particular view of the world as having specific kinds of scientific laws regarding the fundamental particles which make things up.

⟨impossible for human beings⟩ is a modest conclusion from this discussion and, already there is an answer for the antiskeptic. First of all, if it's impossible that human beings experience colocation then it's not clear that any explanation is required. There is a necessity on the level of the phenomena. If one demon were to ask another to produce an experience of colocation, the demon would refuse because it is impossible. There is no need to account for what is necessary already by appealing to other things. The antiskeptic is forced to argue that this necessity is the kind that "calls out for" explanation to proceed with their argument.

Further, if an explanation is required, it is reasonably thought to be on the side of human beings rather than on the side of the causes of experience (in fact, our results suggest that even in cases of imagining in which there is no external cause of our experience, the experience of colocation is impossible for human beings). A quasi-Kantian explanation of this regularity is just as good as the external world explanation. Our brains/minds are just as important links in the causal chain producing experience as are more remote causes. It is natural to suggest that there is something about the human mind that accounts for why we cannot help but experience the world as geometrical and following specific geometric and colocation based rules. Why can't the computer monitor display anything but two-dimensional images? Is it because there is something about the input the monitor is getting that explain this fact? No, no matter what data are sent the monitor couldn't produce anything but a two-dimensional image because it is so constructed as to be only able to produce two-dimensional images. As it is natural in cases like these to posit explanations on the side of that which receives and displays the image, it is natural to posit that there is something about the human being that prevents these impossible co-location experiences. When one is presented with a necessity about some occurrence in an individual, one should first look for explanations that appeal to facts about that individual to so explain.

Now suppose the impossibility is stronger yet. Suppose it's impossible for *any* experiencing thing to experience objects as colocated. If this is true, in addition to the above, there is also a new skeptical scenario that has just the same advantages as claimed by the real world hypothesis. Consider the following.

If the reason that none of the strategies succeed is that it is impossible for any creature to have experiences that are "objects colocated" experiences, then there is another demon scenario that will dodge Vogel's argument. Consider the Imaginative Demon Hypothesis (IDH). On this skeptical scenario a demon imagines a world and then creates a deception based on that imagined world. So, the answer to Vogel's above questions of why the demon produces in me a snow-like experience would be that there is snow located phenomenally in front of me in the demon's imagined universe. But the fact that we don't experience objects as colocated is just straightforwardly accounted for by the fact that it is impossible to experience such objects and so the demon can't imagine such a world in order to produce such an experiential world in us. The Imaginative Demon Hypothesis is meant to be the demon equivalent of the CSH, a demon scenario that gives a theory isomorphic with RWH. This seems to be just as good as the RWH on this second horn of the dilemma as well.

If all of the above is right, then Vogel has either failed (in Vogel (1990)) to show the RWH to be stronger than the modified Computer Skeptical Hypothesis or the exercise isn't a case of phenomenal colocation, in which case the unmodified CSH AND the IDH both do at least as well as the RWH (A demon scenario can also be defended against the real world hypothesis while giving up the assertion that experience is necessarily non-colocatory, simply allowing a different necessary truth to enter into the explanation discussed, but we won't discuss this proposal at length until it, perhaps, comes up later). But Vogel has developed and expanded on this argument in the period since the release of "Cartesian Skepticism and the Inference to the Best Explanation."

Before moving on to attempts to improve Vogel's argument, it is worth summarizing more formally the main thrust of the argument against Vogel which I have been defending in this chapter:

1. Suppose that it is not necessary that my experiences are non-colocating.

2. If it is not necessary that my experiences are non-colocating experiences, then it is possible for a non-colocating world to cause experiences as of colocation.

3. If it is possible for a non-colocating world to cause experiences as of colocation, then the necessary truth about non-colocation of objects only partially explains the fact about experiences being non-colocation experiences.

4. If the necessary truth about non-colocation of objects only partially explains the fact about experiences being non-colocation experiences, then there are skeptical scenarios that seem to have just the virtues alleged of the RWH.

5. Suppose it is necessary that my experiences are non-colocating.[7]

6. If it is necessary that my experiences are non-colocating, then it remains unestablished that the facts about experiential non-colocation require explanation.

7. Suppose that this necessity does require explanation.

8. Either the necessity discussed in 5 is merely a necessity about the experience of all human beings or it is a necessity about the experience of any experiencing thing.

9. If the necessity discussed in 5 is merely a necessity about the experience of all human beings, then a quasi-Kantian explanation of the experiential regularity is as reasonable as an explanation that appeals to regularities in an external world.

10. If the necessity discussed in 5 is a necessity about the experience of any experiencing thing, then there is another demon scenario which has the same virtues as RWH[8]

The above argument structure can be applied to any such regularity Vogel might consider.

---

[7]I argue that it is in fact necessary, which will push one onto the following horn of the dilemma I've constructed. However, the horn presented in 1-4 shows that the argument is just as roundly defeated if the admittedly weak argument for necessity fails.

[8]Gifford (2013) presents an argument that has a structure that is in some ways similar. However, I think there are advantages to my argument in that I don't require any controversial points about whether Vogel's particular explanatory proposal is admissible in this case. There are also a few quibbles that I have with Gifford, which we need not get into at present.

6.3 Skeptical Hypotheses and Geometrical Regularities

Now, the argument of Vogel's discussed earlier is merely Vogel's first published work on the issue. He has since clarified and strengthened the argument in Vogel (2004) and Vogel (2008). This section will consider Vogel's developments of his original argument in Vogel (1990) and evaluate the merits of the argument. In the end, the clarifications and developments by Vogel will help illuminate the responses already given in a very useful way.

6.3.1 Triangle Inequalities, Perfect Solids, and Having a Geometry

Vogel's primary new case in Vogel (2008) is the case of the triangle inequality. He claims that, like the non-colocation of objects (and perhaps even better), isomorphic skeptical hypotheses cannot as well explain the experiential regularities that are explained by the triangle inequality in the real world hypothesis. He says:

> At this point, the question may be asked why the genuine spatial distance $Dist(a, b, c)$ is greater than $Dist(a, c)$. This fact is a consequence of a *necessary* truth about distances among non-collinear points known as the triangle inequality. The corresponding question with respect to ISH would be why the magnitude of $Dist^*(a^*, b^*, c^*)$ is greater than the magnitude of $Dist^*(a^*, c^*)$. According to ISH, the $Dist^*$ relations generally are relations among magnetic patterns at various chunks of the computer disk, including L_A^*, L_B^*, and L_C^*, such that these relations are causally efficacious in certain specified ways. What magnitudes are assigned to those relations is a matter of how the computer happens to work. So, according to ISH, the fact that $Dist^*(a^*, b^*, c^*)$ is greater than $Dist^*(a^*, c^*)$ would be *contingent*. If ISH leaves this fact unexplained, it would fail to explain all that RWH does. So it seems that ISH has to introduce some further empirical regularity otherwise analogous to the triangle inequality that has the consequence that $Dist^*(a^*, b^*, c^*)$ is greater than $Dist^*(a^*, c^*)$. ISH might then explain all that RWH does, but its explanatory apparatus would be encumbered by an extra empirical regularity that has no counterpart in RWH. ISH would then be less simple than RWH. Hence RWH

would enjoy an explanatory advantage over ISH and would be preferable to ISH on that account. (Vogel, 2008, 548)

The new fact then is the triangle inequality explaining regularities in experience, such as the fact that it always seems that traveling from point a to c takes less time than traveling from a to c by way of b when holding one's velocity constant (where a, b, and c are vertices of a triangle.

Vogel makes it clear that the underlying structure of his argument is very similar to his original argument in Vogel (1990) in this geometrical case. Vogel describes his general line of argument in the following::

> The claims about distance relations and their skeptical substitutes are meant only to illustrate a general line of argument, which goes as follows:
>
> 5a. There are necessary truths about spatial/geometrical properties that have no counterparts for non-spatial/geometrical properties.
>
> 5b. These necessary truths enter into explanations with respect to spatial/geometrical properties.[9]
>
> 5c. If non-spatial/geometrical properties were to discharge the explanatory function of genuinely spatial/geometrical properties, empirical regularities would enter into these alternative explanations in place of the necessary truths that figure in ordinary explanations in terms of spatial/geometrical properties.
>
> 5d. Hence the alternative explanations would be less simple than, and inferior to, the ordinary ones.(Vogel, 2008, 548-549)

The point of Vogel's cases are to give examples of how a world that is non-geometrical will always fail to be as simple an explanation of our geometrical-rule-following experience by a world that *is* geometrical.

---

[9]A better wording might be: "enter into all explanations which appeal to spatial/geometric properties." The point here is that necessary truths about spatial/geometric properties will be a part of any explanation which appeals to actual spatial/geometric properties, which will push views which do not appeal to such actual spatial/geometric properties to, perhaps, posit a contingency in the place of this spatial/geometric necessity.

The problem for Vogel is that each of the two strongest skeptical proposals above are perfectly up to this challenge. This is because both the Imaginative Demon Hypothesis and the modified Computer Skeptical Hypothesis posit skeptical worlds with geometries and can thus appropriate any of the necessary geometrical truths that are so helpful to the real world hypothesis. The Imaginative Demon's deception world is a world that has a geometry just like a real world would. It's just a phenomenal geometry. The Imaginative Demon Hypothesis will completely avoid challenges of the general form of 5a-5d.

This is a little more difficult to show for the modified CSH. The complication comes from the distinction between the data structure that is an array and how it will be stored on the disk. On a normal way of thinking of how data are stored on a disk the array will be stored in a way that each object has real spatial relations, but since it is 3D data being stored on a 2D surface the spatial relations of the data-on-the-disk will not mirror the spatial relations of the phenomenal world. This will mean that some, but not all, of the problems of not having a geometry will be solved by arrays, if we are relying on the actual spatial relations of magnetic storage on a classical hard disk to do the work. That is, the modified CSH, on this account, will be able to deal with the non-colocation problem but will not be able to deal with the triangle inequality problem.

There are two ways that one might go about solving this problem, the first of which involves digging one's feet in on the modified CSH and the second of which involves further modifying the CSH. The first way of rescuing the CSH is to call attention to the logical geometry of the array data structure rather than the geometry of physical bits stored on a physical disk. The array data structure is basically a system of coordinates into which to put data and as such a system of coordinates has a geometry just like a set of Cartesian coordinates has a geometry. It's not a phenomenal geometry and it's not what I will call a "physical" geometry, but it is a geometry nonetheless. And

as a geometry, it follows the rules of geometry necessarily just like phenomenal and physical geometries. The modified CSH can use the necessary truths about geometry to account for their experiential regularities just like the RWH uses the necessary truths about geometry.[10]

If this is unsatisfactory, one may simply modify CSH once more to more tightly connect the physical implementation of the code with the properties of the data structure. Suppose that the array is implemented in a way that directly reflects its logical structure. The various bits in the array are stored in their respective places in a three-dimensional storage drive, such that the array item $\{1, 1, 1\}$ is in the $\{1, 1, 1\}$ spot on the drive. With this slight modification, which we will call the 3D-CSH, the "world" that we experience quite literally has a three-dimensional geometry, and can use any of the geometrically necessary truths that are used in the real world hypothesis.

With regard to the 3D-CSH, one may worry that we are getting dangerously close to meeting our minimal stipulated conditions for defending antiskepticism. The 3D-CSH entails that there is a world of 3D objects, and on a reasonable understanding of what it means for the 3D objects to be "roughly the shape that we take them to be" this aspect of antiskepticism will be fulfilled as well. However, the point where it fails to meet the antiskeptical challenge is that I'm not sure it counts as positing a world of "spatial objects," except in the very loose sense that patterns of magnetization on a disk count as "spatial objects." At the very least it's not a world which posits the magnetizations of a disk as the spatial objects of our experience. It's as a result of bits of data stored in thus and such location on the disk that we have our experiences of the object and not a result of the shape or location of the object. The object

---

[10]One might worry about the coordinate system of the array data structure and it's seeming to imply a discrete rather than a continuous geometry, but this need only imply that the objects will be in discrete geometrical units. It need not imply that the same sort of continuous geometrical relations *amongst* the objects fail to hold, just like the fact that we are justified that the objects posited by science have a minimum size (given by Planck's constant, if I remember correctly) need imply that objects don't have distances between each other representable by lines.

causes my experience not in virtue of it's being a material object of a certain shape in a certain location. It doesn't matter to our experience that the array happens to be arranged in a strictly spatial way. But what we experience is a shaped object, not an object or set of objects with their data such that a computer program would give us a shapely objectish experience.[11]

All of the proposals above fare extremely well against Vogel's proposed cases (and promise to do so against future proposed cases precisely because they are scenarios which have geometries and thus can take advantage of any geometrically necessary truth in precisely analogous ways to the real world hypothesis. The success of these skeptical scenarios as just specified is not all there is to the current case. The same argument structure as used above against the colocation case applies just as well to this case. There is even stronger reason to think the regularity Vogel wants is necessary.[12]

6.4   McCain's Strengthening of Vogel's Case

Kevin McCain takes on precisely the sort of response given above in *Explanationist Evidentialism.* He says:

> This virtual space could be a mathematical representation of an imaginary space that has the same mathematical properties as the space of CS [Common Sense]....Although this objection [appealing to geometries of virtual spaces] is initially plausible, it ultimately fails....DH* will still need to posit fundamental regularities that CS does not. DH* will need to posit a fundamental regularity that the demon implement a plan that includes such a virtual space and the demon will not change plans after its deception has begun. Without this regularity there would be no guaran-

---

[11]This may not be a skeptical scenario in the final analysis, but it certainly seems that, if the only difference between this scenario and a regular CSH is where and how the data are stored, one would be hard pressed to say that this has made the difference between the scenario being skeptical and not being skeptical.

[12]The case for this particular property being a necessary property of experience as such is much stronger, and thus the case that this is even a necessity that requires an account is in even greater need.

tee that the distances between $a*$, $b*$, and $c*$ are governed by the triangle inequality theorem because it is possible that they are not arranged spatially at all since they are mental states or features of mental states of the demon.(McCain, 2014, 135-136)

Modified CSH and IDH, despite being improved theses, must still posit extra explanatory regularities that are not necessary truths and thus remain less good theories than the common sense view or the RWH. The structure of this kind of claim about skeptical scenarios involves two main subclaims. The following expresses this structure, leaving open whether one is appealing to entities or regularities.

(Mc1) An isomorphic skeptical hypothesis requires all the contingent entities/regularities (or close analogues) required by the real world hypothesis.

(Mc2) An isomorphic skeptical hypothesis requires contingent entities/regularities that the real world hypothesis does not.

(Mc3)Therefore the isomorphic skeptical hypothesis is less simple than the real world hypothesis.

It is important to notice regarding **(Mc2)** as it occurs in McCain's argument that, even if the demon world (and the modified CSH world) requires certain fundamental regularities, the real world hypothesis requires analogous fundamental regularities. A real spatial world is just as capable of causing experiences as of the failure of the experiential regularities appealed to by Vogel and McCain. We can simply consider the brain state caused by a computer skeptical hypothesis in a case where it produces an experience that doesn't follow the triangle inequality (whatever that would be like, if possible at all) and have such a brain-state produced by an odd quantum event. Further, a computer skeptical world *can be* a real spatial world (e.g., a computer world in which people are recently envatted is a real spatial world), so any experience producible by a computer on a brain in a vat is producible in us by a spatial world.

175

The experiential regularities that are being accounted for by the different accounts all require a geometry plus a bunch of other auxiliary hypotheses a number of which are explanatory regularities. Whereas with the demon, these auxiliary hypotheses will be hypotheses about the demon's executing a plan to give a deception while in the case of the real world hypothesis these will be hypotheses about particular natural laws.

We should return to a point about what is being explained. What is being explained are experiential regularities. There are two kinds of regularities that are accounted for (at least in part) by the triangle inequality. The first kind of regularity is unimportant as it appears to be nonsense for a *visual* experience not to follow this kind of regularity with respect to the triangle inequality. These are the cases where our synchronic experience follows the triangle inequality with a kind of regularity. This is part of what it *is* to have a visual experience, and thus needs not have a special account in any theory. Secondly, there are diachronic experiential regularities. These are regularities that are only partially (and to a very small extent) accounted for by the triangle inequality. The following is such a proposal: "when you have sensations of moving at the same speed your sensation of walking lasts longer when you have the sensation of walking from a to b to c than when you have the sensation of walking directly from a to c."(Vogel, 2014)

The latter requires all sorts of what McCain and Vogel call "contingent regularities". It is worth noting that no demon or anything can produce an experience that synchronically violates the triangle inequality. However, with respect to the diachronic violations, all the ways that the demon world could go wrong are ways in which the real world (meaning simply a spatial world) could go wrong as well. A demon who (McCain is right to point out) has the ability to cause a deception without causing it in accord with a mental geometry is no more capable of producing such nonsensical phenomenal worlds than a spatial world is. For example, whereas a demon could just nonsensically produce experiences and visual stimulations randomly,

so could a set of quantum events in the brain or along the optic nerve. So could slight changes to the behavior of photons(or, for the purpose of simply changing the world to get rid of Vogel's particular case, slight changes in how time perception works). So to the assertion that there are more contingent explanatory regularities required by the demon hypothesis the answer is not necessarily that these are not contingent regularities in the demon case. The point is that there are just such contingent regularities in the hypothesis that there is a spatial world which are required to account for the regularities Vogel and McCain discuss. And just as there is a regularity that explains why the demon continues to cause experiences according to his imaginative demon plan, the universe still behaves (perceptually) in accord with a spatial geometry. It is not clear that there are extra regularities required by demon scenarios or computer skeptical scenarios.

But this is only to provide some initial considerations which call into doubt the worry McCain has for the real world hypothesis. Toward the end of the next chapter I will consider this worry in some more detail.

6.5   Skeptical Scenarios' Dearth of Advantages

In the process of his defense, Vogel also claims that skeptical scenarios lack advantages to make up for disadvantages he takes himself to have pointed out. He says:

> Finally, as I will suggest below, the RWH enjoys certain explanatory advantages over skeptical hypotheses. There is no reason to suppose these advantages are offset by unspecified *disadvantages* attaching to the RWH, which a structurally different skeptical competitor will somehow manage to avoid.(Vogel, 2005, 76)

Vogel thinks that there are no advantages to be had by skeptical hypotheses. This assertion is false, but it will be more appropriate to discuss this question after having set out on the table BonJour's proposal for the advantages of the real world

hypothesis as well as attempts by others to bolster Vogel's and BonJour's arguments. In the following chapter, BonJour's antiskeptical argument will be treated in detail, in which case the question of whether skeptical scenarios enjoy advantages over real world hypotheses will be revisited.

Chapter 7

BonJour's IBE Against Skepticism

7.1 Introduction

Laurence BonJour gives a differently structured inference to the best explanation argument against skepticism. Rather than taking, with Vogel, certain particular geometrical regularities of visual experience to make the difference in the argumentation, BonJour takes a broader view of sense data and considers the crucial part of the argument to be the fact that skeptical scenarios are of a special type that is intrinsically inferior to nonskeptical scenarios. The following chapter will present and criticize this argument and similar attempts.

7.2 The Sense Data

Suppose one takes a normal six-sided die close to oneself (but not too close) and turns it around with one's fingers. There will be a large number of two-dimensional images in one's experience, all of which are consistent with being two-dimensional projections of the cubical object. At any point while the die is being turned around, the image received will be consistent with the two-dimensional projection of the cube at that point on its rotation. These data are what BonJour (with Price) calls "nuclear

sense data."(BonJour, 2003) Further, less core data are understandable as modifications of these nuclear sense data. Crucially, the visual data as of a cubical shape are consistent with our other senses' experience of these qualities. For example, the feeling I get in my hand when I turn the die around also seems consistent with its being a cube.[1] Even other senses, such as our sense of sound, are affected by the shapes of objects. These are all very interesting data that call out for explanation.

7.3 Analog Vs. Digital Explanations: BonJour's Case

BonJour's argument depends on splitting all of the skeptical scenarios into two camps. He distinguishes analog vs. digital explanations. Before getting into the details of what this distinction amounts to, we should outline the work BonJour intends this distinction to do for him. The upshot of his argument is to capture all of the classical skeptical scenarios that we think of as falling under the heading "digital" explanations. He then takes himself to show that there is something wrong with digital explanations as such (when compared with analog scenarios). But then the category "analog" as applied to an explanation doesn't settle whether it is skeptical or not. The real world hypothesis would count as an analog explanation of our experiences. However, BonJour asserts that, in fact, no analog *skeptical* scenarios exist. But if all of the skeptical scenarios are digital and thus inferior to the real world hypothesis which is analog, then BonJour has shown that the real world hypothesis is superior to skeptical competitors.

The case gets more difficult in attempting to characterize the distinction between analog and digital scenarios in a way that will do the work BonJour wishes it to do. Consider this first pass at describing analog scenarios. Analog scenarios are in important ways very similar to primary and secondary quality explanations.[2] Analog

---

[1]The phenomenon of it not being an instantaneous realization that, say, the thing that felt like a cube also looks like a cube for people who have recently gained sight gives me pause about saying this is a necessary connection between the two kinds of sensation.

[2]Primary quality explanations are those according to which the cause of our experiences resemble

scenarios will explain experience in in one of two ways. Analog scenarios may explain our experiences in the way primary qualities (like shape, extension, and distance) do in the real world hypothesis. If such is the explanation of our experiences then, arguably, we are necessarily talking about a real world hypothesis. Secondly, and more importantly, analog scenarios may explain our experiences in the way that secondary qualities (like color, taste, or sound) explain our experiences in the real world hypothesis. The suggestion from BonJour is that one might wish to construct a skeptical scenario that is in some ways very similar to our secondary quality explanations of color, taste, and sound experiences.

Digital scenarios are supposed to take up the rest of the explanations of our sense experiences. Digital scenarios are those which appeal to a representation of a world and a translation mechanism by which those representations cause our experiences as of a world.[3] All the classic skeptical scenarios are meant to fall under this umbrella. Classic cases where we are the subject of a deception by an evil demon or in which we are a brain-in-a-vat subject to the matrix-like experiments of a mad scientist are paradigm digital scenarios for BonJour.

BonJour concludes that the nature of digital explanations makes them intrinsically inferior to analog explanations. He argues that a digital scenario's very plausibility:

> ...depends in effect on the truth of two claims: first, that the corresponding analog explanation could indeed account for the experience in question (that a material world could produce experience of the sort that we in fact have), since the digital explanation works by emulating the action of the cause or causes postulated by the analog explanation; and, second, that

---

the experiences they cause. The secondary quality explanations are those explanations of our experiences which, in the real world hypothesis, don't resemble our experiences, such as the explanation of color experience.

[3]I mean to exclude in this description those explanations which simply amount to a theistic picture of the world. That is, suppose there is a God who represents a world of three-dimensional objects and then creates it. The universe he creates then causes experiences just as in the real world hypothesis. The only difference between this scenario and the real world scenario is that we posit a theistic cause of the universe. If BonJour's distinction is any good at all it will not count such a hypothesis as a digital scenario.

the specific translating mechanism postulated by the digital explanation in question can indeed successfully do the job of emulation (that God or the computer could indeed produce the sort of experience that would be produced by the represented material world).(BonJour, 2003, 95)

In order to test the mettle of this distinction, we will begin by discussing skeptical cases that are easiest for BonJour to address. After presenting the easy cases we will move on to more difficult cases until we've moved past the distinction's (and the argument's) usefulness.

First, suppose there is a universe much like the world in the real world hypothesis. Now, suppose that a copycat demon then represents that universe to himself and causes experiences matching that universe in some conscious being. Suppose that the conscious being is not in that universe. This scenario[4] is clearly an analog scenario. Further, the case is also very clearly far less simple than the real world hypothesis. BonJour would be right to treat the copycat demon case as untenable.

Secondly, in the category of slam dunks for BonJour's argument, consider the classical brain-in-a-vat case. In a classical brain-in-a-vat scenario, there is a physical world in which a mad scientist has decided (for some scientific reason or other) to perpetuate a massive deception against various living envatted brains which he has hooked up to a computer. The scientist, based on the world he experiences, writes a computer program. This computer program simulates (for the envatted brains) the experiences a real world would cause in a normal human being. This case is clearly a digital explanation. It requires the mad scientist to mentally (or at least on a computer disk) represent a three-dimensional world. It also requires both something like a real world hypothesis to be true (except for its causing *our* experiences) *and* it requires the mad scientist to have set up a translation mechanism from his program to our envatted experiences. This scenario also exhibits an intolerable complexity and

---

[4]We will leave aside the question of whether it is in fact skeptical. Nothing important turns on it being so, but this will end up being one of the borderline cases briefly discussed in chapter 2.

should be rejected.

Returning to demon scenarios, we will come to a case that BonJour seems to have as one of his primary targets in the quote provided above. Consider a classic demon scenario, where a demon has represented a world and telepathically produces the experiences as of this world in an individual conscious being. In this case we have a paradigm case of a digital scenario. The kind of case that BonJour makes in his quote above will be more difficult to make than in the previous two cases. It's not that the demon needs an *actual* world and a *represented* world along with an experience producing mechanism.

But the point is that there is an important asymmetry between the real world hypothesis and the demon world hypothesis. The demon world hypothesis seems to require that a real world would produce the experiences which the demon produces. The demon is representing a world and producing experiences based on the world he mentally represents. A counterfactual that is specified enough will turn out necessary (a world with these kinds of laws and those objects, etc.). But the real world hypothesis doesn't seem to require that the demon world would produce the experience which the real world causes. This, it is alleged, makes the demon world less simple than the real world hypothesis.

Further, it's not clear that the demon even *needs* it to be the case that the universe he's representing *would* cause the kinds of experience he's causing. There are all kinds of crazy worlds that a demon might represent and cause experiences with consistent 3D-projection regularities. Suppose the demon thinks about a world in which there are no physical/mental causal laws. Or a demon might represent a dark world and then imagine that world had lights. There are all sorts of cases where the world *wouldn't* produce such experiences which the demon might represent, tweak, and then cause experiences. This kind of world might actually be impossible. For example, the demon might cause experiences in accord with the geometry of a world which has no

physical/mental causal laws.[5]

But the truths about which worlds would cause which experiences are arguably necessary truths. There are a number of worlds which count as real worlds which would produce the same kinds of experiences. That is, there are Newtonian worlds, Einsteinian worlds, Einsteinian worlds with slightly tweaked laws, and so on. But it seems that there is a settled group of worlds that match the description of the world the demon is representing. But since this is a necessary truth (that this or that set of RWH worlds would successfully cause the relevant 3D experiences the demon is mimicking) it need not contribute to the complexity of the demon scenario.

One might argue, instead, that the source of the complexity is the demon's having to *know* this truth. The demon will need to understand that this is the case and perhaps *this* will be what makes the demon scenario less simple than the real world hypothesis. The demon requires his represented world and the causal mechanism by which his representations cause our experiences, but it seems that he also needs to *know* that the world he's representing would cause the experiences he's causing. But if we're to fault the demon scenario for being too complex in virtue of an item of knowledge possessed by the demon, one will need to argue that this knowledge doesn't come free with the existence of the demon. It certainly would come probabilistically free with an omniscient demon's existence. But it *might* also come free with a detailed representation of a world and an even relatively intelligent demon. Or at least it will be highly probable given these two things. So suppose a demon represents and mentally simulates a real world in a very detailed way. But running this mental simulation will involve mentally simulating humans and brains and the way that physical objects affect brains. If we posit demons that are omniscient with respect to their own mental

---

[5]Although troubles will come up here for a classical understanding of counterfactuals, since every counterfactual about what would happen in an impossible universe will be trivially true. This will mean there will be no unique true counterfactual about what experiences a universe with no physical/mental causal laws would bring about. Perhaps this is evidence against the classical understanding of counterfactuals. Then again, perhaps the suggestion that a demon can cause impossible world experiences is misguided.

states and add this to the current representative demon case we have a case where the demon's knowledge of the important facts are not an additional complexity in the case.

But suppose that the argument makes the real world scenario simpler than the demon scenario. An easy out for this kind of view is to make our demon scenario instead a god scenario. Appealing to a god to rescue this kind of skeptical scenario will make concerns of lack of simplicity go away. This is because a god would be a necessary being, who is omniscient and omnipotent. It's a necessary truth that God exists and it's a necessary truth that he knows any true proposition.[6] This includes all the propositions about what experiences the real world would cause in every particular situation for every particular orientation of physical objects. In fact a god need not even imagine a world or represent it in more than perhaps one time slice. A god would know, from the god's first representation of a world with me in it all the possible ways that my life could go and the exact experiences such a world would feed me after every decision. But a deity that is also omnipotent would also have no need of a contingently existing "translation mechanism" because it's a necessary truth that an omnipotent deity's actions happen when willed. One need not appeal to extraneous contingent entities or processes to make this decision come true. A god world seems to be substantially simpler than a real world.

From here on the cases will become more difficult for the distinction to capture. Suppose, for example, that there is a computer which is running a simulated matrix world program. Suppose further that this computer is hooked up to envatted brains, feeding them experiences based on its simulation program. In contrast with the standard brain-in-a-vat scenario above, the computer is not designed or programmed. Instead it has always existed (or, say, it arose as a result of some quantum event). The eternal computer scenario, so constructed, is not a digital scenario. That would

---

[6]There is important Chisholming to be done with the concept of omniscience, but none of that will have an effect on the point here.

require representations. But the computer no more has representations of a world than a colony of ant which happens to form in a way that looks like Winston Churchill represents a person.(Putnam, 2000) But it's difficult to intuitively characterize the eternal computer scenario as an analog scenario. It seems much closer to digital scenarios than it is to ordinary secondary quality explanations (like our ordinary explanation of color experience). All that's different from this case and the classic brain-in-a-vat case is that the computer in this case wasn't designed. But the digital scenarios are intentionally more precisely characterized than analog scenarios. In fact it is the trait BonJour picks out for them (scenarios that produce experiences by representing a world) that is meant to be their downfall according to his argument. If we are to modify the analog/digital into something that is more probably exhaustive we should hold the notion of a digital scenario to its technical characterization. We should allow the eternal computer scenario (and any scenario that fails to meet the criteria for a digital scenario) to fall into the category "analog scenario" simply in virtue of failing to be a digital scenario.

But this scenario will escape BonJour's argument. There are no representations of a world or a need for a translation mechanism. There is a mechanism by which our experiences are caused, but the real world has this as well. Below, in the discussion of Hasan's development of BonJour's argument, we will consider an argument that this scenario is a less good explanation than the real world hypothesis.

There is also a demon scenario that doesn't appeal to representations of a world in order to explain our experiences. Suppose there is just a demon who has phenomena and uses those phenomena to cause our experiences. That is, suppose a demon just wants to produce interesting and orderly experiences in some other conscious beings, so he produces a massive orderly set of images a la Berkeley's God. He then feeds this set of experiences into conscious beings as a sort of interactive painting for our enjoyment. One can even conceive of a demon having a very detailed 3D

186

phenomenal imagining[7] that follows laws which he then uses to cause experiences. One might also imagine the demon enjoying getting individuals who only get 3D vision by combining 2D images with parallax (along with shading tricks etc.) to experience his interactive 3D artwork. This demon scenario is not digital, and thus is not covered by BonJour's argument, though it may have vices that make it less good than the real world hypothesis. One might, for example, accuse this kind of hypothesis of failing to achieve the the explanatory "richness" of the real world hypothesis. Or one might allege that this fails somehow to explain why the experiences we have aren't much wonkier if there is no picture of a world limiting the demon's work of art. But suppose the demon somehow has intuited a Berkeleyan understanding of sense experience and takes it as a challenge to create such an experience without representing a world. This might mitigate the force of such a worry.

But there are also more *classically* analog skeptical scenarios which BonJour hasn't adequately discussed. The first thing that BonJour dismisses as implausible analog skeptical hypotheses are two-dimensional and one-dimensional worlds, since such worlds are not complex enough. However, I'm not sure that a two-dimensional world is all that problematic to construct. Suppose we are wanting a two-dimensional world to model a finite three-dimensional world.[8] All we need is an infinite two-dimensional world, with portions of the two-dimensional world being cross-section like portions of what a 3D world would be like. Further, suppose there are rules of interaction such that each slice (all of which are on one two-dimensional plane) act in the same way as a corresponding cross-section of the quasi-common sense world. This would be a two-dimensional world that is structurally such that it would produce experiences

---

[7]A phenomenal imagining in the Humean sense that an imagining here is just a less vivid phenomenal experience than ordinary sensation.

[8]Such a three-dimensional world may also need to be a quantized world. That is, it may need to be that matter in the three-dimensional world mimicked in our two-dimensional world will need to not be infinitely divisible. Further, if there are worries about constructing a purely two-dimensional world that can mimic a three-dimensional world one may simply think of the world as a very thin, flat universe that nevertheless is in three dimensions.

just like the quasi-common sense hypothesis despite being analog and (reasonably, since there are no three-dimensional objects) being a skeptical scenario.[9]

Despite the fact that there are skeptical scenarios which fail to fall in the desired place in BonJour's taxonomy, BonJour has at least given an interesting, if flawed, argument that a certain class of skeptical scenarios should be considered inferior to the real world hypothesis. Nevertheless, there are a number of analog scenarios to contend with, including one which should be called a classically analog scenario.

But isn't at least the two-dimensional world scenario above as good a candidate as the three dimensional world scenario? If there is anything to be said against the two-dimensional view it will have to be the fact that, intuitively, the natural laws in the two-dimensional world will have to be more complex and thus that this view will be less probable than the quasi-common sense hypothesis. There will need to be laws relating to the interaction of each slice with each other, along with analogues of the physical laws in the three-dimensional world. But it's not entirely clear *how* much more complex this will be, since the requirement is merely for *analogues* of the laws in the real world hypothesis plus the extra law relating slices. I say extra law instead of laws, since we can make a fairly simple systematic law that will relate each slice to its systematically related 2D slices ordered in the world in a systematic spatial way (e.g., There is a bottom plane starting in position $\{0,0\}$, with plane 2 (mimicking the slice of the 3D world just above the bottom plane) starting in position $\{0, 3\,000\,000\}$, and plane 3 starting in position $\{0, 6\,000\,000\}$ with a *single* law relating everything in a slice to the slice starting three-million away from it). It's possible that the two-dimensional world will have to be more complex in the final accounting, but it's not a case that has been made. BonJour or defenders would have to argue that

---

[9]If the finally constructed case has these piecemeal laws making each cross-section behave as it would in a three-dimensional world, then perhaps there would be a case that the laws are less simple. However, it's not an open-and-shut case against this scenario. It may seem that the laws would be less simple, but I'm not sure why it can't be the case that there are just analogously simple 2D laws in this 2D universe, but with the result that this world and its laws end up mirroring a 3D real world.

(1) the laws which are analogues to the laws in the 3D world will necessarily be at best equally simple (perhaps in the mathematical sense described in chapter 5) to the laws to which they are analogues. If this is the case, then the added laws relating the 2D slices to one another will push the real world hypothesis ahead in terms of simplicity. But another strategy may be more interesting. One might instead just adopt the quantitative approach and count the laws, asserting that there's always going to be *one* more law in the two-dimensional world than in the three-dimensional world. But suppose we were perfect scientists looking down on this two-dimensional world? Would it really be the case that the simplest way to describe the behavior of the 2D world as a collection of 3D world analogues plus a bridge law? I suspect that this is not the case. Probably for such a universe there will be a more simple way to express the world's behavior that puts the real world and the 2D world in league with the 3D world in terms of *number* of laws.

The 2D world scenario, though new and interesting, is not quite as clearly good as the sophisticated demon/god scenarios and scenarios like the eternal computer scenario, which are clearly more difficult to dismiss for the BonJourian.

But there is an interesting point and problem glossed over in many of the discussions (and especially in this one). This is the problem of *how,* on the real world hypothesis, subjective experiences *are* explained. Especially if we do not rely on complex science in making an inference to the best explanation, there is a real problem of how the real world hypothesis explains subjective experiences. This is evidenced by the vast number of different attempts to either explain away or deny the existence of matter/mind causation. This is the hard problem of consciousness. In the real world hypothesis there is a gap between the world and our experiences that needs to be bridged by a translation mechanism. Whatever this translation mechanism in the real world hypothesis, this doesn't need to be simulated in the digital explanation. All the digital explanation needs is its own translation mechanism. It doesn't need to

simulate the translation mechanism in the non-skeptical world. This will be discussed more closely after a discussion of Hasan's interesting addition to the debate.

7.4   Hasan's Additions to the IBE

In Hasan's forthcoming article (Hasan, 2015), he gives a very detailed presentation of the data in antiskeptical IBE's as BonJour understands them, and then takes a new approach to the Vogel/BonJour antiskeptical argument. He points out several flaws in various other defenses of such IBE's, finally giving his understanding of the real reason to prefer the real world hypothesis to skeptical hypotheses.

Hasan raises several worries for the versions of the antiskeptical argument presented above, some of which mirror worries raised above. In the end he argues that there are *different* aspects of the skeptical scenarios which fail to match up with the real world hypothesis in terms of explanatory quality. He concedes the point that the analog/digital distinction fails to wipe away all demon/computer scenarios with one swipe. However, he does take the distinction to significantly help the real world hypothesis along. He uses the analog/digital distinction to distinguish demon/God scenarios from computer scenarios and addresses each in turn. Hasan spells out the problem with the computer skeptical hypothesis as follows:

> Compare the CSH to the RWH. Corresponding to each three-dimensional object invoked by the RWH, the CSH posits a file or portion of a computer disc with certain magnetic properties. Corresponding to each projection produced by each object, the CSH posits another file, activated by the first, with different magnetic properties. Let O be an object with shape S and varying locations or positions $L_1...L_n$ relative to the observer, corresponding to which the CSH invokes file F(O) with fixed magnetic properties E(S), and "inputs" $I(L_1)...I(L_n)$. The latter "inputs" could be regarded as further magnetic properties of the computer. Let $P_1...P_n$ be projections of O resulting from changes in position $L_1...L_n$, corresponding to which the CSH invokes files $F(P_1)...F(P_n)$, which are produced or activated by F(O) given $I(L_1)...I(L_n)$.

To explain why $P_1...P_n$ result from O and inputs $L_1...L_n$, the RWH posits straightforward laws or processes of projective geometry; the very same laws or processes govern how all the projections are produced. To explain why $F(P_1)...F(P_n)$ result from $F(O)$ and inputs $I(L_1)...I(L_n)$, what laws or processes could the CSH posit? One possibility is for the CSH to posit that each of $F(P_1)...F(P_n)$ results from $F(O)$ and the inputs $I(L_1)...I(L_n)$ respectively. Each $F(P_x)$ would be explained piecemeal by $I(L_x)$ and a causal or lawful relation between $I(L_x$ and $F(P_x))$. The RWH, on the other hand, allows us to regard each P_x as a projection of a particular, more or less stable three-dimensional shape in motion relative to the observer; the existence of each P_x results from L_x and a relation between L_x and P_x, *but each instance of the latter relation is accounted for in a unified way, as a projection relation.* We are assuming that CSH and RWH are isomorphic in Vogel's sense (see section 3.1): each particular object, projection, event, etc., has a counterpart in the CSH; the difference is that the RWH can treat the many token processes or relations between these particulars as exemplifications of projection. The skeptic who infers from this isomorphism that the two hypotheses share all explanatory virtues ignores the virtue of being able to regard various phenomena as exemplifications of the same kind or type of process or relation.(Hasan, 2015, 32-33)

Hasan argues that a real world hypothesis naturally places all cases of 2D projection of 3D objects as instances of the same kind of process. Before addressing the specific argument for the RWH having this specific virtue, some comments are in order regarding the nature of this virtue.

Earlier in his article, Hasan asserts that the virtues the RWH will have over skeptical scenarios will be one of four main types of virtue (explanatory breadth, depth, and type and token simplicity). There are two different virtues which Hasan might be here attributing to the RWH. One natural option is to say that Hasan's argument, shows that the RWH has greater type simplicity than the CSH because the RWH appeals to fewer "types...of contingent entities, processes, causal or lawful regularities"(Hasan, 2015, 22) than the CSH.

But this can't be the virtue of the RWH on this account, especially in a world where a significant portion of our experience of the world involves experiencing rendered 3D images on computer screens. The RWH picture includes processes of the type which Hasan suggests for the CSH, so in the RWH the world is no more type simple than the CSH.[10] The real world is a world of computers that engage in 3D rendering, just like the CSH. We have experiences as of computers displaying 3D images on a two-dimensional screen, engaging in processes mimicking the way 2D projections of 3D objects arise in the real world.

However, perhaps it is being alleged that the RWH, instead, has the virtue of explanatory depth. On this way of interpreting the argument, the RWH, by making all of the instances of supposed projection exemplify {being a projection}, thus allowing the real world hypothesis to be deeper than the CSH. On Hasan's characterization of the virtue, this must mean that the RWH now adds another explanation of a contingent explainer that is unavailable to the CSH. The RWH's cases of projection are (partially) explained by their being exemplifications of the universal {being a projection}. But however many types the processes in the CSH are, presumably they will all be explained by the universal they each exemplify. This can't be the alleged virtue of the RWH either. Neither of the remaining virtues in Hasan's account are reasonably construed as what Hasan is here alleging as the RWH's advantage.

But suppose that what is alleged of RWH *would* in fact be a virtue. It should be discussed whether the RWH has such a virtue over against the CSH. It seems that the computer skeptical hypothesis projection cases all are exemplifications of the computational process called 3D-rendering. A computer skeptical hypothesis posits a computer doing a sophisticated version of the process many computers engage in today called 3D-rendering. This is a process-type that all the cases of projection in the CSH exemplify just as projection is what is exemplified in the RWH case. The real

---

[10]It might, in fact, make the CSH *more* type simple than the RWH, but there's no need to argue this in detail here.

world hypothesis does not have such a virtue advantage over the computer skeptical hypothesis. But Hasan dismisses demon and God scenarios by appealing to what I call the final simplicity challenge. More should be said about the allegation that a stable act of volition to produce a deception is required by demon scenarios and thus pushes demon and God scenarios into intolerable complexity.

7.5   The Final Simplicity Challenge

The objections that remain to skeptical scenarios will, in one way or another, amount to an assertion that a demon scenario will always be in some respect at least slightly less simple than real world hypotheses. Ali Hasan and Kevin McCain will point to the demon's intention to produce the deception, while Tim McGrew will point to the *existence* of the demon as that one extra thing which makes the demon/god case less simple than the RWH. As this is a very important point to consider, it deserves its own section.

7.5.1   Things Get Complicated, and That's Okay

McGrew acknowledges that complexity can be okay in cases when the more complex explanation explains what it explains better than the more simple explanation. In order to show how this applies to the case at hand, consider the following vignette:

The Teleportation Trick: Suppose Hugh attends a magic show. He's a fairly seasoned magician, so he knows his stuff. He sees a magician preform a teleportation trick. That is, he sees the magician step into a box and in a very short time step out of another box on the other side of the stage. Hugh has eliminated ordinary magician's tricks in this case. Now Hugh could believe either that the magician on stage has invented a teleportation machine, or that the magician on stage is one of a pair of (very) cleverly concealed twins both of whom are skilled at doing magic tricks.

Famously, in the movie *The Prestige*, Hugh Jackman's character believes the former and goes on a long self destructive quest to duplicate the same trick. But this is an irrational conclusion assuming Hugh also has the option of concluding that there is a twin who is just as good at magic as the magician on stage helping to

make this trick work. The twin hypothesis better explains the case, despite being less quantitatively simple. This is because, despite arguably being simpler according to **QuantSimp6** than the twin hypothesis, the teleportation hypothesis has other vices. The teleportation hypothesis appeals to a new type of process for which we don't already have precedent. That is, we don't already have reason to think that teleportation happens.[11] But the twin hypothesis is *much* less simple, to be sure. In the case in the movie, the magician that Hugh observes even comes out of the second box with the same injuries as the first magician, including amputations and scars. It is a *very* elaborate deception by the twins. But we still think the twin case is the better explanation primarily because the teleportation process is one for which we have no precedent (and probably have a lot of defeaters for this process's existence). But, it seems the twin hypothesis would be preferred even if we took away some of the defeaters (like the fact that if the magician had invented a teleportation device he would have sold it or been abducted by the government rather than performing in some two-bit magic show). Still the teleportation theory is less good in the final reckoning because of its affront to what I call qualitative simplicity above in chapter five.

But the principle illustrated by the teleportation trick case applies to the case of demon/god scenarios granted (for the sake of argument) that demon and god hypotheses are slightly less quantitatively simple than the real world hypothesis. The causation of phenomenal mental states by the physical is needed for the real world hypothesis. Such causation is meant to be analogous to teleportation in the case of the teleportation trick. Prior to being justified in believing the real world hypothesis we have no precedent for physical/mental causation. But we do have precedent for mental/mental causation and even for experience of images as a result of the decisions of a mind (i.e., our own when we choose to imagine or daydream). We already have

---

[11]This follows our characterization of qualitative simplicity above.

good reason to believe there are mental/mental causal laws, but without already having justification for believing in a physical world, we have no antecedent reason to believe there are physical/mental causal laws.

But what about the *tu quoque* for the demon hypothesis? Isn't the demon hypothesis just as bad as the real world hypothesis? This is the classic kind of response to such simplicity-based arguments for idealism. One simply tries to point out ways in which the process, e.g., mental/mental causation in which a demon causes an experience in another mind, is *also* a very exotic process just like physical/mental causation. Here's the case that it isn't. The demon scenario draws on a process for which there is precedent. There is precedent for minds causing phenomenal experiences. We can cause ourselves to have phenomenal experiences by our volition. One's deciding to imagine a green star shape has caused one to imagine a green star in the past, etc. This is a very ordinary experience.

The natural rebuttal to this is to point out that we don't have a precedent for minds causing *other* minds to have phenomenal experiences, and this makes the demon hypothesis on equal footing with the real world hypothesis. They're both cases of processes that lack precedent. But I'm not sure this is the case. There is at least precedent for the *kind* of causality that is posited by the demon hypothesis. That is, there is at least precedent for mental things causing other mental things. At the very least we can see that mental things can explain phenomenal stuff in a very ordinary way. It's perfectly acceptable to have mental explanations of the mental. I'm not here to say that it wouldn't be better for the skeptic if we somehow had prior justification for believing that demons have caused phenomena in human beings, but it remains that having precedent for mental/mental causation makes the demon scenario more probable. That a demon causes us to have phenomenal states is made more probable by the fact that there are minds that have causal powers to produce phenomenal states. The demon scenario explains better than the real world

hypothesis.

But what of charges that there are mitigating factors to the demon scenario's goodness? The Demon Scenario, it is hypothesized, would require the demon to have such a powerful and different mind that all bets are off regarding it's being a good explanation. The idea would be that the precedent that exists wouldn't apply to such a very different case. I've always found this kind of response odd. I think that the point of them is to suggest that there's a problem of specifying what type we should apply to a case to which one is applying this sort of argument. That is, why shouldn't we think of the type of the demon causation as demon-mind-causation instead of as falling under the more general mind-causation type. We don't have antecedent reason to think that there is demon-mind-causation. But I think there is a solution to this problem. The more specific the type one has precedent for which also applies to the case at hand, the greater the advantage conferred. So, for example, in the demon case we get closer by having antecedent reason to think that there are mental causal laws than we can in the real world hypothesis case since we have no precedent at all for physical-mental causal laws.

7.5.2 The Qualitative Simplicity Response: Or Simplicity and Probability

Another response that doesn't require digging one's feet in on the quantitative simplicity issue is to appeal to the classic charge that scenarios such as the demon scenario are *qualitatively* simpler (in a different way) than scenarios like the real world hypothesis. The charge is that idealistic scenarios are qualitatively simpler than the real world scenarios and thus preferable on that front. There is some reason to think that demon scenarios are simpler than the real world scenario. After all demon scenarios appeal only to minds, whereas the real world scenario must appeal to physical objects. But this kind of response has gone out of favor, in some part, due

to the fact that few hold really dualist views of the world nowadays. However, the precise final metaphysics of mind won't matter quite so much on the understanding of qualitative simplicity discussed in chapter 5.

The version of qualitative simplicity in Chapter 5 is more a reinterpretation of what's happening when one appeals to qualitative simplicity. My assertion was that when one reasonably appeals to qualitative simplicity, it is because one already has reason to believe that a certain kind of entity exists which makes explanations appealing to this kind of entity preferable to explanations which do not so appeal. For example, if one already has reason to believe that there are physical objects, it raises the probability that the explanation of some new phenomena will be physical as opposed to appealing to some other kind of entity for which one has no independent reason to believe exist.

In the context of discussing whether one is rational to conclude to the existence of the non-skeptical external world, we clearly are giving ourselves independent reason to believe in the existence of minds. Thus, one should prefer mental explanations(at least without countervailing reasons) of new phenomena until such time as one has justification for believing that there are material objects.[12] The challenge facing this advantage is the question of whether it can avoid McGrew's "greater confirmation for simpler theories" argument. The foundation of McGrew's probabilistic point is as follows:

> Let $0 < P(X/B) < 1$, and $0 < P(Y/B) < 1$, where B is a coherent set of background information and X is independent of Y. Then
>
> $P1$ $P(X/B) > P(X\&Y/B)$
>
> or, to put it verbally, the probability that X is true, given B, is greater than the probability that *both X and* some other claim Y are both true, given that B. And if, in addition to the above conditions, $P(X/B) <$

---

[12]For example, we forgo explanations appealing to teleportation devices without really good reason.

$P(X/B\&e)$, while $P(Y/B) = P(Y/B\&e)$, then evidence e confirms X to a greater extent than it confirms the conjunction of $X\&Y$:

P2 $P(X/B\&e) - P(X/B) > P(X\&Y/B\&e) - P(X\&Y/B)$.

Under the relevant conditions, theorems $P1$ and $P2$ hold for any coherent probability function.(McGrew, 1995, 134)

This is a trivial probabilistic point, but it's important to the case as McGrew sees it. McGrew, to leave out some complications for now,[13] treats the requirements of the real world hypothesis and the requirements of the demon world scenario as roughly equally probable on our background information, excluding the fact that the demon exists (or, as Hasan would prefer, that the demon has a stable intention to deceive). Suppose we substitute for X above the requirements of the demon scenario (minus the demons existence) and for Y the demon's existence. We will get that the fully fledged demon scenario will be less well highly confirmed than a real world hypothesis (which we are treating as probabilistically identical to X for this argument).

The problem is that whatever the background information, the probability of the demon hypothesis will be *at least* slightly raised by the fact that we have good reason to believe that there are such and such mental states (which, presumably, e gives). The better evidence we have that there are mental states, the more likely we are in a demon world, since a demon world is only possible in a world where there are mental states. Even if we grant that the real world hypothesis is more simple than the demon scenario, the greater confirmation point has not been established, since $P(Y/B) \neq P(Y/B\&e)$.

Further, $P1$, which requires the assumption that X is independent of Y, is suspect when applied to this case as well. In McGrew's case $X\&Y$ is the full Ramsey sentence for the demon hypothesis. Roughly speaking X is what the Ramsey sentence for the real world hypothesis and the Ramsey sentence for the demon hypothesis have in

---

[13]In his actual case McGrew first applies the probabilistic point to the Ramsey sentences for RWH and demon scenarios and then discusses whether this makes the RWH theory (whose Ramsey sentences he's discussing) more highly confirmed.

common . Y is what is added to X to yield the demon hypothesis (i.e., the hypothesis that there is a demon). But surely X isn't independent of Y in this case. Surely the existence of a deceiving demon makes it *more* likely that we have not only experiences in general but experiences that are orderly in the very kind of way that the real world hypothesis' experiences are. $P(X/Y)$ is higher than $P(X)$, so McGrew (and the simplicity fan who wishes to take advantage of McGrew's insight) has also failed to show that the probability of the (admittedly) simpler real world hypothesis is greater than the probability of the demon scenario. The real world hypothesis is in no great position by being one unit simpler than the demon scenario. This second point is true in spades if we substitute the demon's intent to deceive for the demon's existence. The probability of having real-world-hypothesis-style experiences is higher supposing a demon intends to create a deception. But then it remains unestablished that the demon hypothesis is less well confirmed than the real world hypothesis.

7.5.3   Complicated Demons? Some Doubts Raised

There are a couple of places one might pick to argue against the simplicity of demon or God hypotheses. McGrew puts the added complexity in requiring the Demon to exist, but Hasan and McCain, rightly, place the complexity rather in the Demon's intention to deceive. Arguably where the real world hypothesis requires a demon to exist, the real world will require the universe or some other equivalent exist. The demon world does in fact require demon intentions on top of the demon's existence and this, at least *prima facie*, gives good reason to think that the demon scenario is less probable by being less simple than the real world hypothesis.

However, I'm not convinced that this is the right way to cut the theories. Yes, the demon needs to have stable intentions and those intentions need to be successfully implemented, but the analogues of the laws and entities of the real world hypothesis are made more probable once the intentions are in place (this is especially so in the

case of a God hypothesis or a hypothesis involving an infallible demon with an unfrustratable will). The demon/god hypotheses are cases where not all is equal. It remains unestablished that the demon hypothesis is less virtuous, all things considered, than real world hypotheses.

But this, along with the advantages of the various skeptical hypotheses discussed in the previous chapter, seems to indicate that the real world hypothesis and skeptical hypotheses are on par, at best. At worst skeptical scenarios are better off than the real world hypothesis. Perhaps it's time to start worrying about where the failure of IBE's against skepticism leaves us.

Chapter 8

Exploring Other Options

We've failed to vindicate the real world hypothesis with an IBE from our experiences. All hope may seem lost, but in this chapter we will devote our attention to finding the faintest glimmer of epistemic hope in this barren wasteland of skeptical arguments. If one remembers, in Chapter 1, the skeptical argument which is the target of our considerations goes as follows:

(P1) If S's evidence regarding p makes p no more probable than it does not-p, then S is not justified in believing that p.

(P2) S's evidence regarding p makes p no more probable than it does not-p.

(P3) S is not justified in believing that p.

But the reader will notice that all that has been argued in the above chapters is that the inference to the best explanation arguments given in the literature to date fail to establish the probability of the real world hypothesis. The careful reader will immediately notice that this doesn't mean that the skeptical argument of this form is unassailable. One might naturally take the argument to in fact be something like the following:

(P1) If S's evidence regarding makes p no more probable than it does not-p, then S is not justified in believing that p.

(P2*) If S has no strong inference to the best explanation argument against skepticism, then S's evidence against skepticism makes it no more probable than its negation.

(P3*)There is no strong inference to the best explanation argument against skepticism.

(P4) S is not justified in believing that skepticism is false.

This argument, with true premises, would get the skeptical conclusion from the failure of the IBE's established above. But putting the argument in this way will make it easier to make clear the avenues of response left by the preceding work for those of us strongly committed to antiskepticism.

8.1 Against (P3*): Making IBE work

One might, in light of a Moorean commitment to antiskepticism, choose to reject the thought that the previous chapters have shown that IBE's against skepticism currently on offer fail. Hopefully, for my case, it turns out that the above is well argued. But there are other ways one might reject (P3) and avoid a skeptical solution to the argument.

First, one might propose a new addition to classical IBE's that avoids the difficulties with the above arguments. Perhaps there's something else about digital scenarios that makes them radically implausible, or perhaps there's another virtue that the real world hypothesis has over competitors. The antiskeptic may spend his energy here to avoid the skeptical conclusion.

But one might also attempt to first level the metaphysical playing field by modifying one's view of material objects. If we adopt something like a causal theory of objects, then perhaps we can establish antiskepticism by making the goal something less hardcore than a full-bodied real world hypothesis. Some remarks are in order regarding the prospects of a causal theory of objects on this front.

8.1.1 Causal Theory of Objects

If we can define material objects as those objects, whatever they are, which cause certain experiences under certain conditions, we might be in good shape. This will make the antiskeptical project much easier, it seems. As long as there is a consistent cause of certain kinds of experiences then we will be able to conclude that there are material objects that are roughly the shape that we take them to be and avoid skepticism.

Now there are a number of issues that could come up here. One might expect a lengthy treatment on the proper conceptual or linguistic analysis of the notion of a material object. I claimed earlier, in chapter 2, that a skeptical scenario is one in which material objects (as we experience) must be eliminated. One might expect discussion of any view which is "easier" epistemically to collapse into discussion of which reductions of physical objects are plausible and which not. Being an antiskeptic is centered closely around whether one's view allows an elimination of material objects.

But I don't think it's necessary to take this question on directly, even granting that in giving a philosophical analysis of physical objects we are giving a linguistic or conceptual analysis.[1] But there is a point about meanings that I'd like to make. If we self-reflect, it is striking how much better we know how to *use* and *apply* terms than we know the definitions or the meanings of our terms. This is why, for example, the Norman thought experiment (BonJour, 1985) should be seen as a good argument for epistemic internalism. But then this will allow us to decide, at least, on the antiskeptical prospects of views which would lower the metaphysical standards for antiskepticism.

But then there is a very simple argument to be made:

1. A view of the nature of objects will either (a) call cases like demon scenarios and

---

[1]I'm personally not particularly enamored with the linguistic and conceptual construals of the philosophical project, but here is not the place for this discussion.

matrix scenarios worlds in which there are material objects which are roughly as we take them to be or (b) it will not.

2. If (a) is true on a view of the nature of material objects, then the view of material objects is false.

3. If (b) is true on a view of the nature of material objects, then it doesn't avoid the issue of creating an IBE that will conclude against demon and matrix scenarios.

We can rest premise 2 on the strong intuitions about demon scenarios and matrix scenarios as scenarios which eliminate the material objects we experience. The strong intuition itself, like our intuitions in Gettier cases and in Norman-style cases, is enough to provide strong reason to accept that such scenarios are skeptical.

In epistemology there is a view called particularism, which is the view that, *prior* to knowing what method we ought to use to distinguish the good beliefs from the bad beliefs, we have access to the fact that certain particular judgments are good ones. On this view, then, the method of distinguishing good beliefs from bad beliefs must preserve the result that the particular judgments are good which we naturally judge so. There is also a distinction between particularism and methodism in the analysis of knowledge, which Fumerton (2008) discusses. But there are two kinds of view which might be described as particularist regarding the analysis of knowledge. Some particularist approaches take specific particular cases of our judgments that we know or are justified and require an analysis of knowledge to get the correct answer with regard to these particular judgments. But this can be problematic. Suppose I look at a broken clock and thereby conclude that the time is 2:30. It would be unfortunate if we required our analysis of knowledge to call *that* belief knowledge, despite it being a rather intuitive case from the internal perspective. But there's another kind of particularism which uses particular thought experiments to aid in the development of an analysis. In this kind of case one fills out a scenario in which one would or would not categorize a belief as justified or known. This kind of particularist would think that particular thought experiments are more fundamental to the project

of analysis for us. There is a similar approach that one might make in metaphysics where what one prioritizes is the capturing of the truth of particular judgments before one prioritizes the status of general principles.

In the end, the person interested in defending antiskepticism will have to defend a general view of the nature of the physical objects we experience. This will be needed so that we may judge the various borderline cases and decide whether, in the end, such scenarios are skeptical or not. But I think it's reasonable to think that there are some skeptical cases which we can tell are skeptical scenarios[2] prior to having a top-down theory of the nature of material objects.

This is just like the fact that one may be justified in believing that Norman, in the following case, doesn't have a justified belief that the president is in New York, prior to being convinced of the truth of internalism:[3]

> *Case 4.* Norman, under certain conditions which usually obtain, is a completely reliable clairvoyant with respect to certain kinds of subject matter. He possesses no evidence or reasons of any kind for or against the general possibility of such a cognitive power or for or against the thesis that he possesses it. One day Norman comes to believe that the President is in New York City, though he has no evidence for or against this belief. In fact the belief is true and results from his clairvoyant power under circumstances in which it is completely reliable.(BonJour, 1985, 41)

We should take ourselves to have strong evidence that Norman isn't justified in the above case, whether or not we have a worked out view of justification in mind. Similarly, we have strong evidence that the classic demon and matrix-style scenarios discussed in the above chapters are skeptical scenarios, whether or not we have worked out views of the nature of physical objects. In fact, the worked out views of physical objects will have to capture the intuition that such scenarios are skeptical. They are scenarios which eliminate the physical objects which we experience.

---

[2]and thus scenarios which eliminate the material objects which we experience

[3]That is, this isn't the special problematic sense of particularism which Fumerton says internalists can't maintain. It's the reasonable particularism of analysis by way of thought experiments.

But then the argument holds. Proposing a view of objects which would make anti-skeptical arguments easier will undermine the view of the nature of objects proposed. But then no view can make it *easier* to get out of skepticism. A view might make it harder (say a view which would call some of the borderline scenarios skeptical), but we have enough to worry about as it is.

One set of arguments against this assertion will be instructive, however. Berkeley argues that the idealist God scenario is not a skeptical scenario.

> Phil. I entirely agree with you, as to the ill tendency of the affected doubts of some philosophers, and fantastical conceits of others. I am even so far gone of late in this way of thinking, that I have quitted several of the sublime notions I had got in their schools for vulgar opinions. And I give it you on my word, since this revolt from metaphysical notions to the plain dictates of nature and common sense, I find my understanding strangely enlightened, so that I can now easily comprehend a great many things which before were all mystery and riddle
>
> Hyl. I am glad to find there was nothing in the accounts I heard of you.
>
> Phil. Pray, what were those?
>
> Hyl. You were represented in last night's conversation, as one who contained the most extravagant opinion that ever entered into the mind of man, to wit, that there is no such thing as material substance in the world.
>
> Phil. That there is no such thing as what philosophers call material substance, I am seriously persuaded: but if I were made to see anything absurd or skeptical in this, I should then have the same reason to renounce this, that I imagine I have now to reject the contrary opinion.
>
> Hyl. What! Can anything be more fantastical, more repugnant to common sense, or a more manifest piece of skepticism, than to believe there is no such thing as matter?
>
> Phil. Softly, good Hylas. What if it should prove, that you, who hold there is, are by virtue of that opinion a greater sceptic, and maintain more paradoxes and repugnancies to common sense, than I would believe no such thing?[4] (Berkeley, 1979)

---

[4] We'll need to be careful here, since Berkeley often takes the term matter as a term of art

Notice that Philonous' statement of his initial argumentative strategy does *not* attempt to show that Berkeley's view is *the* antiskeptical view. Instead, Philonous attempts to argue that the idealist view is able to retain *more* of common sense than the view that there are material objects. I think this is the right instinct to have (in different places he seems to be more enamored with the idea of the *right* view of the nature of material objects being his idealist view). The idealist view should be thought of as eliminating material objects and thus not antiskeptical in our sense. It may, nevertheless, be that an idealist picture is "less skeptical" in that it can preserve more of common sense than views which posit material objects.

Now this core datum will make something like the particular kind of IBE above necessary if we are to engage in an IBE defense of antiskepticism. This is because different views of material objects will either "make it easier" or they won't. If they would make it easier, they hold a faulty view of the nature of material objects. If they don't make it easier (by make it easier I mean make the conclusion we have to reach with an IBE a less risky thesis), then we have to use IBE to rule out these scenarios anyway.

8.2   Against (P2*): Disjunctivism and Direct Realism

Disjunctivism and direct realism are both options for how changing our theory of perception may effect the prospects of getting out of skepticism without the need for an IBE (or at least for the specific IBE traditionally given in the literature). Disjunctivism and direct realism are theories of sense perception that each propose that what we get in perception is something far richer than what's been granted to the antiskeptic so far.

Direct realism,[5] on my view, claims that in veridical sense perception the subject is

---

referring to a specific set of philosophical commitments regarding bodies. Nevertheless, the points that Berkeley makes regarding the distinction between skepticism and a view being *more* or *less* skeptical propose a useful lesson in the discussion here.

[5]Here I am concentrating on metaphysical direct realism. There are a number of views called

directly acquainted with material objects(or their parts, or their surfaces).[6] Disjunctivism, on the other hand, is often seen as a way of rescuing the kind of metaphysical direct realist view above. The disjunctivist, in response to a particular problem with direct realism which will be briefly discussed below, adds two theses to the direct realist's assertion. The disjunctivist asserts (i) that the state of awareness which occurs in veridical experience is a radically different from the state of awareness which occurs in non-veridical experience and (ii) that the states of awareness in veridical experience and in non-veridical experience are introspectively indistinguishable. That is, the disjunctivist will say that in hallucination one is in a very different kind of mental state (say, direct awareness of an experience with a certain character) than the one which would occur in veridical perception (direct awareness of a physical object for the direct realist). But one is incapable of introspectively distinguishing these different states.[7]

In the rest of this section we will first discuss realism's and disjunctivism's prospects for aiding the antiskeptic, granting their truth for the sake of argument (there would be no need for the antiskeptic to get bogged down in the details of whether the view is true if it won't hep address skepticism). After doing so, for any view that *will* help, I will raise some of the challenges for the view(s) in question. I will suggest briefly where the defender might go in giving a defense and give an initial prognosis for each particular view.

epistemic direct realism which defend the thesis that external world propositions are justified non-inferentially (e.g., Huemer (2001) and Goldman (1979) should each be considered epistemic direct realists). We are here discussing metaphysical direct realism because many of the epistemic direct realists who are not also metaphysical direct realists have been discussed above.

[6]This is a contentious claim as to the nature of metaphysical direct realism, but it's the version that has the best prospect, if true, of helping the antiskeptic.

[7]Now, characterizing disjunctivism as an attempt to rescue metaphysical direct realism is contentious. It's uncontroversial that disjunctivism may be a refuge for the erstwhile direct realist, but disjunctivists often characterize themselves in different ways. Further, contemporary direct realists and disjunctivists may dislike the use of acquaintance language, but I'm discussing here the particular versions of these views which I think have some hope of helping the antiskeptic out in our contexts. It may very well be, as far as we're concerned, that other versions of these views will have prospects which are not appreciated here, but that is no big problem for our discussion.

8.2.1 Antiskeptical Prospects

On metaphysical direct realism, I'm directly acquainted with physical objects when I have veridical perception. It's tempting to say that the only question left to ask is whether I've had any veridical perceptions. But there are skeptical scenarios consistent with my having had veridical perceptions in the past. For example, suppose a demon has since destroyed the world except for my mind and has begun feeding me a deception that seamlessly integrates with my past experiences.

Further, establishing that there has been veridical perception isn't enough to get the view that more straightforwardly implies epistemic direct realism. Richard Fumerton argues that several similar candidates for such views are consistent with views which are not so realist. He says:

> None of [the above proposals] entails ENR [which states that on at least some occasions we are non-inferentially justified in believing a proposition asserting the existence of a physical object]. (R1a) asserts that in all, and (R1b) asserts that in some, sense experience we are directly or immediately acquainted with constituents of physical objects. We can be directly acquainted with the existence of a part or a constituent of a physical object without being acquainted with the fact that the entity in question *is* a part or a constituent of a physical object, and thus without being acquainted with the fact that a given physical object exists. Put another way, our acquaintance with the fact that a certain entity exists may yield non-inferentially justified belief that the entity in question exists, but such acquaintance does not thereby yield knowledge that the entity is a constituent of a physical object, even if it is. The idealism of Berkeley, the early phenomenalism of John Stuart Mill and A.J. Ayer, and the views of the later Russell...are all consistent with the naïve realism defined by R1b and incompatible with naïve realism defined by ENR. (Fumerton, 1985, 74-75)

On the metaphysically but not necessarily epistemically direct realisms more work is needed to establish an antiskeptical view. If we are confident that we have acquain-

tance with such objects or constituents there still seems to be inferential work to do.

But it's not clear that this argument establishes that direct realism doesn't make the antiskeptic's job much easier, if true. Fumerton's R1a and R1b as well as the statement of direct realism talk of direct acquaintance with objects or their constituents. Facts aren't listed in the definitions, and as such, there is reason to expect that more justificatory work is needed before one establishes antiskepticism. One will need to be acquainted with or otherwise justified in believing that certain facts obtain. But Fumerton takes the views he describes to be asserting rather that we are at least occasionally acquainted with the fact that parts or constituents of physical objects exist. The barrier to justification as stated is that it's open what kind of thing exists in the fact with which we are acquainted. But this doesn't mean no advantage is conferred.

If we are acquainted with a fact, we are acquainted with a structured entity that mirrors the structure of the proposition it makes true (at least in simple cases). The fact is just that {there exists such and such an entity}. Whether the acquaintance here is to get justification for the existence of physical objects depends on what exactly is there in the "such-and-such" portion of the fact. If the "such-and-such" that I'm aware of in that fact is something with the properties we are requiring of the antiskeptical material object, then it seems the work is (at least very nearly) done. There may still be a gap depending on how general and what exact propositions are required to establish antiskepticism. It may be that an induction or IBE needs to be made (e.g., so that we can go from an individual case of a material object existing to the broader antiskeptical thesis), or it may be that some minor introspection on what counts as a physical object is needed for the direct realist to get justified antiskeptical belief (e.g., that this extended surface of a particular shape with which I'm acquainted meets the description of the kind of material object needed). Whatever the case may be regarding the extra epistemic work required, the antiskeptic's case seems to be in

a better position on this view than on a view that requires the IBE's discussed in Chapters 6 and 7.

Disjunctivism will treat the good case (i.e., veridical perception) as a case of acquaintance. Now, the disjunctivist, on this way of thinking, says that in veridical experience we are acquainted with physical objects but that these experiences and the experiences one would have in cases of non-veridical experience are introspectively indistinguishable. Despite claiming such indistinguishability, the disjunctivist claims that one has better evidence or justification in the veridical case than in the non-veridical case.

For example, consider the experience one has when one is in front of a red rose on a relatively sunny day. The disjunctivist will say that my experience of the red rose *highly justifies* me in believing that there is a red rose in front of me in virtue of the fact that I'm acquainted with this fact. But, with the critic of direct realism, the disjunctivist takes the experience of the rose to be indistinguishable to *us* from the experience we might have if we were, say, dreaming of a rose.

These claims don't seem to jibe well. Several critiques have been leveled against the disjunctivist which call attention to the fact that it seems incoherent to call radically different mental states which justify to different degrees indistinguishable.[8] But we shall, at least for now, hold back from critiques of disjunctivism as such and simply consider whether disjunctivism can help with skepticism if true.

At first, one will notice that all of the same initial worries for direct realism will apply equally to disjunctivism. The added epistemic work to justify an antiskeptical thesis will again vary depending on what exactly we are acquainted with in veridical experience. There may be some minimal introspection or trivial logical steps required like above.

There is a worry about philosophical assurance that is similar to the worry raised

---

[8]See, for example, the critique of disjunctivism and its purported advantages in addressing skepticism in Conee (2007)

earlier for the externalist's position with respect to skepticism. Earlier I discussed how an acquaintance internalist might respond to Bergmann's assertion that externalism and internalism are on par with respect to the charge. However, if the disjunctivist is correct that acquaintance is the sort of thing that can be indistinguishable from non-veridical (and, if the arguments in chapter 6 and 7 are correct, non-evidence providing) experiences, then it's hard to see how philosophically assured the individual in the good case will be with respect to external world skepticism.

But this charge of assurance-failure on the part of disjunctivism gets down to the view's big epistemic innovation. The most contentious epistemic point on disjunctivism, I think, is its needed insistence that a mental state can be more highly evidential than a state which is indistinguishable from it. This is despite urgings from some that the view is internalist.(Pritchard, 2011) This is at the very best an improbable consequence of disjunctivism, but it is also the very heart of disjunctivism's very ability to help the antiskeptic. Disjunctivism will need to concentrate on this assurance issue in order to develop disjunctivism as a reason to maintain hope.

Now, as most will agree and as we've briefly discussed above, direct realism makes the antiskeptical project much easier, if true. But then we would expect more people to be direct realists and short circuit skepticism in the process. The problem is that direct realism is nearly universally thought to be false, and there's a pretty good argument for it. Richard Fumerton nicely summarizes a strong version of this argument (usually called the argument from illusion or hallucination).

1) Any sensation or sequence of sensations S I have is such that it is possible that I have a hallucinatory experience so rich as to be phenomenologically indistinguishable from S.

2) My justification (if any) for believing some proposition P about the physical world as a result of my having S is the same as I would have were S a hallucinatory experience.

3 If S were a hallucinatory experience, my justification for believing P would not consist in my being directly acquainted with the fact that P (since by hypothesis

the fact that P would not exist.

It follows that:

4)My justification for believing P whether S is hallucinatory or not is not my being directly acquainted with the fact that P.(Fumerton, 1985, 79-80)

The point is if it's possible that there are experiences that are indistinguishable from the alleged direct acquaintance involved in veridical experience, then our evidence can't be better than the hypothetical hallucinatory experience it is indistinguishable from. Suppose I claimed to be acquainted with triangularity but then I discovered that it were possible to have experience phenomenologically identical to this acquaintance that was non-veridical. This along with an internalist commitment will get us that we have the same justification in both cases for whatever one claims is justified in the first case. But this means the justification one has for such triangular facts in our alleged acquaintance is no better than that in the non-veridical case. But this is the very point of the argument from the possibility of hallucination.

This is an important problem for the direct realist, and would eliminate our last ditch effort to find a reason for hope that the skeptical problem can be solved. There is room for fruitful research in the area though. One might attempt to hybridize direct realism and disjunctivism in such a way that retains direct realism's advantage against skepticism but is able to countenance the possibility of robust illusions or hallucinations. That is, disjunctivism seems to be a way of rescuing direct realism from problems like the argument from illusion, but then results in a very epistemically implausible view. The suggestion is that perhaps there is a way to *partially* move toward disjunctivism which will help mitigate some of the problems of each view. The key for the antiskeptic would be to successfully defend a kind of unidirectional indistinguishability of the good case and the bad case. That is, the hybrid view will have to defend that hallucinations are indistinguishable from veridical experience, but not the other way round. If this can be defended, I suspect that the radical

epistemic implausibility of disjunctivism can be mitigated while saving the advantages of direct realism. It's a faint glimmer of hope, but it's at least a direction for future antiskeptical research.

There will be other problems to solve. The most troublesome problem is the objection to direct realism and a disjunctivist development of direct realism which comes from a very plausible account of the physiology of perception. Perception is a long causal chain, and there are brain states that are the proximate causes of sensory experiences. It seems radically implausible to say that the proximate cause might remain the same (namely, thus-and-such a brain state firing) while the effect can be so radically different in hallucinatory versus veridical experience. The intuition is that if we have the same cause, we should have the same effect.(Fumerton, 2006) Any plausible defense of even the kind of hybrid metaphysical direct realism suggested above will have to find a way of addressing this worry.

Despite having ruled out the major IBE's against skepticism that are presented in the literature (and having raised worries for a number of other responses to skepticism), there remains at least a small light at the end of the tunnel. It will take future research to determine whether that light represents legitimate hope or merely the light from an oncoming train.

Appendix: Teleological Scenarios

One might try to set oneself up for success in IBE's against skepticism by giving reason simply to dismiss demon and Berkeleyan God scenarios as intrinsically inferior to scenarios appealing primarily to physical mechanisms. Vogel, for example, takes his main arguments to primarily address Brain-in-vat scenarios, and in later work hopes to dismiss mental (or teleological) explanations of our experiences as intrinsically inferior to causal explanations.

Since Vogel prefers the locution 'teleological' explanation, and since the use of such a locution will be particularly instructive here, it will be useful to do a little shadowboxing with the assertion that teleological explanations are inherently inferior to causal explanations (and thus that demon scenarios are inherently inferior). One line to take is to argue that teleological explanations merely stand in for causal ones and thus a causal explanation is always to be preferred if available. This is an interesting argument, but a set of distinctions is in order so that the argument can be adequately evaluated.

The teleological explanations that are not causal explanations at all are what we will call "functional explanations." For example, we say that a bird has wings so that it can fly. In this kind of explanation we explain something about the world by appealing to some function it serves. These sorts of explanation are most usually applied in biological contexts and in explanations about artifacts. A bird has wings so that it can fly and a toaster has heating elements so that it can warm bread.

Reasonably, this kind of explanation can be construed as merely standing in for

causal explanations. When we discover the factory process of toaster making or are given the evolutionary history of a bird's wings we no longer need appeal to the aforementioned teleological explanations and if the corresponding causal explanations are available one should reject the teleological explanation over the causal explanation now on order (at the very least this can be granted for the sake of argument).

But one might think that, even if we have the evolutionary history of the bird or the factory history of the toaster, there is a lingering teleology that isn't gotten rid of by giving the thing's causal history. It seems to be that "so that it can fly" is a perfectly adequate answer to the question of why a bird has wings. Further, it seems to be a perfectly adequate answer to the question of why a toaster has heating elements is "so that it can warm toast." The two cases, however, should be distinguished.

The bread case is reasonably thought to stand in for *another* kind of teleological explanation to be discussed below, which are those explanations which are called teleological because they appeal to the motivations or desires of conscious beings to explain facts about the world. Such explanations will be discussed briefly below.

But the wing case still seems to fundamentally be the kind of case we're discussing in this section. "So that it can fly" seems to be precisely the future-directed teleological explanation in question. There are various different things that one might say on this topic. For example, one might say, with Hempel, that this kind of explanation should be translated into assertions about what a creature needs to survive (or at least about the survival-conduciveness of a trait).(Hempel and Oppenheim, 1948, 144) However, one might, with Ruth Millikan, argue that functional explanations like the above wings case are actually stand-ins for a very specific kind of evolutionary history.[9] But either of these kinds of view will be consistent with the assertion that teleological explanations should be replaced by causal explanation. In the Hempelian

---

[9] See, for example, Millikan (1984) and Millikan (1989) for a more detailed outworking of Millikan's view of teleology.

case they should be so replaced because they're not explanations at all, while in the Millikan case they should be so replaced because such explanations are most fundamentally an assertion of a kind of causal history.

A second kind of teleological explanation that, though technically causal, is reasonably thought of as merely a stand-in for a causal explanation is what I will call "hidden powers" explanations. In these explanations some *explanandum* is 'explained' by appealing to some causal mechanism that is only spelled out teleologically. For example, the classic explanation of something's causing people to sleep by appeal to its 'dormitive powers' is an explanation of this kind. This kind of teleological explanation is merely leaving a gap for one later to insert a *bona fide* causal mechanism.

The third and final kind of explanation that are called teleological are explanations that appeal to the intentions and desires of an agent to explain some fact or entity. We can contrast the case of the toaster and a bird's having wings to develop a contrast of the two kinds of explanation in question. The toaster's explanation is fundamentally teleological in the current sense: the toaster's existence is caused by a human being with beliefs and desires coming up with a design plan and implementing that plan. The bird's having wings is rather explained (on the Millikan-style view) by various facts about the bird's history, with biological mechanisms and chance mutations accounting for the properties of the bird. In the case of the toaster, the explanation for why it has heating elements has to do, rather, with the desires and beliefs of its designers.

It's more difficult to make the crucial argument about this particular kind of 'teleological' explanation. It's important to point out that it's not really correct to describe mental explanations as anything other than causal explanations. They're simply causal explanations that appeal to different kinds of causes than the causal explanations that we describe as non-teleological.(Hempel and Oppenheim, 1948) On certain views of the world the design plans and thoughts and desires of intelligent

agents will yet be replaceable by reference to scientific and causal laws, that is, it might be that all this mental talk can just as well be replaced by talk of brains and brain processes.

In fact, one might argue, everything about the world that may seem to have an explanation that appeals to desires of conscious minds has a physical causal explanation that will replace any reference to the mental. This kind of view is fine and may be defensible in the end as far as we're concerned here, but the problem comes in attempting to use it to get around demon scenarios in skeptical arguments. If this is justified at all it would have to be justified posterior to the justification of our ordinary beliefs about external world objects (and science). It is science that can justify the belief that the things that happen in the world will always have a full physical causal explanation. But dismissing mental explanation from the universe depends on sophisticated brain study at the *very* least. In any case one will have had to already be highly justified in rejecting demon and other such 'teleological' explanations of our experiences antecedent to defending such a view of the world. Such arguments are thus inadmissible in ruling out mental explanations of our sensory experiences.

Bibliography

Bamford, Gregory (1999), "What is the problem of ad hoc hypotheses." *Science and Education*.

Barnes, Eric (1992a), "Explanatory unification and scientific understanding." *PSA: Proceedings of the Biennial meeting of the Philosophy of Science Association*, 3–12.

Barnes, Eric (1992b), "Explanatory unification and the problem of asymmetry." *Philosophy of Science*.

Beebe, James (2009), "The abductivist reply to skepticism." *Philosophy and Phenomenological Research*, LXXIX, 605–635.

Bergmann, Michael (2006), *Justification without Awareness*. Oxford University Press.

Bergmann, Michael (2008), "Externalist responses to skepticism." In *The Oxford Handbook of Skepticism*, Oxford University Press.

Berkeley, George (1979), *Three Dialogues between Hylas and Philonous*. Hackett.

BonJour, Laurence (1985), *The Structure of Empirical Knowledge*. Harvard University Press, Cambridge.

BonJour, Laurence (1999), "Foundationalism and the external world." *Philosophical Perspectives*, 13, 229–249.

BonJour, Laurence (2003), "A version of internalist foundationalism." In *Epistemic Justification: Internalism vs. Externalism, Foundations vs. Virtues* (Laurence BonJour and Ernest Sosa, eds.), 3–96, Blackwell Publishing.

Briesen, Jochen (2008), "Skepticism, externalism, and inference to the best explanation." *Abstracta*, 4(1), 5–26.

Butler, Keith (2000), "Problems for semantic externalism and a priori rrefutation of skeptical arguments." *Dialectica*, 54, 29–49.

Byerly, T. Ryan (2012), "It seems like there aren't any seemings." *Philosophia*, 40, 771–782.

Byerly, T. Ryan (2013), "Explanationism and justified beliefs about the future." *Erkenntnis*, 229–243.

Byerly, T. Ryan and Kraig Martin (2015), "Problems for explanationism on both sides." *Erkenntnis*, 773–791.

Carnap, Rudolf (1967), *The Logical Structure of the World: Pseudoproblems in Philosophy*. University of California Press.

Cartwright, Nancy (1979), "Causal laws and effective strategies." *Nous*, 13, 419–437.

Church, Ian M. (2013), "Manifest failure failure: The gettier problem revived." *Philosophia*, 41, 171–177.

Churchland, Paul M. (1981), "Eliminative materialism and the propositional attitudes." *The Journal of Philosophy*, 78, 67–90.

Conee, Earl (2007), "Disjunctivism and anti-skepticism." *Philosophical Issues*, 17, 16–36.

Cullison, Andrew (2010), "What are seemings?" *Ratio*, 23, 260–284.

Dougherty, Trent (2012), "Reducing responsibility: An evidentialist account of epistemic blame." *European Journal of Philosophy*, 20, 534–547.

Dummett, A. E. (1954), "Symposium: Can an effect precede its cause? I." *Proceedings of the Aristotelian Society*, 28, 27–62.

Eddington, Sir Arthur (1935), *The Nature of the Physical World*. Dent.

Ehring, Douglas (1985), "Simultaneous causation and causal chains." *Analysis*, 45, 98–102.

Fumerton, Richard (1980), "Induction and reasoning to the best explanation." *Philosophy of Science*, 47, 589–600.

Fumerton, Richard (1985), *Metaphysical and Epistemological Problems of Perception*. University of Nebraska Press, Lincoln.

Fumerton, Richard (1992), "Skepticism and reasoning to the best explanation." *Philosophical issues*, 2, 149–169.

Fumerton, Richard (1995), *Metaepistemology and Skepticism*. Rowman & LIttlefield Publishers, Inc.

Fumerton, Richard (2006), "Direct realism, introspection, and cognitive science." *Philosophy and Phenomenological Research*, 73, 680–695.

Fumerton, Richard (2008), "The problem of the criterion." In *The Oxford Handbook of Skepticism* (John Greco, ed.), 34–52, Oxford University Press.

Fumerton, Richard (2011), "Epistemic internalism, philosophical assurance, and the skeptical predicament." In *Knowledge and Reality: Essays in Honor of Alvin Plantinga* (Thomas M. Crisp, Matthew Davidson, and David Vander Laan, eds.), Springer.

Gendler, Tamar Szabó (2008), "Alief and belief." *Journal of Philosophy*, 105, 634–663.

Gifford, Matthew B (2013), "Skepticism and elegance: problems for the abductivist reply to cartesian skepticism." *Philosophical Studies*, 164.3, 685–704.

Goldman, Alvin I. (1979), "What is justified belief?" In *Epistemology: An Analogy*, 340–353, Blackwell Publishing.

Harman, Gilbert (1965), "The inference to the best explanation." *Philosophical Review*, 74, 88–95.

Harman, Gilbert (1973), *Thought*. Princeton University Press, Princeton.

Harman, Gilbert (1975), "Moral relativism defended." *The Philosophical Review*, 84, 3–22.

Hasan, Ali (2013), "Phenomenal conservatism, classical foundationalism, and internalist justification." *Philosophical Studies*.

Hasan, Ali (2015), "Skepticism and spatial objects." Forthcoming.

Hempel, Carl G (1965), "Aspects of scientific explanation." In *Aspects of Scientific Explanation and Other Essays*, Free Press.

Hempel, Carl G. (1966), *Philosophy of Natural Science*. Prentice Hall.

Hempel, Carl G. and Paul Oppenheim (1948), "Studies in the logic of explanation." *Philosophy of Science*, 15, 135–175.

Huemer, Michael (2001), *Skepticism and the Veil of Perception*. Studies in Epistemology and Cognitive Theory, Rowman & LIttlefield Publishers Group.

Huemer, Michael (2003), "Causation as simultaneous and continuous." *The Philosophical Quarterly*, 53, 556–565.

Huemer, Michael (2007), "Compassionate phenomenal conservatism." *Philosophy and Phenomenological Research*.

Huemer, Michael (2009), "When is parsmony a virtue?" *The Philosophical Quarterly*, 59, 216–236.

Jaynes, Edwin T. (1973), "The well posed problem." *Foundations of Physics*, 4, 477–492.

Jefferys, William H. and James O. Berger (1992), "Ockham's razor and bayesian analysis." *American Scientist*, 80, 64–72.

Kitcher, Philip (1981), "Explanatory unification." *Philosophy of Science*, 507–531.

Kline, A. David (1980), "Are there cases of simultaneous causation?" *PSA: Proceedings of the Biennial meeting of the Philosophy of Science Association*, I, 292–301.

Lange, Marc (2004), "Bayesianism and unification: A reply to wayne myrvold." *Philosophy of Science*, 71, 205–215.

Lehrer, Keith and Thomas D. Paxson Jr (1978), "Knowledge: Undefeated justified true belief." In *Essays on Knowledge and Justification*, Cornell University Press.

Lipton, Peter (2004), *Inference to the Best Explanation*, second edition. International Library of Philosophy, Routledge: Taylor and Francis Group, London.

Lycan, William G. (2007), "Moore's anti-skeptical strategies." In *Themes from G.E. Moore: New Essays in Epistemology and Ethics* (Susana Nuccetelli and Gary Seay, eds.), Oxford University Press.

Marinoff, Louis (1994), "A solution to bertrand's paradox." *Philosophy of Science*, 61, 1–24.

McCain, Kevin (2014), *Evidentialism and Epistemic Justification*. Routledge.

McCain, Kevin (2015), "Explanationism: Defended on all sides." *Logos and Episteme*.

McGrew, Timothy J. (1995), *The Foundations of Knowledge*. Littlefield Adams Books.

Millikan, Ruth Garrett (1984), *Language, Thought and Other Biological Categories: New Foundations for Realism*. MIT Press.

Millikan, Ruth Garrett (1989), "In defense of proper functions." *Philosophy of Science*, 56, 288–302.

Moore, G.E. (1966), "Four forms of skepticism." In *Philosophical Papers*, 220–222, Collier.

Mulnix, Jennifer Wilson (2011), "Explanatory unification and scientific understanding." *Acta Philosophica*, 175–194.

Myrvold, Wayne (2003), "A bayesian account of the virtue of unification." *Philosophy of Science*, 70, 399–423.

Nagel, Ernest (1961), *The Structure of Science: Problems in the Logic of Explanation*. Harcourt, Brace & World, Inc.

Nozick, Robert (1981), "Knowledge and skepticism." In *Philosophical Explanations*, 172–217, Harvard University Press.

Popper, Karl (1974), "Replies to my critics." In *The Philosophy of Karl Popper* (Paul Arthur Schilpp, ed.), volume 2, Open Court, La Salle, Illinois.

Poston, Ted (2014), *Reason and Explanation : A Defense of Explanatory Coherentism*. Palgrave Macmillan, Basingstoke, GBR.

Pritchard, Duncan (2011), "Evidentialism, internalism, disjunctivism." In *Evidentialism and its Discontents* (Trent Dougherty, ed.), 235–253, Oxford University Press, New York.

Putnam, Hilary (1973), "Meaning and reference." *The Journal of Philosophy*, 70, 699–711.

Putnam, Hilary (2000), "Brains in a vat." In *Knowledge: Readings in Contemporary Epistemology* (Sven Bernecker and Fred I. Dretske, eds.), Oxford University Press.

Quine (2004), "The nature of natural knowledge." In *Quintessence: Basic Readings from the Philosophy of W.V. Quine*, 287–300, The Belknap Press of Harvard University Press.

Reid, Thomas (1997), *An Inquiry into the Human Mind on the Principles of Common Sense*. Penn State Press.

Russell, Bertrand (1940), *An inquiry into Meaning and Truth*. W.W. Norton and Co., Inc.

Salmon, Wesley (1974), "The pragmatic justification of induction." In *The Justification of Induction* (Richard Swinburne, ed.), 85–97, Oxford University Press.

Salmon, Wesley (1988), "Statistical explanation and causality." In *Theories of Explanation* (Joseph C. Pitt, ed.), Oxford University Press.

Salmon, Wesley (2006), *Four Decades of Scientific Explanation*. University of Pittsburgh Press.

Sandberg, T. A. (2004), *Thomas Reid's Providentialist Epistemology*. Ph.D. thesis, University of Iowa.

Schupbach, Jonah N. (2005), "On a bayesian analysis of the virtue of unification." *Philosophy of Science*, 72, 594–607.

Shackel, Nicholas (2007), "Bertrand's praradox and the principle of indifference." *Philosophy of Science*, 74, 150–175.

Sober, Elliot (2002), "Instrumentalism, parsimony, and the akaike framework." *Philosophy of Science*, 69, S112–S123.

Sosa, Ernest (2007), *A Virtue Epistemology: Apt Belief and Reflective Knowledge, Volume I*. Oxford University Press.

Steup, Matthias (2011), "Evidentialist anti-skepticism." In *Evidentialism and its Discontents* (Trent Dougherty, ed.), 105–122, Oxford University Press, New York.

Taylor, Samuel (2015), "What seemings seem to be." *Episteme*, 12-3, 363–384.

Thagard, Paul (2007), "Coherence, truth, and the development of scientific knowledge." *Philosophy of Science*, 74, 28–47.

Turri, John (2012), "Manifest failure: The gettier problem solved." In *Virtue Epistemology: Contemporary Readings*, Massachusetts Institute of Technology.

Unger, Peter (1968), "An analysis of factual knowledge." *Journal of Philosophy*, 65, 157–170.

VanFraasen, Bas C. (1977), "The pragmatics of explanation." *American Philosophical Quarterly*, 14, 143–150.

Vogel, Jonathan (1990), "Cartesian skepticism and inference to the best explanation." *The Journal of Philosophy*, 87, 658–666.

Vogel, Jonathan (2004), "Skeptical arguments." *Philosophical Issues*, 14, 426–455.

Vogel, Jonathan (2005), "The refutation of skepticism." In *Contemporary Debates in Epistemology* (Matthias Steup and Ernest Sosa, eds.), 72–84, Blackwell Publishing, Malden, MA.

Vogel, Jonathan (2008), "Internalist responses to skepticism." In *The Oxford Handbook of Skepticism* (John Greco, ed.), 533–556, Oxford University Press.

Vogel, Jonathan (2014), "E&¬ H." In *Skepticism and Perceptual Justification* (Dylan Dodd and Elia Zardini, eds.), 87–107, Oxford University Press.

Wolterstorff, Nicholas (2001), *Thomas Reid and the Story of Philosophy*. Yale University Press.

Wronski, Leszek (2010), *The Common Cause Principle: Explanation via Screening Off*. Ph.D. thesis, Institute of Philosophy, Jagiellonian University.

Zagzebski, Linda (1994), "The inescapability of gettier problems." *The Philosophical Quarterly*, 44, 65–73.