

Modal Knowledge, in Theory

BY

ROBERT WILLIAM FISCHER

B.A., State University of New York at Geneseo, 2004

M.A., Trinity International University, 2006

M.A., University of Illinois at Chicago, 2010

THESIS

Submitted as partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Philosophy
in the Graduate College of the
University of Illinois at Chicago, 2011

Chicago, Illinois

Defense Committee:

W. D. Hart, Chair and Advisor
Colin Klein
Walter Edelberg
Daniel Sutherland
Karen Bennett, Cornell University

ACKNOWLEDGEMENTS

Marya Schechtman is doubly to blame for the existence of this dissertation: first, her seminar on personal identity raised the questions that got me interested in the epistemology of modality; second, it was at her suggestion that I began to study with Bill. I'm thankful that she led me down this path.

By and large, I owe the shape of this dissertation to my committee members. Bill pushed me to do constructive philosophy at a time when I was content with taking potshots at other people's work; I'm glad that he did. And Bill let me run with an idea that he still regards as essentially wrongheaded; I appreciate that freedom. Conversations with Colin helped me to figure out both my central thesis and the overall structure of my argument (not least by encouraging me to set aside a number of issues that were, ultimately, distractions). In particular, I'm grateful that he guided me through van Fraassen's *The Scientific Image*, which continues to exert a major influence on my thinking. Finally, Walter has been – and continues to be – a model of philosophical charity and rigor. My view would be far less sophisticated were it not for his example and explicit direction.

There are a number of other people who deserve mention: Karen Bennett, Valerie Brown, Brian Casas, Bucky Farley, Sam Fleischacker, Dave Hilbert, Peter Hylton, Charlotte Jackson, Chris McNiven, Josh Norton, Burkey Ozturk, Daniel Sutherland, John van Dyke, and Aleks Zarnitsyn. In ways direct and indirect, large and small, they all had a hand in bringing this project to its present stage. But I should single out Brian and Burkey. I would have left the program after my first year were it not for Brian's encouragement, and I would be a much less well-rounded philosopher were it not for my discussions with him. And without the tireless

sounding board that is Burkay, I would have struggled to generate and precisify all the ideas that wound up in this document. I'm very grateful for their friendship.

Finally, I thank my parents, my wife, and my son: my parents, for their unending support and encouragement; my wife, for her steadfast love; Jeremiah, for a reason to write.

RWF

TABLE OF CONTENTS

<u>CHAPTER</u>	<u>PAGE</u>
1. INTRODUCTION	1
1.1 Modal Knowledge, in Theory	1
1.2 Some Considerations	2
1.2.1 Benacerraf's Dilemma	3
1.2.2 <i>A Posteriori</i> Necessities	5
1.2.3 Taking Stock	8
1.3 An Overview	8
2. THE THEORY THEORY	11
2.1 Introduction	11
2.2 Motivating TT	11
2.2.1 OPTIMISM vs. PESSIMISM	12
2.2.2 The World's Rules	14
2.3 TT in Detail	17
2.4 Clarifying TT	25
2.5 Some Objections	30
2.5.1 The Cheating Objection	30
2.5.2 The No-Theory Objection	34
2.5.3 The Logic Objection	41
2.5.4 The Modal Collapse Objection	44
2.5.5 The Conflict Objection	46
2.6 Looking Ahead	51
3. CONDITION A	53
3.1 Introduction	53
3.2 Modal Content and Explanatory Virtues	53
3.3 A Deeper Challenge	57
3.4 Epistemic Value Monism	61
3.4.1 If Premise 1 Is True, Then Epistemic Value Monism Is True	61
3.4.2 Epistemic Value Monism Is False	62
3.4.2.1 Utilitarianism in Ethics and Epistemology	63
3.4.2.2 The Problem with TRUE	65
3.4.3 Back to van Fraassen and IBE	68
3.5 Some Models of IBE	69
3.6 Particularism and IBE	75
4. CONDITION C	79
4.1 Introduction	79
4.2 Benacerraf's Dilemma	79
4.3 Replies Not Pursued	86
4.4 Rejecting the Demand for Explanation	88
4.5 Hawthorne's Objection	91

4.6	Handling Hawthorne's Objection	92
4.7	Taking Stock	96
4.8	Conclusion	98
5.	MODAL SKEPTICISM.....	99
5.1	Introduction.....	99
5.2	Van Inwagen's Argument.....	101
5.3	Problems for the Disjunctive Argument	106
5.4	The 'Direct Evidence' Version of Modal Skepticism	111
5.5	An Alternative Proposal.....	114
5.6	Some Cases	121
5.6.1	The Problem of Evil.....	121
5.6.2	The Ontological Argument	125
5.7	Conclusion	130
6.	THE OPTIMIST'S OBJECTION	131
6.1	Introduction.....	131
6.2	Thomasson	134
6.2.1	Thomasson's View <i>qua</i> Modal Epistemology	135
6.2.2	Problems for Thomasson's View <i>qua</i> Modal Epistemology.....	139
6.3	Yablo	145
6.3.1	Imagining without Images	148
6.3.2	Imagining with Images	153
6.4	Williamson.....	156
6.5	Conclusion	162
7.	CONCLUSION.....	164
7.1	The Case for TT	164
7.2	Modal Skepticism and the World's Rules	164
7.3	Conservatism.....	167
7.4	Disagreement	170
7.5	Taking Stock	172
	REFERENCES	175
	VITA	182

1. INTRODUCTION

1.1 Modal Knowledge, in Theory

This dissertation is a contribution to the epistemology of modality. Consider the following proposition: *I went out for coffee this morning, but I could have stayed home*. That I could have stayed home is a modal claim. I am realist about modality,¹ so I think that the truth or falsity of this claim does not depend on what anyone thinks or believes or agrees to (whether implicitly or explicitly); indeed, I think that this claim true or false in whatever sense the following statements are true or false:

- Hydrogen has an average atomic weight of 1.00794 u.
- My mother is current president of the United States.
- Four is greater than two.

As it happens, I think that it's true that could have stayed home this morning and – what's more – that I am justified in believing as much. My aim here is to develop and defend a theory about the nature of this justification. In short, I maintain that the justification enjoyed by our modal claims is parasitic on the justification had by our best theories; we are justified in believing a modal claim only if we believe it because it's true according to a theory that we justifiably believe.² I call this view 'the Theory Theory' (TT). If we think of TT as a negative thesis, then it boils down to the slogan, 'no theory, no justification'.

(To be clear, TT is *only* concerned with justification. I take no stance on how (or whether) you would need to supplement TT to get an account of modal knowledge. For stylistic reasons, I occasionally use the language of entitlement, warrant, knowledge, etc. (as in the title of

¹ ...but not (necessarily) a Lewisian modal realist. By 'realism about modality', I simply mean any view on which there are modal claims with mind-independent truthmakers; this includes Lewis' view, versions of ersatzism that posit abstract objects of various kinds, etc.

² I am ignoring derivative sources of justification (e.g., testimony, memory, etc.) and any properly basic modal beliefs that we may have.

this dissertation). However, unless the context very clearly indicates otherwise – as in, e.g., Chapter 3 – this language is not meant to direct attention toward a different epistemic notion.)

I have two goals in this introduction: first, to outline general considerations that shape the chapters that follow; second, to give an overview of those chapters.

1.2 Some Considerations

This project – like any other – is shaped by a number of antecedent commitments. I have arguments for some of them; others I register as assumptions. In the interest of full disclosure, I should indicate what these commitments are.

Let me set the stage with four questions that fall within the purview of the epistemology of modality. (I'll put these questions in terms of our beliefs about what's possible, although they can all be asked in terms of our beliefs about necessities and counterfactuals.)

1. Under what conditions are we justified in believing that p is possible?
2. What constitutes evidence for p 's possibility?
3. What are we justified in believing about modal matters?
4. How can we resolve disagreements about modal matters?³

No doubt these questions are among the most important in the epistemology of modality (with no particular concern about the order). Success in modal epistemology – as in philosophy generally – is saying something true. But since it's hard to know whether you've said something true, saying something plausible often has to suffice. What makes a modal epistemology plausible? In short, a modal epistemology is plausible insofar as it answers (a) all these questions (b) in an elegant way (c) that fits well with our antecedent commitments.

In essence, I've just said that a theory is plausible to the degree that it scores well on

³ There are, of course, other questions that you could add to this list. Here are two interesting ones that are beyond the scope of my project. First, what's the relationship between our justification for believing that p is possible and our justification for believing perceptual, moral, mathematical, and scientific claims? Second, what (if anything) can be said to someone who denies that we are justified in believing at least some modal claims?

three explanatory virtues: generality, simplicity, and conservatism. And I do think that these virtues ought to be front and center when constructing a modal epistemology. As important as these considerations are, though, they are generic: every theory in every domain is plausible insofar as it scores well on the explanatory virtues (or so say I). What domain-specific assumptions shape my epistemology of modality? I'll answer this question by taking a quick look at two standard puzzles: first, the modal analogue of Benacerraf's Dilemma; second, the problem of the necessary *a posteriori*.

1.2.1 **Benacerraf's Dilemma**

Benacerraf's Dilemma is a familiar problem in the philosophy of mathematics. Paul Benacerraf argues that the semantics and the epistemology of mathematics pull in two different directions: the platonist offers a semantics for mathematics that's in line with the one that we prefer for other domains, but in so doing posits entities with which we can't interact; the logicist has a straightforward epistemology, but gets it by way of a paltry account of mathematical truth (Benacerraf 1973). There is an analogous dilemma concerning modal truth. Realists about modality offer the most plausible semantics for modal discourse (we will suppose), but they invoke entities with which we can have no causal contact; antirealists can offer straightforward epistemologies, but they manage this only by analyzing modal claims in ways that undermine much of their interest. In mathematics and modality, the demands of truth do not appear to match the demands of knowledge.

Whatever the appeal of realism about modality, it's tough to show that we *ought* to be realists. Given my purposes, I leave that project for others⁴: I will simply assume that we should

⁴ See (Divers 2007) for a good overview of the hurdles in front of the case for realism. For helpful accounts of the costs of antirealism, see (Hart 1989) and (Rea 2002). W. D. Hart points out that many of the classic problems in

be realists come hell or high water. In particular, I assume that no consideration from the epistemology of modality should lead us to give up realism about modality.

So here is the realist's horn of Benacerraf's Dilemma – or, rather, its modal analogue. Gettier cases show that knowledge is incompatible with (a certain sort of) luck. Perhaps the most attractive solution to the luck problem requires some causal commerce between the knower and the known, where this interaction explains the knower's epistemic success. But if realism about modality is correct, then we bear no causal relations to the truthmakers for modal claims. Therefore, if the realist can't provide an alternate solution to the luck problem, she makes our epistemic success unintelligible; on her view, it is unclear how we can have any modal knowledge whatever.⁵

There are three ways out of this argument: (1) you can deny that we are causally isolated from the modal facts, (2) you can find an alternate, non-causal explanation of our epistemic successes, or (3) you can reject the explanatory-*cum*-causal requirement on knowledge. In their own ways, (Armstrong 1989), (Koons 2000), (Rea 2002), (Molnar and Mumford 2003), and (O'Connor 2008) opt for one of the first two strategies. I think that the acrobatics required are not worth the result. So, I go the third route.

philosophy (e.g., the mind / body problem) are modal in character. To give up on modality is to deny that those problems are genuine. Michael Rea contends that it's hard even to understand the thesis that there are objects without taking on some modal commitments.

⁵ What, exactly, does the object of the belief need to explain? The belief's truth? I could be convinced that truthmakers cause propositions to be true, and if this is so, then possibilities can figure appropriately in the relevant causal explanations. Perhaps it's not the belief's truth that the object should explain, but its genesis. Recall, though, that even if the evil demon is deceiving me, I can have justified beliefs about trees and train stations. The objects of these beliefs are, *ex hypothesi*, non-existent objects (the demon is fabricating only their appearances). Since non-existent objects cannot cause anything, they cannot figure appropriately in causal explanations of my beliefs. So, the object of your belief need not figure appropriately in a causal explanation of that belief in order for it to be justified. One way out of this objection is to make a *de re* / *de dicto* distinction: if we are in a skeptical scenario, then we justifiably believe that certain propositions are true – i.e., we have justified *de dicto* beliefs – but we have no (or very few) justified beliefs about objects – i.e., we have no justified *de re* beliefs. But there are normal, non-skeptical cases in which it looks as if we have justified beliefs about non-existent objects: e.g., when we justifiably believe that Sherlock Holmes smokes a pipe. So, if this reply is based on the assumption that we never need to appeal to non-existent objects, and so they're an aberration in the present case, then it won't pass muster.

In so doing, I follow David Lewis. I will discuss the details of his view in Chapter 3. For now, let it suffice to say that Lewis does not respond by postulating some special faculty of rational intuition, or Russelian direct acquaintance, or what have you. Essentially, he puts Chisholm's particularism to work: we begin with the set of those beliefs that we take to be justified and we argue for epistemological principles by showing that they best explain the set's constitution. This *modus operandi* makes it wildly improbable that we'll reject paradigm cases of justified belief, since no principle is likely to be more secure than the exemplars for which it is supposed to account. So given platonism, we have to choose between our mathematical beliefs and a causal requirement on knowledge. It's obvious which to choose. Likewise, given realism about modality, we have to choose between our modal beliefs and a causal requirement on knowledge. No doubt we will choose the former every time, and rightly so.

Of course, all this is in need of defense, and I say much more about it in Chapter 3. I mention it now just to be clear about the paths that I *won't* pursue. I am not interested in tinkering with the metaphysics of modality or mind until the two fit together neatly. I grant that either Lewis' modal realism or some version of ersatzism is true; moreover, I grant one of the assumptions behind Lewis' response to Benacerraf: namely, that we have no special acquaintance with or immediate access to the truthmakers for modal claims. My aim is to develop an epistemology within these constraints.

1.2.2 A Posteriori Necessities

Let's turn now to the second standard puzzle in the epistemology of modality: the problem of the necessary *a posteriori*. Kant said, "[e]xperience teaches us that a thing is so and

so, but not that it cannot be otherwise.”⁶ If Kant was correct – as many seem to believe⁷ – then our modal knowledge must be *a priori*. But Kripke raised a serious problem for this view. Identities are necessary, and we know some identities *a posteriori*; it seems to follow that we know some necessities *a posteriori*. If Kant was right, then how can this be? And if Kant was wrong, then how does experience provide the knowledge in question? Surely we don’t *observe* necessities. What’s more, even if we did, would it be at all plausible that we observe non-actual possibilities?

These may not seem like legitimate worries. We establish particular identity claims via *a posteriori* methods and we establish the necessity of identity via *a priori* methods. Slap the two together and you have *a posteriori* necessities. Where’s the problem?

Let’s grant that we know *a priori* that identities are necessary. We are now trying to determine whether *a* and *b* are identical. If we were just trying to determine whether *a* and *b* are perfectly correlated in the actual world – i.e., that it is actually the case that *a* is always accompanied by *b* and vice versa – then there would be no issue (or none peculiar to this debate). But once we know that we are not simply trying to determine whether *a* and *b* are perfectly correlated in the actual world, but in every world, then it does seem reasonable to question whether the evidence for perfect actual correlation is strong enough to justify perfect correlation in all worlds: the change in the scope of the hypothesis surely warrants questioning the data on hand. Consider an analogy. Imagine that we encounter a group of birds in an otherwise empty field; we immediately note their bright orange plumage. This is probably evidence enough to justify the following biconditional: for all *x*, *x* is a bird in the field iff *x* has bright orange plumage. But Is this evidence enough to show that this sort of bird always and everywhere has

⁶ CPR, B Introduction, §II.

⁷ See, e.g., (Hart 1988), (Rea 2002), and (Tahko 2009).

bright orange plumage? Plainly not. (All the observed birds could be of one gender, with the other colored quite differently; or the color of the birds' plumage could change seasonally; etc.) But if our scant evidence won't justify the more general hypothesis here, then why is it enough when we are considering necessary identity statements?

The situation is worse yet. Many of the identities that Kripke highlights are scientific discoveries: water is H₂O; gold is the element with atomic number 79; heat is mean molecular kinetic energy. If these are *a posteriori* necessities, then no doubt there are others waiting in the wings, heretofore undiscovered – after all, no one takes present science to be complete. And if there are unknown others, then what does this fact do to our claims to *a priori* knowledge? If all our modal knowledge is *a priori*, then it seems that we always have a defeater for any claim to the effect that *p* is possible: for all we know, the claim runs afoul of a yet-undiscovered necessity.

What can we learn from these problems? First and foremost, not to side with Kant. If realism about modality is true, then I think that we should expect much – if not all – of our modal knowledge to be *a posteriori*. When a class of truths can't claim this distinction, we reasonably begin to suspect that the truths are really about us (our concepts, our language, our practices, our etc.), not about their ostensible subject matter.⁸ Given this, it's pretty clear how to manage the identity problem above. You can either (a) maintain that we *do* observe possibilities and necessities, thereby rejecting (a narrow construal of) Kant's position or (b) expand your conception of 'justified by experience' in such a way that, though possibilities and necessities aren't observable, experience is somehow essential to their justification. Since I don't see how to make the first option work, I think that the latter option is the better of the two. On my view,

⁸ This may partially explain the longstanding appeal of Quine's indispensability argument for mathematical objects; see, e.g., (Quine 1981a) and (Colyvan 2001).

experience can underwrite our modal knowledge, albeit indirectly: experience provides the data to be explained, and those explanations – in the form of well-confirmed theories – are what stand immediately behind our modal knowledge.

1.2.3 Taking Stock

So here, in short, are the commitments that shape what follows. Realism about modality is non-negotiable, and I am not questioning the view that we are causally isolated from the modal facts. Moreover, I have no interest in views that clear these two hurdles by postulating a special faculty or cognitive capacity or divine intervention. And even if we have some *a priori* knowledge, I have a strong preference for a view on which our modal knowledge is *a posteriori*. My aim here is to construct a general, simple, and conservative modal epistemology that respects these considerations.

1.3 An Overview

It's time to provide an outline of what's to come. Chapter 1 sets out TT in detail. I begin by explaining why, given realism, we should opt for a theory-based modal epistemology. I then introduce a particular account of theories – the so-called semantic view. The key component of this view is that a theory includes a set of models, many of which purport to represent possible states of the target system. If we are justified in believing that the theory is true, then we are justified in believing that those possibilities are genuine. This is the foundation of TT. Ultimately, I maintain that, where *p* is any modal claim:

[TT] You justifiably believe that *p* iff (a) you justifiably believe a theory according to which *p* is true, (b) you believe that *p* on the basis of this theory, and (c) you have no defeaters for the belief that *p*.

I close Chapter 1 by responding to a first round of objections.

TT places three conditions on justified modal belief. I have very little to say about Condition B, but Conditions A and C deserve considerable attention. Chapter 2 concerns Condition A: plainly, TT won't fly unless we have a suitable account of theory justification. I turn here to a defense of inference to the best explanation (IBE), the mainstay of scientific realists. (Although TT could be married to a different form of ampliative inference, IBE is the natural pairing.) After considering Colin McGinn's argument to the effect that modal claims can't be justified on empirical grounds, I turn to the two major complaints about IBE: (1) that employing the explanatory virtues amounts to making unjustified – and probably false – assumptions about the world and (2) that the explanatory virtues cannot provide epistemic reasons to believe that a theory is true, but only pragmatic reasons to accept a theory as true for certain purposes. I contend that these complaints are based on either (a) an untenable commitment to epistemic value monism, the thesis that there is but one valuable quality that beliefs can have, or (b) a simplistic view of what the virtues are and the role that they play in IBE. I conclude by sketching a positive story of IBE's epistemic credentials.

Chapter 3 looks at Condition C. We are not causally related to the truthmakers for modal claims, and many think that this provides a general defeater for the realist's claim to modal knowledge (see, e.g., (Sidelle 1989), (Nozick 2001), (LaPorte 2004), and (Thomasson Forthcoming)). What can the realist say in the face of this challenge? I argue that we ought to think about the challenge in terms of the threshold of epistemic risk that each side advocates. In this light, it becomes clear that if you level this objection against the realist, then you will probably have to make a number of other concessions to the Cartesian skeptic: e.g., denying that we know that we aren't brains in vats. And if the realist's claim to modal knowledge is no worse

off than her claim to know that we aren't in a skeptical scenario, then the realist needn't be too worried.

Chapters 4 and 5 set out and defend perhaps the most significant implication of TT: namely, modal skepticism. 'Modal skepticism' is Peter van Inwagen's term for what we might call modal *modesty* – i.e., the view that, while we have plenty of modal knowledge about ordinary matters, there are a great many extraordinary modal matters about which we are ignorant. If modal skepticism is true, then we probably don't know whether there could be naturally purple cows, or three-inch thick sheets of transparent iron, or complete brain transplants, or a maximally perfect being (where necessary existence is a perfection). Chapter 4 explores van Inwagen's version of modal skepticism and argues that it rests on a poor analogy between modal and perceptual knowledge. I then show that TT does a better job of securing modal skepticism.

Chapter 5 is a reply to those who think that we know more than modal skepticism allows, and hence that – if TT indeed leads to modal skepticism – TT ought to be rejected. To turn back this objection, I offer some considerations in TT's favor, since any argument for TT is thereby an argument for modal skepticism. Then, I contend that the three prominent modal epistemologies – (Thomasson 2007, Forthcoming), (Yablo 1993), and (Williamson 2007) – do not plausibly increase the scope of our modal knowledge beyond what TT provides.

Finally, in a brief conclusion, I take stock of the case for TT. I argue that it has a serious claim to be the best modal epistemology for realists about modality.

With this introductory material behind us, it's time to begin setting out TT.

2. THE THEORY THEORY

2.1 Introduction

This chapter has three parts. After offering a very quick sketch of TT, the first part is an attempt to motivate this view. In the second, I explain how TT works. In the third, I defend TT against a first round of objections.

2.2 **Motivating TT**

In short, TT says that the justification attaching to our modal beliefs depends on the justification attaching to our theories. Let's begin with our beliefs about what's possible.⁹ TT assumes that our theories have modal content; it assumes that theories say how things could be. (For now, let's not worry about what it is for a theory to 'say' that something is possible; I'll cash out the idea later.) Given this starting point, and setting aside justified inferences from actuality to possibility, TT offers the following account of justifiably believing that p is possible.¹⁰ Call the story 'POSS':

[POSS] You justifiably believe that p is possible iff (a) you justifiably believe a theory that says that p is possible, (b) you believe that p is possible on the basis of this theory, and (c) you have no defeaters for the belief that p is possible.¹¹

⁹ TT tells analogous stories for beliefs about necessities and counterfactuals, but those stories are complicated by a number of factors. I will discuss them after getting the basic framework on the table.

¹⁰ As I said in the introduction, TT is a view about justification. You could come up with an analogous view about knowledge, but TT is not it. I will often talk about knowledge, but unless the context clearly indicates otherwise, the terminological change is either for reasons of style or convenience.

¹¹ This is not exactly right, but it is close enough for present purposes. First off, POSS doesn't say anything about beliefs justified derivatively – say, through testimony. The right thing to say here is that POSS is an account of non-derivative justification – i.e., that more basic justification on which testimonial justification is parasitic. Second, POSS doesn't allow that we may have some properly basic modal beliefs – i.e., beliefs that we hold justifiably, but not based on anything else that we believe. I don't know that we have any properly basic beliefs (or whether any of them would be modal if we did), but I have no argument to the effect that we don't. So, for all I know, we justifiably believe some propositions to be possible despite the fact that we do not satisfy the right side of POSS with respect to those propositions. I am prepared to concede that POSS should be a more complicated biconditional that allows for properly modal knowledge, but I defend the above formulation on pragmatic grounds: I can't amend POSS to deal with this exception without making it rather unwieldy, and I doubt that this exception is what makes POSS contentious; so, it seems safe to ignore this complication in the present context.

I think that most philosophers would be willing to grant a weaker view that we might call ‘POSS-’:

[POSS-] If (a) you justifiably believe a theory that says that p is possible, (b) you believe that p is possible on the basis of this theory, and (c) you have no defeaters for the belief that p is possible, then you justifiably believe that p is possible.

POSS- offers only a sufficient condition for justifiably believing that p is possible, and it’s a pretty plausible one. POSS, by contrast, says that this sufficient condition is also necessary. In other words, it says that the justification attaching to our modal beliefs depends on the justification attaching to our theories.

2.2.1 OPTIMISM vs. PESSIMISM

Why adopt POSS over POSS-? I’ll answer this question in two stages. First, it’s worth noting that an epistemology of modality is inevitably going to be shaped by our take on the *difficulty* of acquiring modal knowledge. Here are two approaches to this question:

[OPTIMISM] We are in a fairly *good* epistemic position vis-à-vis the modal facts.

[PESSIMISM] We are in a fairly *poor* epistemic position vis-à-vis the modal facts.

Which of the above is most appropriate: OPTIMISM or PESSIMISM? Well, OPTIMISM seems reasonable if we are peculiarly suited to discovering modal facts. We might have this ability because of a fit between one of our faculties and the way that the modal facts are organized; you might take this to be the epistemic upshot of Hume’s view of the imagination – which is basically combinatorial – plus his ban on necessary connections between distinct existences (the modern version of which is the so-called “principle of recombination” defended in (Lewis 1986b) and (Armstrong 1989).)¹² Alternately, we might be suited to discovering

¹² Of course, if we are very bad at determining whether a candidate object is a real object – and therefore a distinct existent – then this won’t be the case. Let’s just assume – probably incorrectly – that Hume thought that we

modal facts if there is some fortuitous correlation between our modal judgments and modal reality; this is the essence of the theistic reliabilism that is defended in (O'Connor 2008), based on (Plantinga 1993). Finally, we might be suited to discovering the modal facts because they depend on us in some way or other – think, for example, of those who would reduce modal facts to semantic or conceptual ones.

I am realist about modality: I take the facts about what's possible to be just as sturdy and mind-independent as are the facts about the half-lives of isotopes. Moreover, I take it that there is an important sense in which these facts are inaccessible to us: I make no claim here about whether they are concrete or abstract, but whatever their nature, I do assume that we bear no causal relations to them; they are not among the things with which we are (causally) acquainted. I also deny that God – or any other supernatural force – is responsible for coordinating our faculties with modal reality. And finally, given these commitments, I see no reason to think it likely that modal reality is organized in a way that would be epistemically convenient for us – as it would be if, e.g., metaphysical possibility is just logical (or broadly logical) possibility.

With the above in the background, it's hard to see how we could be in a good epistemic position with respect to the modal facts. I should think that, *at best*, our epistemic situation vis-à-vis the modal facts is akin to our epistemic situation vis-à-vis the facts outside our light cone¹³: we will be able to make a few claims with a high degree of confidence, but most claims we will need to hold rather loosely, and there will be a great deal about which we must confess ignorance. If this is correct, then PESSIMISM looks like the right starting point: we should not expect to know a great deal about modal matters. Modal knowledge is hard to come by, and

are fairly good at carving reality at its joints.

¹³ On the assumption that no cause can travel faster than the speed of light, we can say, very roughly, that our light cone is the region of spacetime containing both (a) the events that could causally affect us and (b) the events on which we might have some causal impact. See (Huggett 2010).

much of what we have is tentative.¹⁴

If PESSIMISM is the prudent starting point, then we should be hunting for a duly cautious epistemology of modality. So, if there is a tendency to look askance at POSS for its conservatism, the above should serve as gentle corrective. And, of course, POSS is a more cautious epistemology than POSS-, since POSS- allows that there might be routes to knowledge of possibility other than the ones provided by our best theories. Still, we need more to show that POSS is preferable to POSS-; we should also want some reason to think that POSS- is *too* liberal, not just *more* liberal than POSS. Here, then, is the second stage of my reply.

2.2.2 The World's Rules

At this juncture it will help to think about *Clue*, that old murder mystery board game. The object of *Clue* is to be the first one to determine who killed Mr. Boddy. The game has six characters, all of whom are suspects. There are six weapons with which Mr. Boddy might have

¹⁴ N.B., I should note that there may be some temptation to reject PESSIMISM based on a mistaken understanding of the way that we employ the possible worlds framework. Possible worlds are often put to work to model broadly psychological phenomena: e.g., the meanings of sentences, the contents of beliefs, conversational presuppositions, agents' epistemic positions, etc. Moreover, it's hard to deny that we can utter meaningful sentences about naturally purple cows, or believe things about transparent bars of iron, or assume – in a given conversation – that there are such cows or iron samples. All this leads quite quickly to the assumption that there are possible worlds at which these objects exist – or at least that a commitment to these worlds should be the default position, placing the burden of proof on the detractor. Since the possibilities must be plentiful to account for these broadly psychological phenomena, and since PESSIMISM indicates that we are not in a position to assess whether the possibilities are plentiful, PESSIMISM gets the boot. But this line of reasoning is confused. True, we need to suppose that the worlds are plentiful. However, we needn't suppose that the *possible* worlds are plentiful; it will suffice if most of the worlds are impossible. Where these broadly psychological matters are concerned, the possibility is not what does the work. Rather, it is (a) a more-or-less combinatorial conception of the worlds, which allows the account – of sentence meaning, say – to be sufficiently fine-grained, plus (b) a notion of truth-at-a-world, possible or not, which allows you to distinguish the sentences by way of the worlds at which they are true. The possibility is dispensable. Of course, we would need to say that the possible worlds are plentiful if we couldn't make recourse to impossible worlds. However, if we restrict ourselves to possible worlds, there is a serious worry whether our accounts can be fine-grained enough to capture the structure that the relevant phenomena exhibits: e.g., we seem to be able to utter many different mathematical truths, all of which are necessary; however, without recourse to impossible worlds, it's hard to see how the possible worlds approach can distinguish them. (On this and related issues, see (Yagisawa 1988) and (Nolan 1997).) Impossible worlds allow us to dodge this worry. And, once we invoke them, it's hard to see why it matters that the possibilities are plentiful, since the impossible worlds can fill in wherever possible worlds would otherwise have been needed. This suggests that, where we are focused on broadly psychological phenomena, the possibility of the worlds is an extraneous assumption, and therefore one that we can do without. And, if we can do without it, then it doesn't create a serious problem for PESSIMISM.

been killed, and nine rooms in which the murder might have occurred. There are no restrictions on how suspects, weapons, and rooms can be combined. So if asked at the outset of the game whether Colonel Mustard could have killed Mr. Boddy, we have no trouble answering the question: it's certain that he could have. Moreover, we know how the murder could have occurred: Mustard could have ended Boddy in the Ballroom with the lead pipe, or in the Library with the noose, or whatever; there is no shortage of ways for Mustard to be guilty. Likewise, suppose that we are asked whether Mustard could have killed Boddy in one of the hallways that connect the various rooms. The answer in this case is equally clear: certainly not. The murder can only occur in a space on the board to which there corresponds a card in the deck; and, while there is a card for each room, there is no card for any of the hallways. Where *Clue* is concerned, we have modal knowledge in spades.

But consider this: if you didn't know the rules of *Clue* – and if, for some sad reason, no one would share them with you – how could you determine whether Mustard could have killed Boddy? I see only two options. On the one hand, you could try to figure out the rules by watching game after game; if you were successful, then you could use the rules to answer your question. On the other hand, you could watch game after game until you happen to see one in which Mustard is the guilty party. Put differently, you could either (a) develop and employ a theory of *Clue*'s gameplay or (b) make an inference from actuality to possibility.

Why are these options the only ones? In short, the rules of *Clue* fix what is and isn't possible in the game; there are no other factors at work. So, if you know all the rules, then you can assess the modal facts of the game. If you don't know any of the rules, then you have no grip on the factors that determine what is and isn't possible in *Clue*. So, if you don't know any of the rules, it's very hard to see why your judgments about unobserved states of the game

should have any epistemic credibility whatever.¹⁵ To be clear, I don't deny that it may *seem* to you that Mustard could have killed Boddy long before you can articulate your theory of the rules, nor do I deny that such seemings could be tolerably reliable. However, insofar as your belief about Mustard is justified, the justification is best explained by a partial grasp of the rules of *Clue*: the partial grasp explains both (a) why you're right when you're right, since your judgments are formed based on actual rules (or close approximations) and (b) why you're wrong when you're wrong, since your judgments are not based on all the rules, and so aren't sensitive to all the factors that fix the game's possibilities. So it's plausible that, insofar as you are justified in making a modal claim about *Clue*, your justification is tied to either (a) your developing theory of *Clue*'s gameplay or (b) an inference from actuality to possibility.

We've just outlined a modal epistemology for *Clue*, and I submit that it's a good model for our modal knowledge generally.¹⁶ The world has rules too: logical, mathematical, natural, and metaphysical. Like the rules of *Clue*, these rules define possibility spaces. Different accounts of these rules will posit different relationships between the corresponding spaces: on some views, the natural and the metaphysical possibilities are coextensive; on others, they are not; on some views, the logical possibilities outstrip the metaphysical possibilities; on other, they do not. And, of course, no one tells us the world's rules; there is no tablet on which they were written. So, we have to do our best to figure them out. If we've discovered some, then we can use them to assess modal claims; we can try to determine what the rules require and what they

¹⁵ What if your knowledge of the rules is partial? In this case, your judgments about unobserved states may have some epistemic credibility. However, for any given judgment, that credibility depends almost entirely on (a) whether you've discerned any of the rules that are relevant to the question at hand and (b) the proportion of the relevant rules that you've managed to uncover thus far.

¹⁶ ...although, when we compare our actual situation to the one that we have in *Clue*, things look grim: (1) we can't watch the world over again, so it's much harder to tell whether we are being deceived by coincidences, and (2) for all we know, there are no 'rules' that define the space of genuine, metaphysical possibilities – or at least there may not be any rules that are simple enough for creatures like us to comprehend them. But I take these as reasons for greater epistemic humility, not to deny that we have any modal knowledge.

allow. If we haven't discovered any, then it's hard to see why we should lend any epistemic weight to our modal judgments: we have no grip on the factors relevant to whether p is possible, necessary, or impossible. So, since we can't just divine the world's rules, it's plausible that our access to them is much like our access to the rules of *Clue*: we theorize about them based on whatever data we have available. It's therefore plausible that, with the exception of inferences from actuality to possibility, our modal beliefs are justified by being based on our best theories about the world's rules. Thus, POSS- is indeed too liberal, whereas it looks as though POSS strikes the right balance. The task now is to spell out TT – of which POSS is a component – in more detail.

2.3 TT in Detail

The Theory Theory (TT) is a modal epistemology that's based on the semantic view of theories.¹⁷ I think of it as the epistemology that Bas van Fraassen would defend if he were a scientific realist. Accordingly, let's begin with a suggestive passage from *The Scientific Image*:

Guided by the scientific theories we accept, we freely use modal locutions in our language. Some are easily explicated: if I say that it is impossible to observe a muon directly, or to melt gold at room temperature, this is because no counterpart to such events can be found in any model of the science I accept. But our language is much subtler and richer than that; its modal locutions reflect the fact that in the models of our theories we see structures that correspond to alternative courses of events, not all of which could be jointly actualized.

On the view of acceptance of theories which I have advocated under the name of constructive empiricism, it does not require belief that all significant aspects of the models have corresponding counterparts in reality. This applies to many aspects discussed by philosophers of science: space-time, elementary particles, fields, and, finally, alternative possible states and courses of events. The locus of possibility is the model, not a reality behind the phenomena (Van Fraassen 1980, 201-202).

Van Fraassen is drawing on both the semantic view and his constructive empiricism here. Let's

¹⁷ For arguments in favor of the semantic view, see (Lloyd 1994), (Suppe 1977, 1989), (Suppes 1993), (Thompson 1989), (Van Fraassen 1989).

set aside the latter for the moment. The characteristic feature of the semantic view is that it takes a theory to be a family of models. This claim is what distinguishes it from the syntactic view of theories.¹⁸ On the syntactic view, a theory is a set of statements – ideally, ones that are expressible in first-order logic. These statements are the theory’s laws, and they are interpreted by “bridge principles” or “correspondence rules” that link the theory’s theoretical terms (i.e., the predicates) to observable conditions. Unfortunately, this requires a distinction between the theoretical language and the observation language – a distinction that is notoriously difficult to draw.¹⁹ Moreover, the commitment to bridge principles or correspondence rules is based on the assumption that it’s possible to translate sentences from the theoretical language into the observation language. This too is no small hurdle. (Carnap’s *Aufbau* appears to be the best attempt, and the consensus is that it does not succeed.) To make the job easier, some allowed that there could be ‘partial’ interpretations of the theoretical language. But now there is reason to worry about the uninterpreted terms. Are they meaningless? If so, then scientists are simply making noise when they utter these terms, and this is a hard pill to swallow. However, if the uninterpreted terms *are* meaningful, whence comes their meaning?

Admittedly, it’s possible to formulate the syntactic view in a way that rejects the theoretical / observation term distinction, and so does not involve correspondence rules: you can interpret the theory’s statements so that their terms refer to – and their quantifiers range over – objects and properties ‘out there’ in the world. This eliminates the problem to which the correspondence rules lead. However, once you have the interpretation of the theory – the set of structures or models – it becomes unclear why it’s important that theory be axiomatized in one way rather than another. Wouldn’t you have the same theory regardless of how you defined the

¹⁸ I have learned a great deal about the syntactic view (and the semantic view) from (Suppe 1989). I am also indebted to (Thompson 2007).

¹⁹ On this point, see Putnam’s “What Theories Are Not” in (Putnam 1979).

set of models? Why is any particular linguistic formulation essential to it? Isn't the real question about the relationship between the models and the target system? These questions lead straight to the semantic view.

Drawing on (Giere 1979) and (Van Fraassen 1980), we can articulate the semantic view in the following way. A theory involves a theoretical definition and a theoretical hypothesis: the former defines a set of models; the latter articulates a relationship between the models and some phenomena. The theoretical definition stands to the models roughly as sentences stand to propositions. Many sentences can express a single proposition, and it's the proposition – i.e., what's said – that's of interest. Likewise, the proposed theoretical definition is the standard way to define the set of models, but there are other theoretical definitions that would do the same work, and it's the models that matter. Each model represents a state of a system; jointly, the models represent every 'possible' state of the system – i.e., every structure that satisfies the theoretical definition.²⁰ The theoretical hypothesis specifies the relationship between the models and the target system – e.g., that one or more of its members is similar to that system.²¹

In *Laws and Symmetry*, van Fraassen illustrates the semantic view by applying it to a theory of shadow (in its mass-noun sense – i.e., 'There is a lot of shadow in that picture' – rather than its count-noun sense – i.e., 'The shadows danced among the trees' (Van Fraassen 1989, 217-218)). We begin with a theoretical definition. Where x is any physical object:

1. If x casts any shadow, then some light is falling directly on x .

²⁰ It's important that p -models not *vacuously* satisfy the theoretical definition, since that will let in too many models: it will reduce to the suggestion that a theory specifies a model just in case that model is consistent with the theoretical hypothesis, which will include countless models that are completely irrelevant to the subject at hand. Unfortunately, I don't have a suitably strict account of satisfaction to offer. You might be able to get the right result by allowing the irrelevant models, but then saying that the theoretical hypothesis directs us not to take a stand on those models that vacuously satisfy the theoretical definition (i.e., we should neither affirm nor deny that they represent possible states of the target system).

²¹ Many other relationships might be posited: identity, isomorphism, partial isomorphism, etc. Van Fraassen breaks with Giere in preferring isomorphism over similarity. Isomorphism creates some complications that can safely be ignored in the present context; so, for simplicity's sake, I will preserve Giere's formulation.

2. x cannot cast shadow through an opaque object.
3. All shadow is shadow of something.

The above defines a set of models. Call the set '**S**'. The theoretical hypothesis ('**H**') is implicit, but in calling (1) – (3) 'a theory of shadow', the intent is probably that there should be no examples of shadow that are not represented by a member of **S**. So:

- (**H**) For any shadow-involving phenomenon s , there is a member of **S** to which s is similar.

The theory, then, is **S+H**. Given this, it's easy to see how the theory sanctions certain modal claims and rejects others. Let's say that a ' p -model' is a model that represents the target system as being such that p is true of it, and let's say that a 'representing model' is a model that is *supposed* to represent a possible state of the target system (i.e., a model that the theoretical hypothesis does not write off as, e.g., a spandrel of the idealization). Corresponding to each theory there are some 'theoretical necessities' (N_T), 'theoretical possibilities' (P_T), and 'theoretical counterfactuals' (C_T):

- (N_T) The theory says that p is necessary if and only if every representing model in **S** is a p -model.
- (P_T) The theory says that p is possible if and only if **S** includes a representing model that's a p -model.
- (C_T) The theory says that if p were the case, then q would be the case if and only if (a) **S** includes a model that is supposed to represent actuality (the '@-model'), (b) **S** includes representing p -models and representing q -models, and (c) the representing p -model most like the @-model is a representing q -model.²²

(In the shadow example, of course, every model in **S** is representing; but it won't be so in other contexts: everything depends on how and how severely the theory idealizes the target system.)

N_T , P_T , and C_T reveal that, on the semantic view, theories have modal content: they say how things must be, could be, and would be in various circumstances. So, for example, by the lights

²² This is based on Stalnaker's semantics for counterfactuals, but it's obvious that you could use Lewis'; the only difference is that Stalnaker makes the simplifying assumption that there will be a unique closest world.

of our theory of shadow:

- it is impossible for x to cast a shadow unless light is falling directly on x ,
- it is possible for x to cast a shadow through a non-opaque object, and
- if light were falling directly on x , then x would cast a shadow.

If there is any doubt about this, reconsider the theoretical definition and theoretical hypothesis. Jointly, they make it quite clear that the above are examples of a theoretical necessity, a theoretical possibility, and a theoretical counterfactual.²³

N.B., I have simply been assuming that the models represent possibilities thus far. Why should we think this? There was, at one time, a very elegant answer to this question. Early proponents of the semantic view simply *equated* a theory with a class of models. So, if you wanted to say that theory T represents reality, then you had to say that the class of models represents reality. But it's can't be the case that all the models represent actuality, since – on this version of the view – the models represent by being isomorphic to the target system, and (most of) the models aren't isomorphic to one another. Hence, most of the models must represent possibilities.

This line is not available once you introduce the theoretical hypothesis and abandon the tight connection between representation and isomorphism. (See (Frigg 2002) for arguments against the older view.) However, there is another route to the same conclusion. Recall a difference between the semantic view and its predecessor. On the syntactic view, a theory is a set of sentences; so, we can use our preferred theory of truth to cash out the specific case of theory truth. But on the semantic view, a theory is largely the set of models, and models aren't true or false of anything. Instead, what can be true or false is the theoretical hypothesis, which claims that (at least) one of the models is similar to the target system in certain respects and to

²³ Van Fraassen goes on to describe a geometrical theory that (of course) lends itself to being articulated with greater precision. But while it's important to note that the semantic view can handle mathematically sophisticated theories, it is just as adept at expressing mathematically unsophisticated ones (as the shadow example makes plain).

certain degrees. What of the other models? Plainly, they can't all be similar to the *actual* target system: if they were, then we would get the paradoxical result that the theoretical hypothesis is true just in case the world contains a system that is similar to many, many structures that are importantly dissimilar to one another. So, either we ignore the other models or we interpret them differently. Now, van Fraassen seems to think that scientific realists *will* interpret the other models as representing possibilities, taking them to be on an epistemic par with unobservables (see the passage quoted above). He does not unpack this thought, but he does hint at the reasoning behind it. One of the (alleged) virtues of constructive empiricism is that it “does not require belief that all *significant* aspects of the models have corresponding counterparts in reality” (Van Fraassen 1980, 202, my emphasis). In context, the implication is that non-actual states affairs are significant aspects of the models. Significant how? He doesn't say, but I think that we can fill in the story – or, at least, a story that the realist could tell. It's plausible that we take one of a theory's models to represent actuality because of the explanatory power of that model. But the other models are valuable for explanatory purposes too, since we often explain a system's behavior by noting how it would behave differently in different circumstances. Indeed, one powerful account says that *all* causal explanations detail patterns of counterfactual dependence between explanans and explanandum: (Woodward 2003). In short, we say that parts of the models represent unobservable objects and properties because of the explanatory power of this hypothesis. Likewise, we should regard many of the models as representing possibilities.

So, if you have reason to believe that a theory is true, then you have reason to believe that certain models correspond to genuine possibilities. There are, of course, those who will balk at this; indeed, van Fraassen's constructive empiricism is designed precisely to avoid such modal

commitments. On his view, the aim of science is to construct theories that are empirically adequate, not to discover theories that are true. Accepting a theory involves nothing more than believing that the goal of empirical adequacy has been achieved. Since accepting a theory does *not* involve believing that the theory is true, it also doesn't require believing that the possibilities it represents are genuine.²⁴ His suggestion is that constructive empiricists should use the semantic view to offer a reductive account of modal locutions (recall: "[t]he locus of possibility is the model, not a reality behind the phenomena"). By his lights, there are contexts in which it is perfectly appropriate to talk *as if* the models of our theories represent genuine possibilities – e.g., when we are presupposing that our conversation partners accept a well-confirmed theory and we are discussing the merits of a new hypothesis ("...this hypothesis seems to suggest that *p*, which is impossible..."). This amounts to (or is a gesture toward) an account of the assertability conditions for modal claims.

If we are realists, though, then we can use the semantic view as the basis of a modal epistemology. We will, of course, be obliged to defend our preferred account of theory selection – whether inference to the best explanation, or conditionalization, or what have you – but we had this obligation already. And once we discharge it, then the semantic view does the rest: our account of theory selection explains how we are justified in believing that certain theories are true; the semantic view explains how our theorizing generates modal commitments.

Now on the semantic view, when you believe that a theory is true, you believe that a class of models represents a system.²⁵ We can say, therefore, a theory commits us to a modal claim if

²⁴ Stathis Psillos argues that constructive empiricism is an unstable halfway house between (some form of) scientific realism and full-on skepticism about ampliative inference. See (Psillos 1999), particularly Chapter 9.

²⁵ I should emphasize that we need not say that every model represents a genuine possibility: for example, there is probably a way to formulate a theory of inclined planes so that you needn't believe that there could be frictionless planes; the trick would be to formulate the theoretical hypothesis so that it takes no stance on whether the limit cases represent the target system. Nevertheless, there are going to be many possibilities that we surely *will* regard as genuine if we believe the theory. It's hard to imagine why we would claim to believe Galileo's theory about the

it ‘says’ that p is necessary, that p is possible, or that if p were the case, then q would be the case. If a theory commits you to a modal claim, then your reasons to believe the theory are reasons to believe the modal claim. So, for example, if you justifiably believe a theory that says that p is possible (in the sense of P_T , above), then – absent any defeaters and assuming that you believe that p is possible on this basis – you are justified in believing that p is possible. And based on the lessons gleaned from our reflections on *Clue*, we’ll suppose that we are not justified in believing more about modal reality than our theories suggest: i.e., if you are justified in believing that p is possible (for example), then you justifiably believe a theory that says that p is possible. We now have the makings of the following biconditional:

[POSS] You justifiably believe that p is possible iff (a) you justifiably believe a theory that says that p is possible, (b) you believe that p is possible on the basis of this theory, and (c) you have no defeaters for the belief that p is possible.

POSS, of course, is not a complete modal epistemology; we need corresponding accounts for necessities and counterfactuals. Here they are:

[NEC] You justifiably believe that p is necessary iff (a) you justifiably believe a theory that says that p is necessary, (b) you believe that p is necessary on the basis of this theory, and (c) you have no defeaters for the belief that p is necessary.

[CF] You justifiably believe that if p were the case, then q would be the case iff (a) you justifiably believe a theory that says that if p were the case, then q would be the case, (b) you believe that if p were the case, then q would be the case on the basis of this theory, and (c) you have no defeaters for the belief that if p were the case, then q would be the case.

mechanics of inclined planes while denying that objects can change their locations. Would it make any sense to affirm this theory while denying that a body placed at the top of a slope could come to occupy a position at the bottom? If we believe that the theory is true, then surely we are committed to this claim and many more like it. So, with the appropriate caveat about the theoretical hypothesis, we’ll say that when you believe that a theory is true, you believe that *some members* of a class of models represents a system. Also, on a different note, I should point out that it is contentious to say that believing a theory commits you saying that its models represent a target system. Many theories idealize their target systems, and some people think that idealized models do not represent their target systems: e.g., (Klein 2008). For my purposes, though, this disagreement boils down to a debate about how complicated it is to determine the genuine possibilities based on (what I’ll call) the theoretical possibilities. At first blush, I see no reason why I couldn’t complicate my story to accommodate this view of idealization. And, in any case, I take it that very few scientific realists (my main audience) think that *all* theories idealize their target systems in this way, and it is always open to them to restrict the scope of my comments to those theories that they take to represent their target systems more straightforwardly.

TT is the conjunction of POSS, NEC, and CF. It's plain, though, that these three propositions share a single form. So, we can state TT more succinctly if we gloss over the differences between necessity claims, possibility claims, and counterfactuals. Where p is any modal claim whatever:

[TT] You justifiably believe that p iff (a) you justifiably believe a theory that says that p , (b) you believe that p on the basis of this theory, and (c) you have no defeaters for the belief that p .

2.4 Clarifying TT

I think that it is fairly clear how to understand TT's Condition A: the only technical terminology is explained by N_T , P_T , and C_T – i.e., what it is for a theory to 'say that' p . Condition C is also straightforward. On the standard view, there are two types of defeaters: rebutting and undercutting. A rebutting defeater for your belief that p is a reason to believe that $\sim p$; an undercutting defeater for your belief that p is a reason to think that your belief was formed improperly – e.g., on the basis of false testimony, or sloppy reasoning, or what have you. Condition C just says that you have no such defeaters. The only question, then, is how to understand Condition B's language. What is to believe that p has a particular modal status 'on the basis of' a theory?

If we take this as a properly epistemic question, then it is a version of a much more general query: what it is for belief b_1 to be based on belief b_2 in such a way that b_1 acquires the epistemic merits of b_2 , whatever they may be? This is, of course, a weighty epistemic matter. But so interpreted, we need not reply. TT is probably compatible with a number of positions on the nature of the basing relation, and I see no reason to have it stand or fall with any one of them. So, we can avoid wading into those deep epistemological waters.

An alternate reading, though, takes the question to be about the mental links between our modal beliefs and our beliefs about theories. What does TT say about how justified modal beliefs are formed? In other words, what does it say about the natural history of justified modal belief? This is a descriptive question: the issue is *how in fact* our modal beliefs are based on our theories, and not how this relation *justifies* those beliefs. This issue comes up because – unlike many of the prominent modal epistemologies on offer – TT does not associate justified modal belief with any particular mental activity: e.g., conceiving, or imagining, or intuiting. (If you think, for example, that conceivability is the only guide to non-actual possibility, then your story is readymade: insofar as we understand what it is to conceive that *p*, we understand quite a bit about what’s going on when you come to believe that *p* is possible.) Hence, it is legitimate to wonder whether there is a descriptive account of justified modal belief that fits with my prescriptive account of justified modal belief. I am afraid that I cannot offer an answer that is as detailed as I would like it to be. Here, however, is the beginning of a response.

There is an approach to cognition known as ‘the theory of mental models’.²⁶ On this view, we reason by way of what are, essentially, iconic versions of the classes of models that the semantic view employs. Here is how P. N. Johnson-Laird summarizes one of the theory’s key tenets:

A principle of the modern theory of mental models is that a model has the same structure as the situation it represents. Like an architect’s model, or a molecular biologist’s model, the parts of the model and their structural relations correspond to those of what it represents. Like these physical models, a mental model is also partial because it represents only certain aspects of the situation. There is accordingly a many-to-one mapping from possibilities in the world to their mental model (Johnson-Laird 2004, 181).

In the above passage, Johnson-Laird may mean only that there is a one-many relationship between the mental model and the target systems: i.e., the mechanic does not have one mental

²⁶ See (Gentner and Stevens 1983) and (Johnson-Laird 2004) for helpful overviews.

model for each engine he repairs, but one mental model that applies to a number of engines. However, it's plausible that he intends more: namely, that on the side of the many, some of the relata are non-actual states of affairs – ways that the target systems could be. This suggests a natural relationship between theories and mental models: mental models probably aren't *identical* to theories, at least as the semantic view interprets them, but they may still be best *represented by* theories. We can suppose that a mental model specifies (more or less exactly) a set of models by way of the structural similarities that Johnson-Laird indicates. Some of those models represent actual states of affairs, but many will represent non-actual states of affairs. And we can interpret our attitudes toward our mental models as roughly equivalent to theoretical hypotheses: if we endorse every implication of a given mental model, then we suppose that there is an exact correspondence between our mental model and the target system; if we withhold judgment in some cases, then we suppose that our mental model is imprecise in some respect or other. The upshot is that the theory of mental models appears to sit comfortably with the semantic view of theories; it seems to provide an attractive account of what's 'in the head' when you believe a theory – namely, a mental structure that corresponds to a class of models.

Here is an example. Suppose that we're trying to determine how many bottles of beer the fridge will hold. According to the theory of mental models, what we have in mind is a structure that in some way resembles the fridge and bottles. Some features of the structure are invariant, but others are, essentially, 'moving parts' – i.e., features of the structure that can be altered, or that are removable, allowing new features to be introduced. Presumably, the walls of the fridge correspond to the invariant aspects of the mental model; the moving parts – i.e., the shelves, the drawers, the bottles – correspond to the variable components. This structure represents – more or less precisely and more or less accurately – all the ways that beer can be arranged in the

fridge. If the mental model theory is correct, then we reach a conclusion about how best to pack the fridge by manipulating this mental structure.

Let's suppose that the mental model theory is indeed true. Then, we can say that you believe p 'on the basis of a theory' if you believe p as a result of manipulating the mental structure that corresponds to appropriate class of models. I suggest that we take this as an account of the cognitive machinery behind conceiving, imagining, and intuiting – the standard routes to justified belief about modal matters. Very roughly, then, to conceive or imagine that p is to manipulate a mental structure so that it corresponds to a p -model. Depending on a variety of factors, this may or may not involve a number of mental images; hence the distinction between conceiving and imagining. Intuiting that p , on the other hand, might be either (a) a species of conceiving or imagining – perhaps just the limit case where speed is concerned – or (b) to come to believe p based on some heuristic device, itself just a rough and ready approximation of a more complex structure.

I am not the first to maintain that conceiving, imagining, and intuiting are epistemically dependent. Many naturalists, for example, have argued that intuitions confer justification only because they are appropriately related to the mechanisms that underwrite our semantic or logical knowledge (DePaul and Ramsey 1998). My own stance, however, is more like Michael Devitt's. Here is his take on intuition:

On my view, the intuitions that concern us [i.e., philosophical and linguistic intuitions] are judgments that are empirical theory-laden central-processor responses to phenomena, differing from many other such judgments only in being immediate and unreflective, not based on any conscious reasoning. These intuitions are surely partly innate in origin but are usually and largely the result of past reflection on a lifetime of worldly experience (Devitt 2010, 294).²⁷

²⁷ Devitt quotes Elkhonon Goldberg approvingly, who writes: "...intuition is the condensation of vast prior analytic experience; it is analysis compressed and crystallized [...] It is the product of analytic processes being condensed to such a degree that its internal structure may elude even the person benefitting from it" (Goldberg 2005, 150).

For Devitt, intuitions offer derivative justification. Intuiting doesn't generate the justification whole cloth; the justification that it provides is not the product of some special faculty that is independent of our other sources of knowledge. Instead, it is an especially quick sort of judgment based on "a lifetime of worldly experience" – in my view, experience consolidated in theories. And we can take essentially the same line on conceiving and imagining: they are quite like intuiting in providing derivative justification, although they do so more slowly, more deliberately, and involving more or less sensuous imagery as the case may be. So, I am happy to say that conceiving, imagining, and intuiting play important roles in the epistemology of modality, although I deny that they play the most important role: if TT is correct, then they are downstream from the basic source of justification – namely, the arguments that we give for our best theories.²⁸ In short, we apply mental models to particular cases by conceiving, imagining, and intuiting; when all goes well, the results are justified modal beliefs.

Of course, much more needs to be said about the theory of mental models, its purported fit with the semantic view, and the picture of conceiving *et al.* that I've just sketched. Still, this seems like a promising way to forge connections between TT, cognitive psychology, and the notions that have traditionally held pride of place in the epistemology of modality.

²⁸ Not incidentally, David Lewis holds a view like this – at least with respect to the imagination. "We get enough of a link between imagination and possibility, but not too much, if we regard imaginative experiments as a way of reasoning informally from the principle of recombination. To imagine a unicorn and infer its possibility is to reason that a unicorn is possible because a horse and a horn, which are possible because actual, might be juxtaposed in the imagined way" (Lewis 1986b, 90). (I reject Lewis' principle of recombination, but I agree with the epistemic priority that he gives it over the imagination.)

2.5 Some Objections

In what remains, I consider five preliminary objections to TT.

2.5.1 The Cheating Objection

TT claims that our modal beliefs depend for their justification on our best theories. So, TT needs to defend an account of how our theories are justified. I intend to do this in the next chapter. However, even before we consider any of the details, you might worry whether there is any account that could work for my purposes. The problem is this. While the process of justifying a theory may involve showing that it has certain intrinsic virtues, it also involves showing that the theory is superior to its alternatives. However, you can't demonstrate that a theory is superior to its alternatives until you know what its alternatives are. How do we identify these hypotheses? One important criterion is *possible truth*. If a hypothesis is not even possibly true, then it is immediately out of the running: all hypotheses that *are* possibly true are therefore superior to it. Recall, though, that theories countenance possibilities. So, if the theory is possible, then what it says to be possible is possible. (This follows from the attractive thesis that, if it is possible that it is possible that p , then it is possible that p .²⁹) The upshot is that, if we start out with justified beliefs about which hypotheses are possibly true, then we already have the justification that TT was supposed to provide. In other words, TT appears to cheat: it claims that our modal beliefs gain their justification from our theories, but it turns out that TT doesn't work unless we have that justification at the outset.

This objection asks too much of candidate hypotheses. We can distinguish between (a) knowing that a theory is possibly true, (b) knowing that a theory is *not* possibly true, and (c) not

²⁹ This inference is valid as long as the accessibility relation is transitive, which means that it holds in S4; you don't need the stronger S5 to make it go through.

knowing whether a theory is a possibly true. Clearly, you should not take seriously those hypotheses that you know to be impossible. However, if you don't know whether a theory is possibly true, then you don't yet know that it is off the table. It is still a live hypothesis. And, assuming that it fares well enough on the other criteria for theory candidacy, why shouldn't it be evaluated alongside the competition? If no reason is forthcoming, then there is room for the order of justification that TT recommends.

Still, this objection raises an important issue. TT implies that, if you justifiably believe that theory T is possibly true, then you needn't justifiably believe that it is actually true to glean modal information from it: if you justifiably believe that T is possible, and if you justifiably believe that T says that p is possible, then you (can) justifiably believe that p is possible. Now let's set aside theories that you justifiably believe to be possible because they have been 'vouched for' – i.e., those cases where you justifiably believe that T is possible because you justifiably believe another theory, T^* , that says that T is possible. Our interest here is in cases where (a) you don't justifiably believe any theory that vouches for T and either (b₁) you do not justifiably believe T to be true or (b₂) you justifiably believe T to be false. In such cases, TT has to deny that you justifiably believe that T is possible. Let's say that TT is 'guarded' to indicate that it has this feature.

Given the assumption that we don't justifiably believe any theory that vouches for the following theoretical claims, here are some consequences of the fact that TT is guarded. If TT is true, then we do not have justified beliefs about the following claims (i.e., we do not justifiably believe that they are true or that they are false):

- there are Newtonian worlds;
- there could have been phlogiston;
- atoms could have been indivisible;
- it could have been the case that all is water;

- it could have been the case that the function of the brain is to cool the blood.

Some people will regard these consequences to be unpalatable. I could try to make them seem more palatable by elaborating on my earlier comments about our limited access to modal reality.

I doubt that this will sway any true believers. So let me try a different tack.

To begin, we should distinguish between ‘conditional’ and ‘unconditional’ modal knowledge. So, for example, you have conditional modal knowledge that p is possible just in case you know either that:

- if q is possible, then p is possible, or
- if q is necessary, then p is possible.

By contrast, you have *unconditional* modal knowledge that p is possible just in case you know that:

- p is possible.

TT – like virtually all modal epistemologies – is a theory of unconditional modal knowledge, and it takes a relatively dim view of the scope of that knowledge. However, conditional modal knowledge is much easier to come by, and I suspect that it’s sufficient for many purposes. For each component of TT – POSS, NEC, and CF – there will be a corresponding conditional version, but I’ll just indicate here how we might formulate the conditional version of POSS:

[C-POSS] You justifiably believe that if T were true, then p would be possible iff (a) you justifiably believe that T says that p is possible and (b) you are not aware of any other theory T^* such that (b₁) if T were true, then T^* would be true and (b₂) T^* says that necessarily, $\sim p$.

Much of our modal reasoning is hypothetical: it doesn’t matter whether it’s *really* possible that p ; it just matters whether it’s possible that p given certain assumptions, or on such-and-such a view, or what have you. I submit that C-POSS – and the corresponding versions of NEC and CF – can helpfully be deployed in these cases; we can now say when that reasoning results in a

justified belief. And in particular, they should allow us to handle cases where we are interested in the implications of discredited (or not yet credited) theories.

So TT is guarded: it requires denying some possibility claims that we may be tempted to affirm. However, apart from the intuitive pull of these claims, are there any circumstances in which we need to believe that such claims are unconditionally true (or false)? Are there any circumstances where we can't get by with the conditional analogues of these claims? I think that there are two troubling cases. One involves an interlocutor who proposes a thesis to which we would like to object, but our objection involves a contentious modal claim (e.g., someone proposes the mind-body identity theory, and we object with the possibility of disembodied existence). We might think that we are in a better position if we justifiably believe the modal claim on which our objection relies. But on closer inspection, it turns out that this isn't so. On the one hand, if the critique is supposed to be internal – i.e., if our interlocutor is expected to believe the modal claim – then we need not believe the modal claim at all. We can simply point that her thesis implies (say) that p is impossible, though she believes p to be possible; that's enough to cause her some trouble. On the other hand, if the critique is supposed to be external – i.e., if we do not expect that our interlocutor will believe the modal claim – then our objection doesn't carry much weight to begin with, and it doesn't seem to be relevant whether or not our belief is justified. (Or, if it is relevant, it's only because we've created a worry about peer disagreement. However, disagreement cuts both ways; so, if it undermines our interlocutor's justification, then it undermines ours as well.) Either way, it doesn't seem to matter whether we have conditional or unconditional modal knowledge.

The second worrisome case involves trying to develop a hypothesis about the general shape of modal reality. Given the scope of the hypothesis, we want as many data points as we

can get, and a belief only qualifies as a data point if it's justified. (It's of no use to frame a hypothesis based on our guesses about what is and isn't possible.) But while this is a worthy aim, it's far from clear that we should expect to be able to develop a well-confirmed theory of the shape of modal reality. If we are realists about modality, we should expect this to be a difficult task, and this for all the reasons mentioned at the beginning of the chapter. Moreover, there is no obvious reason why a proponent of TT must agree with her critic about the space of worlds. If the proponent and critic disagree about what's possible, then they may also disagree about the general principles that organize the modal facts. But in the absence of an argument to the effect that our beliefs about those principles are more secure than the arguments for a particular modal epistemology – and an argument to the effect that those principles are the ones that the critic defends – it's hard to see where the difficulty lies.

So TT is guarded: it requires denying some possibility claims that we may be tempted to affirm. But in the absence of reasons *why* we should affirm them, I think that TT can hold its ground.

2.5.2 The No-Theory Objection

I see no problem with the fact that TT is guarded. But you might worry that TT isn't merely guarded, but also amounts to a fairly wide-ranging skepticism. TT says that, where p is any modal claim, you justifiably believe that p only if you justifiably believe a theory that says that p . This condition seems to imply that many of our modal beliefs are unjustified, since it's implausible that there are justified theories standing behind those beliefs. What theory says that I could have had coffee with breakfast this morning, or that my desk could be a foot to the left, or that my shirt could be striped instead of plaid? This problem is not limited to our beliefs about

mundane modal matters; it also crops up for many of our metaphysical beliefs. What theory says that I could be disembodied, or that I could not have been born to different parents, or that it is impossible for an object to have a gappy existence? If none can be found, then TT says that the corresponding beliefs are unjustified.

Let's begin our beliefs about mundane modal matters: I could have had coffee with breakfast this morning; my desk could be a foot to the left; my shirt could be striped instead of plaid. What theories say that these things are possible? We might take a first step toward a solution by taking a more generous view of what counts as a theory. Indeed, we might say that *all* our beliefs somehow 'amount to' a theory. (Quine takes a position like this.³⁰) Of course, it matters what 'amount to' amounts to in this context. Is it the case that our beliefs *are* theories (i.e., each belief is a mini-theory)? Alternately, do our beliefs jointly *constitute* a theory – or a family of theories? Either way, perhaps we believe the theory in virtue of having the beliefs. But maybe 'amount to' should be construed more broadly. Are our beliefs merely *an expression* of a theory – and hence are distinct from it? Or are our beliefs simply *evidence* of a theory – perhaps the one that we would accept if we were presented with an explicit version of it? (These interpretations seem to fit the way that 'descriptive metaphysicians' think about our beliefs, since they take our theory of the world to be something about which we need to be informed; see, e.g., (Strawson 1959)). On the broader reading, it's less clear whether we believe the theory in question – much less whether we're *justified* in believing it. Hence, it's not obvious that these interpretations are of use to TT, at least in its present form. However, if pressed to opt for one of these interpretations, then I see no reason why TT couldn't be modified accordingly – maintaining, say, that you justifiably believe *p* only if you stand in the appropriate relationship to a theory that says that *p*, where that relationship will depend on the interpretation that wins the

³⁰ For similar perspectives, see (Lewis 1970), (Churchland 1986), (Churchland 1989), and (Baker 2007).

day. However these details are sorted out, let's say that 'folk theory' is the theory that is alleged to stand in the relevant relation to all those beliefs that we do not obviously hold on the basis of some scientific theory,³¹ and let's say that 'the folk theory hypothesis' is the suggestion that those beliefs *do* stand in the relevant relation to folk theory.

To be clear, some think of the folk theory hypothesis as an *empirical* hypothesis about parallels between the acquisition, structure, or function of our ordinary beliefs and the acquisition, structure, or function of our scientific theories; see, e.g., (Gopnik and Meltzoff 1997). I don't. First, I mean it as an epistemological claim: however we acquire our ordinary beliefs, they are *justified* (or at least justifiable) in the same way as our scientific beliefs. On this view, there is no difference in kind between the arguments that we offer for scientific theories and the arguments that we (could) offer for folk theory. (To my mind, the exemplar here is the way that Jonathan Vogel argues that the Real World Hypothesis is a better explanation of our experience than is any of the scenarios proposed by the Cartesian skeptic; see (Vogel 1990).) Second, I mean it as a claim about how we can think about the content of folk theory: folk theory can be represented just as we represent scientific theories, which is to say that we can treat folk theory as a set of models. The epistemological aspect allows me to say that folk theory can also be justified however our scientific theories are justified; the bit about representing folk theory lets TT handle even our ordinary modal knowledge.

The next step is to argue that folk theory has something to say about the examples with which we began. After all, surely there are matters about which folk theory is silent, or ambiguous, or even flatly inconsistent. If the theory is silent with respect to (say) *p*'s possibility,

³¹ Or, if you prefer, let's say that 'folk theory' is that part of our global theory that stands in the relevant relation to all those beliefs that we do not obviously hold on the basis of a scientific theory. Alternately, let's say that our 'folk theories' are those theories that stand in the relevant relation to various sets of beliefs not obviously held on the basis of a scientific theory. I'll ignore these variants in what follows, since I don't think that much turns on how we individuate theories.

then TT says that we ought to abstain from belief about it; if the theory is vague, then we will need to make it more precise before we take a stand; if the theory is inconsistent, then we will need to repair it before we believe one way or the other. That said, I do think that folk theory has something to say about the examples in question:

- *I could have had coffee with breakfast this morning.* We appeal to beliefs and desires to explain our actions, and behind these explanations is the assumption that different belief/desire combinations would have led to different actions; the belief/desire pairs explain because of the way that actions (supposedly) depend counterfactually on them. It's plausible, therefore, that folk theory includes some folk psychology, part of which is a primitive libertarian theory of the will. So, my guess is that folk theory has models in which (a) I wanted coffee with breakfast this morning and (b) I successfully acted on that desire. Hence, folk theory has models in which I have coffee with breakfast this morning.
- *My desk could be a foot to the left.* We spend our lives navigating a world full of medium-sized objects. Accordingly, we have good reason to develop a working theory of what they can do and what can be done to them. This theory will be no good if it says that, at a time, no object can occupy a location other than the one it in fact occupies. Either the theory says that I have the ability to move my desk, or it doesn't. If I have this ability, surely I can exercise it at any given time (all other things being equal). And if I don't have this ability, how do we explain the cases in which I successfully slide my desk to a different spot? So, I think it quite likely that folk theory has models in which my desk is a foot to the left.
- *My shirt could be striped instead of plaid.* This sentence is ambiguous. On perhaps the

simplest reading, the claim is that I could be wearing a different shirt – namely, one with stripes. This makes it much like the claim that I could have had coffee with breakfast this morning; presumably, our theory of the will allows that I have these sorts of options available to me. Read another way, the sentence says that the plaid pattern could be replaced with stripes. This interpretation makes it akin to the assertion that my desk could be a foot to the left: we know that enough bleach will remove the existing pattern; fabric markers will do the rest. But there is yet a third reading: understood this way, the allegation is that my shirt could *always* have been striped rather than plaid. This is perhaps the most interesting case, since it's not obvious what folk theory says – if anything at all. This is no objection, though, because I think that silence is the right result. Could my shirt have had a different pattern? Could it have been made from different cloth? Could some of its threads have been of a different color? Could it have been made elsewhere by different people? Could it have been made by a different company? I don't know how to answer these questions. Or, more accurately, I don't know how to answer these questions without invoking a contentious metaphysical thesis (e.g., the necessity of origin, or mereological essentialism, or what have you). But the present view lends itself to an explanation of why this is so: we are rarely confronted with situations in which we need to settle questions like these, and so folk theory is not well-equipped to handle them. Of course, we could *extend* folk theory so that it *does* answer them – and alongside precisifying and repairing folk theory, that is just the sort of work to which metaphysicians devote themselves. At any rate, there are some readings of *My shirt could be striped instead of plaid* that folk theory seems to affirm fairly unambiguously, and there is at least one where matters are much less clear. And that, I

think, is just what we should expect.

To be perfectly clear, we are free to reject any of the theories that I just used to get these results; no doubt many will want to reject the libertarian component of folk psychology (assuming that I am right in thinking that there is such a component). And this is as it should be. At the same time, if we want to avoid the No-Theory Objection, then we can do our best to defend these theories. We might say, for example, that since we have no better way to understand and predict macroscopic events than by appealing to ships and shoes and sealing wax, we have reason to believe that this theory is true. Likewise for our theory of the will. And, if we are justified in believing it, then TT allows us to use these theories to justify our mundane modal beliefs. So, TT need not deny that our mundane modal beliefs are justified.

The above solution to the No-Theory Objection is designed to preserve our mundane modal beliefs. But the No-Theory Objection is not limited to our mundane modal beliefs: it also asks whether we justifiably believe theories that could underwrite our metaphysical beliefs. Here I am less willing to accommodate the terms of the objection. The No-Theory Objection has force because we demand that an epistemology agree with our considered judgments about those beliefs that are and are not justified. But an epistemology can't be expected to agree with our judgments about *every* case, since there are cases where we disagree, and there is no reason to suppose that we are all correct even where we believe differently. So, insofar as this demand is a legitimate one, it has to be qualified: an epistemology ought to deliver verdict v where there is a general consensus on the rightness of v , and it ought *not* to deliver v where there is a general disagreement about v . Since it is hard to think of a metaphysical topic on which there is a general consensus, an epistemology is under very little pressure to accommodate one set of metaphysical claims over another.

Of course, if you have some metaphysical commitments, then you'll want your epistemology to say that the relevant beliefs are justified. This TT can do, assuming that you take yourself to be justified in believing a relevant metaphysical theory. If you are, then TT says that you can be justified in believing the modal claims that it sanctions. So, *if* you think that you have a good argument for physicalism, then you can use TT to say that you are justified in denying that there could be zombies, or that we might be disembodied. And *if* you think that you have a good argument for dualism, then you can use TT to say that you are justified in taking the opposing stances. Crucially, though, TT puts the weight on the arguments for the metaphysical theories, which – if TT is correct – are rarely going to be modal arguments. If the justification of your modal beliefs depends on the justification of (some of) your theories, then you will rarely be in a position to defend a theory based on modal claims; that would be to put the cart before the horse (unless, of course, those claims are sanctioned by some other theory that you justifiably believe). Usually, the arguments for theories are going to instances of IBE: some features of actuality are taken as the data, and the claim is that they are best explained by such-and-such hypotheses. The modal claims then fall out of the relevant theories –not vice versa.³² And given all this, the No-Theory Objection is no serious threat to TT.

³² Incidentally, all this points to a virtue of TT. An epistemology is better if it can *explain* both the consensus on some cases and the lack thereof on others. With the help of the folk theory hypothesis, TT can do this. We accept folk theory – else it wouldn't have that honorific – and hence we accept the mundane modal claims that it sanctions. But most people are not forced to engage seriously with metaphysical questions, and hence folk theory does not answer them – or doesn't answer them clearly and consistently. So, when pressed, most of us offer our gut reactions, our off-the-cuff judgments. And since these are untutored replies – shaped as much by temperament as anything else – there is no reason to expect uniformity across individuals. Hence, the disagreement.

There are, of course, those of us who worry at length about metaphysical questions, and I think that it is wrong to write off our replies as gut reactions. Still, we cannot agree about how to interpret the data. But I do not want to suggest that the mere presence of this disagreement is enough to undermine any individual's justification about metaphysical matters: I am perfectly happy to concede that there are people with justified metaphysical beliefs. What I want to suggest is that this disagreement is plausibly explained by either the paucity of the data, the difficulty of interpreting the data there is, or some combination of the two. We can put the point by saying that, while we always face the problem of underdetermination of theory by evidence, the problem is particularly acute where metaphysical theories are concerned. Hence, even if we are justified in believing certain metaphysical theories, we probably aren't in a position to convince others that our views are true: there are too many other

2.5.3 The Logic Objection

If my reply to the No-Theory Objection works, then it's because it's legitimate to expand our conception of what qualifies as a theory – that's the upshot of the folk theory hypothesis. Now a different problem crops up. Either the logic I believe counts as a theory, or it doesn't. If it does, then TT is at risk of implying that I'm justified in believing that whatever is logically possible is possible *simpliciter*. In other words, it seems that I am justified in taking consistency to be a guide to possibility. But that does not seem to sit well with the considerations that led us to this modal epistemology: TT is supposed to be cautious, but this would make it completely reckless. On the other hand, if logic doesn't count as a theory, then what is it? And more importantly, how are our logical beliefs justified? After all, for any theory, we need know whether it is consistent in order to know whether it has any models at all, and hence we need some justified logical beliefs for TT to get off the ground.

Russell and Frege had a 'universalist' conception of logic: they regarded it as the most general theory, the theory that applies to absolutely everything. I would prefer not to give up this view of logic; I'll grant that the logic we believe qualifies both as a theory and as something that we justifiably believe. So, I need to find a way to block the suggestion that, if TT is true, then consistency is a good guide to possibility.

Logic may well be the most general theory, but this alone is not enough to show that TT leads to the view that consistency is a guide to possibility. To get there, you also need the assumption that, for any given possibility claim sanctioned by our logic, we have no defeaters for it. Is that the case?

plausible hypotheses that are compatible with the same data. That there would be disagreement, then, is all but a foregone conclusion.

I don't think so, but I don't think that I can defend this view in full detail. Let me simply outline the argument. It may well be that consistency is coextensive with (absolute, unrestricted) possibility, but even if so, consistency is not of much epistemic use. Suppose – plausibly enough – that we can establish identity claims based on *a posteriori* considerations and that we somehow know that identities are necessary. (The latter claim is easy to establish given TT: we are justified in believing standard first-order logic with identity, and that theory has no model in which $a \neq a$; therefore, we are justified in believing that identities are necessary.) We now have the old epistemological worries created by *posteriori* necessities. To be clear, *a posteriori* necessities do not create trouble because they lack the requisite form³³; rather, they are awkward because the form of the proposition is not obvious from the form of the sentence that expresses it. The sentence 'Water is not H₂O' may indeed express a proposition that has the form ' $a \neq a$ ',³⁴ but if you don't already know this, then the sentence won't tip you off. (As Nathan Salmon would put it, you believe the proposition under a guise that does not presume the identity of water and H₂O; hence you have no cognitive access to (the relevant aspect of) the form of the proposition. See (Salmon 2007) for details.) The upshot is that if you are judging consistency based on the form of the sentence, then your guide to possibility is highly fallible: it's going to deliver false positives whenever you are unaware of the relevant empirical facts. And given (a) the sheer complexity of the natural world and (b) the division of cognitive labor that characterizes modern society, there should be little doubt that you are unaware of many such facts. (Parallel problems crop up with property identities if we allow second-order quantification.) The obvious way to guard against this difficulty is to limit your use of this guide

³³ ...or fail to entail a proposition with the requisite form. I'll suppress this alternate formulation in what follows.

³⁴ This is surely too simple, since it is far from obvious that 'water' or 'H₂O' are singular terms, each of which refers to the same massive scattered object. However, I don't think that the details matter for my purposes here.

to those cases where you have positive empirical reasons to believe that $a \neq b$, not merely the absence of evidence that they *are* identical. What would constitute having such reasons? First and foremost, having theories about the natures of a and b . But this, of course, is a much-chastened version of the view that consistency is a guide to possibility, and hardly constitutes a threat to the caution that inspires TT.³⁵

What I've just given is a strategy for managing the Logic Objection, but it's a strategy that rests on the case for there being *a posteriori* necessities. As a defensive move, the strategy is quite effective: given TT, it's plausible that we're justified in believing that identities are necessary; given the standard battery of Kripkean arguments, it's plausible that the meanings of some natural kind terms are rigid designators; and given a natural interpretation of what scientists tell us, it's plausible that they have established some identities. However, should it turn out that such necessities are indefensible (and that no surrogates can be found), TT is stuck with the view that consistency is a good guide to possibility. I don't care for this conclusion, but it is no flaw in TT that it permits it. Indeed, it is probably a *virtue* of TT that it permits the view that consistency is a guide to possibility: it would be a mistake to set up a modal epistemology so that it ruled out this position from the start. We shouldn't gerrymander our modal epistemology so that it gives the results we want, bolstering the philosophical positions to which we are most sympathetic. Rather, we should let our epistemology be flexible, able to

³⁵ We might also bring up property incompatibilities, since it doesn't look as though they even exhibit the requisite form. Many have thought that the proposition expressed by the sentence 'Nothing is simultaneously red all over and green all over' is necessarily true. It seems to me, though, that the denial of this proposition is neither contradictory nor such that it entails a contradiction. Similar points could be about, e.g., being God and being omniscient, or being a panther and being reptilian, or being *this* clay pot (right here) and being made of *that* material (over there). If these propositions are indeed necessary, their necessity does not seem to be explicable in terms of their logical form. However we would like to classify this breed of necessity – perhaps 'absolute' or 'metaphysical' necessity are the best options available – it looks as though consistency outstrips it, since there are consistent but metaphysically impossible propositions (e.g., *something is simultaneously red all over and green all over*). However, this property incompatibilities beg the question against the objection. It seems to me that insofar as I have a theory of color, that theory has no model in which a wholly red object is simultaneously wholly green. However, the point of the Logic Objection is that I have some *other* theory – namely, first-order logic with identity – on which that an object *can* be wholly red and wholly green at the same time.

accommodate an array of tighter and looser restrictions on the scope of our modal knowledge. The challenge then is to argue for particular restrictions, and an appeal to our modal epistemology at that juncture would only beg the question at hand. At any rate, I will here just assume that there are some *a posteriori* necessities, and that these constitute the needed restrictions to block the Logic Objection.

2.5.4 The Modal Collapse Objection

The Logic Objection amounts to the charge that TT lets in too many possibilities; the Modal Collapse Objection amounts to the worry that TT lets in too few. It goes as follows. Either we have theories about what's actual, or we don't. Suppose we do. Recall that, according to TT, a theory says that p is necessary just in case every model of the theory is a p -model – i.e., a model that satisfies p . Now let's imagine that I've forgotten where I left my keys. I consider the available information, and I form a theory about where they are: my theory is that they are on the kitchen counter. But *every* model of this theory says that my keys are on the counter, and hence this theory says that it's *necessary* that my keys are on the counter. Alternately, suppose we *don't* have theories about what's actual. But what theories are we justified in believing if *not* theories about what's actual? The worry now is that we never satisfy Condition A – and hence have no justified beliefs about modal matters whatever.

Let's work with the first horn of the dilemma. Suppose that we have a theory about the way things are in some respect: about what my grandfather did for a living, or how many honey-producing beehives there are in the U.S., or the rate at which signals are transmitted over coaxial cable. Now, by the lights of any theory, its positive claims are necessary (i.e., the claims in the theoretical definition), since they are true in each of its models (else they wouldn't be couldn't

among the theory's models). So let's suppose that a theory claims that there are two and a half million honey-producing beehives there are in the U.S. According to TT, if I am justified in believing this theory, then I am justified in believing the modal claims it sanctions. So, I am justified in believing that it is *necessary* that there are roughly two and a half million honey-producing beehives in the U.S. And surely it is implausible that this is so.

There are two ways to handle this problem. The first is to find a way to restrict our conception of theory so that it cannot play the requisite role in the argument. Given my reply to the No-Theory Objection, it would seem that I cannot complain on this score, since there I pointed out that we can take a more expansive conception of theory. But you might argue that any theory ought to be suitably general, or support interesting counterfactuals, or have a certain measure of explanatory power, or what have you. As far as I can see, such a move is compatible with my reply to the No-Theory objection, at least as long as you place our ordinary beliefs at a suitable remove from the theories that they are supposed to evince (e.g., it won't work if you say that each belief is a mini-theory, but it probably will work if you say that our beliefs presuppose certain theories).

I think that this response is viable, but working out the details would be a significant and unnecessary digression, as there is an easier reply available: namely, that the Modal Collapse Objection reduces to the (yet-to-be-discussed) Conflict Objection. I'd wager that there is a model of the logic I believe in which I run a three minute mile; however, I doubt that there is any such model of the biology I believe. In this respect – and no doubt in others – logic and biology are not in agreement about what I can do. Analogously, there is probably a model of the melittology³⁶ I believe in which there are only one million honey-producing beehives in the U.S.,

³⁶ Melittology is the study of bees; apiology is the study of honeybees specifically. Of course, my knowledge of bees of any kind is very limited; if I know anything beyond the platitudes about bees (they are black and yellow;

though there is no such model of my theory about what's actual. I need to say something about how disagreements between theories are to be adjudicated, and unless it's 'Always defer to the more conservative theory', then it should go a long way toward undermining the Modal Collapse Objection. So, on to the Conflict Objection.

2.5.5 The Conflict Objection

Here is the Conflict Objection. There is probably no model of the biology I believe in which a human being lives to her five hundredth birthday; but I'd bet that, from the perspective of the physics I believe, this state of affairs is possible – if wildly unlikely. What does TT recommend in this sort of case? Should I defer to one theory or the other? If so, which one? And why? Alternately, am I justified in believing both claims – indexed, perhaps, to the appropriate 'sort' of possibility? Or, do the two theories 'cancel out', so that I am not justified in believing anything about the matter?

We can make the problem sharper still by recalling POSS and NEC, TT's clauses about possibility and necessity:

[POSS] You justifiably believe that p is possible iff (a) you justifiably believe a theory that says that p is possible, (b) you believe that p is possible on the basis of this theory, and (c) you have no defeaters for the belief that p is possible.

[NEC] You justifiably believe that p is necessary iff (a) you justifiably believe a theory that says that p is necessary, (b) you believe that p is necessary on the basis of this theory, and (c) you have no defeaters for the belief that p is necessary.

TT gets into trouble here because POSS and NEC either (1) allow cases where you are justified in believing both that it is possible that p and that it is necessary that $\sim p$ or (2) lead to the

they can sting you; they make a 'bzzzzing' sound; etc.), then that's not a result of knowing much in the way of theory, but only because I spent an hour browsing Wikipedia entries in the process of writing this section. So most of *my* modal knowledge about bees is parasitic on the knowledge enjoyed by experts. I expect that this is the normal state of affairs, and I doubt that it is good news for the scope of our modal knowledge (see below).

conclusion that you are not justified in having either belief, since each constitutes a rebutting defeater for the other. This is the Conflict Objection.

A natural first step is to introduce kinds of possibility and necessity. The new version of NEC might go as follows:

[ϕ -NEC] You justifiably believe that p is ϕ -necessary iff (a) you justifiably believe a theory that says that p is ϕ -necessary, (b) you believe that p is ϕ -necessary on the basis of this theory, and (c) you have no defeaters for the belief that p is ϕ -necessary.³⁷

We now face the question: under what circumstances will a theory say that p is ϕ -necessary rather than necessary *simpliciter*? We can think about this in terms of the theory's generality. If a theory purports to apply to *all* phenomena, then if it says that p is necessary, it says that p is necessary *simpliciter*. However, if a theory limits itself to *physical* phenomena – or biological phenomena, or terrestrial biological phenomena, or carbon-based terrestrial biological phenomena – then it's modal claims are restricted accordingly: the theory won't say that p is necessary *simpliciter*, but rather that p is physically necessary, or biologically necessary, or what have you. These 'types' of necessity are just restrictions on the space of possibilities: after you set aside the non- ϕ -worlds – i.e., the worlds where the actual ϕ -laws³⁸ don't hold – ϕ -necessity is just necessary *simpliciter* evaluated just with respect to the ϕ -worlds.

Now the Conflict Objection goes as follows. p is ϕ -possible according to Theory #1; $\sim p$ is ψ -necessary according to Theory #2. We want to know whether p is possible or impossible

³⁷ N.B., there are actually two ways to interpret this version of NEC. On one, the theoretical definition is less specific, but the theoretical hypothesis indicates that only certain models are to be taken as representing possibilities – namely, the ones that satisfy certain assumptions about the logic in play, or the physical environment, or what have you. On the other, this information is built into the theoretical definition from the outset. For my purposes, the choice between these options is largely arbitrary.

³⁸ I don't think that a commitment to ϕ -necessity requires a commitment to laws, though the two obviously sit well together. You can get the same result by saying that ϕ -necessity is truth at all the worlds of which the final ϕ -theory is true (i.e., biological necessity is truth at all of the worlds of which final biology is true). You might also be able to get the same result by saying that ϕ -necessity is truth at all the worlds that are relevantly similar to the actual world, where the relevance relation is cashed out in terms of particular functional relationships, or systems, or mechanisms, or whatever (depending on the value of ϕ).

simpliciter. What should we believe? Let's make this concrete. Some botanists think that it is biologically impossible to have genetically engineered (GE) crops adjacent to non-GE crops without cross-contamination within a single generation.³⁹ I take it, therefore, that there is no model of the biology they believe in which GE crops are adjacent to non-GE crops without the relevant contamination. However, I also take it that this is because their models assume certain statistical facts about wind patterns, the behaviors of insect and animal populations, and so on. From the perspective of physics, though, I take it that there is no reason why cross-contamination must occur. It is wildly unlikely, of course, that pollen from one group of plants wouldn't make its way to an adjacent group, but in an indeterministic framework, nothing *rules out* this possibility; the thing to say is just that the odds are very, very low that it will be actualized. But that, of course, is to agree that the possibility is genuine. On the assumption that I've correctly described the situation, the question now is this: is it or is it not possible *simpliciter* that GE crops co-exist alongside non-GE crops without contamination? In this case, I see no reason to defer to the biological theory; the botanists make plausible assumptions for their purposes, but they are assumptions that needn't be true, as the physical theory shows. So, it seems to me that co-existence is possible *simpliciter*.

At this juncture, you might wonder whether we should *always* defer to the more general theory. The answer, in short, is that you should – unless there are specific reasons to doubt its deliverances. (Logic is just this sort of exception, which gives false positives because we are often insensitive to the forms of the propositions expressed by the sentences to which we give

³⁹ If true, this has significant ramifications. The Monsanto Company is a huge biotech firm that owns the rights to a number of GE seeds. Many, many U.S. farmers use these seeds, and a few years ago, Monsanto made the news for suing the smaller farmers who haven't gotten on board. The problem is that these smaller farmers own fields that are adjacent to ones that use Monsanto seed. As a result of cross-pollination, the smaller farmers end up with GE crops for which Monsanto has the patent. As a result, the small farmers are forced to destroy their seeds at the end of each season – else they violate patent law – which means that Monsanto is making small-scale farming prohibitively expensive, at least if it is going to be done without GE seed.

assent.) But this point shouldn't be taken to license too much optimism when it comes to the scope of our modal knowledge – or at least *my* modal knowledge. In the above case, I suggested that our physical theory has a model in which grains of pollen do not make it from one field to another for an entire growing season. In truth, though, I have no idea whether this is the case. First of all, my grip on the relevant physics is pretty weak. For virtually all intents and purposes, I have a Newtonian conception of the physical world, chastened only by a concern to agree with the things said by people who actually understand quantum mechanics. Moreover, because the relevant physics is obscure to me, the *application* of the relevant physics is obscure to me. When I reason about these matters, I assume that the upshot of indeterminism is that just about any momentary or sustained configuration of objects is possible, though most are wildly improbable. But I certainly don't have an argument for this, and for all I know, the thesis of indeterminism is supposed to apply only to certain microphysical particles, and not to things like grains of pollen. Second, insofar as I even *believe* the relevant physics, I probably believe it second-hand, which is to say that I believe it in virtue of having the disposition to defer to experts in my community on physics-related matters. If so, though, then my modal judgments are not really based on *physics*, but rather on the cartoon version of physics that I've pieced together from various (more or less reliable) sources. And I have a readymade defeater for any such judgment: since I have only the weakest grip on the real theory, I have only a very limited ability to tell where my cartoon version deviates from it; so, where my confidence about the parallel is low, I should be accordingly cautious in my modal judgments. In my view, this defeater is an ineliminable part of living in a society where there is such a significant division of cognitive labor. In any case, the essential point here is that it is not nearly so obvious that I believe a physical theory that has the

relevant model – or that if I do, that I’m also justified in believing that the model represents a genuine possibility.

No doubt there are metaphysical theories that I might invoke to get the co-existence result, and in a different context, it would be interesting to explore what they are. Here, though, let it suffice to say this: whatever the package of metaphysical views, the task is to show that I am justified in believing it (if believe it I do). I don’t doubt that this can be done, but it’s worth stating the obvious: it’s hard to provide a compelling argument for any metaphysical view, even if the members of the target audience are (the mythical) disinterested third parties to the debate. Granted, providing a compelling argument may not the only way to justify belief in a theory, but I would think that one is needed here. After all, if metaphysicians have shown nothing else, it’s that wildly different theories can be made to agree with common sense. This makes it much less plausible that we can rely on common sense to determine the metaphysical theory that we believe. But without a tight connection between common sense and a particular metaphysical theory, we shouldn’t assume that our theory will be able to borrow its justification from common sense. To be clear, I am not a skeptic about metaphysical knowledge. My point is just this: it isn’t *obvious* that I am justified in believing that GE crops can co-exist alongside non-GE crops without contamination. I know that I don’t understand the relevant physics, and it isn’t at all clear that I’m justified in believing a metaphysical theory that specifies the relevant model. If TT is true, then I am justified in believing that p is possible *simpliciter* if there is a p -model specified by the most general theory that I justifiably believe. In any given case, though, we need to press to see whether (a) I actually believe the relevant theory and (b) I am justified in believing it. The thrust of the last few paragraphs is that we shouldn’t be blithely optimistic that

I satisfy (a) and (b) – and insofar as I am not an outlier, we shouldn't be blithely optimistic that others do either.

Let's take stock. TT says that, where p is any modal claim, you justifiably believe that p iff (a) you justifiably believe a theory that says that p , (b) you believe that p on the basis of this theory, and (c) you have no defeaters for the belief that p . The Cheating Objection aside, the above objections amount to a dilemma: the claim is that TT must choose between preserving most of our pre-theoretic modal judgments and becoming overly permissive.⁴⁰ Essentially, I've argued that we can navigate this dilemma by adopting these claims:

1. We justifiably believe some folk theories.
2. Logical possibility outstrips metaphysical possibility.
3. We should take p to be possible *simpliciter* if it is possible according to the most general theory that we justifiably believe.
4. We shouldn't be too quick to assume that we justifiably believe a theory according to which p is possible.

There is more to be said in defense of (1) – (4), but they are not wildly implausible as philosophical theses go. With them in hand, the above objections pose no serious threat.

2.6 Looking Ahead

With TT on the table, and with the first round of objections behind us, we turn to two thornier problems. Chapter 2 takes up Condition A, which requires that we justifiably believe a theory that says that p . Is it plausible that we are justified in believing theories to be true – especially if they have the modal content that the semantic view attributes to them? To defend an affirmative answer to this question, I make a case for the epistemic credentials of IBE.

⁴⁰ I haven't questioned the assumption that we *should* preserve many of our pre-theoretic modal judgments. I think that (1) is probably the most contentious of the four theses, and if we were to abandon it, then TT would be much less hospitable to ordinary opinion. I don't know whether it would be worse for that.

Chapter 3 considers Condition C – TT’s no-defeater clause. The question is whether we can acknowledge our causal isolation from the truthmakers for modal claims without acquiring a general defeater for claims to modal knowledge. I contend that we can.

3. CONDITION A

3.1 Introduction

Recall:

[TT] You justifiably believe that p iff (a) you justifiably believe a theory that says that p , (b) you believe that p on the basis of this theory, and (c) you have no defeaters for the belief that p .

This chapter is concerned with Condition A. Note that we can split Condition A into two parts: the first part requires that we justifiably believe a theory (rather than, say, merely accept it as true for certain purposes); the second specifies a feature that the theory must have (namely, that it ‘says that p ’). This chapter is concerned with the first part of Condition A: for TT to work, we need story about how we can justifiably believe our best theories. Ultimately, the aim of this chapter is to defend the epistemic credentials of IBE, which – given its appeal to so many scientific realists – is the obvious companion to TT. However, before addressing that matter, I should begin by clarifying what’s involved in a defense of IBE. We can see this by considering a challenge from Colin McGinn.

3.2 Modal Content and Explanatory Virtues

If our theories have modal content, and if we can choose among candidate theories using IBE, then IBE can lead us to prefer one set of modal claims over another – based, we’ll assume, on broadly empirical considerations. But McGinn argues that this conclusion is false; he thinks that modal questions are ‘empirically conservative’ – i.e., unanswerable on any empirical basis.

Allow me to quote him at some length:

Let T be a theory free of modal expressions: its generalizations apply only to all actual objects of the kind treated by the theory, past, present, and future; and suppose we know

T to be true, presumably by observation and induction. (We can think of T as got by taking a scientific theory containing modal expressions, implicitly or explicitly, and removing its modal content.) T will have certain empirical consequences which are used to verify it. Now add to T some causally modal constructions V – a nomic necessity operator, a dispositional suffix, a counterfactual conditional (possibly embedded in a modal logic appropriate to the modal notions introduced) – and suppose that the resulting theory $T \cup V$ is realistically true. Then the modalized theory $T \cup V$ is, for a modal realist, a factually stronger theory than the non-modal T : it reports, not just what actually happens, but what happens in all causally possible conditions. But clearly $T \cup V$ is a conservative extension of T [i.e., $T \cup V$ has no empirical consequences not shared by T]. So modality [...] is empirically conservative. The reason is obvious: empirical consequences are reported by sentences which can be observed to be true, but what is non-actual cannot be observed to be true. [...] It follows, I think, that we could not plausibly be said to come to know a theory with such modal content by purely empirical means: for the empirical consequences of the theory are compatible with a weaker theory in which modalities do not figure. So [...] modality [is not] properly empirical, and [modal sentences are] different from other sentences whose subject matter is admittedly not itself observable; for example, theoretical sentences. *And this suggests a restriction upon the propriety of epistemological theories of a given type of statement which are based on the idea of inference to the best explanation: namely, that the statements in question should not be empirically conservative.* If this is correct, the liberalized empiricist can insist upon the non-empirical nature of modality while not simultaneously and unwantedly excluding statements whose epistemological credentials he finds (or should find) acceptable [e.g., ones about theoretical entities] (McGinn 1999, 101-102, my emphasis).

At first blush, anyway, McGinn's argument assumes the syntactic view of theories, and one might hope that his argument could be blocked just by the transition to the semantic view. But things are not so simple. True, on the semantic view, a theory's modal content isn't present because of operators or suffixes or conditionals, but that doesn't imply that it can't be stripped out. On the semantic view, a theory has modal content if it has a theoretical hypothesis according to which, if the theory is true, some of its models correspond to non-actual possibilities.⁴¹ If you want a version of the theory *without* modal content, you just need to doctor

⁴¹ Incidentally, McGinn rejects the possible worlds framework, instead taking modality as the primitive modes in which objects have properties (see (McGinn 1999, 84-97) and (McGinn 2000, 69-86)). Hence, one might think that he denies that there are non-actual possibilities to which the models could correspond. Not so. McGinn objects to worlds on ontological grounds, but he has no worries about their heuristic value. So, unless McGinn also has worries about abstracta (and he doesn't), then we're free to be ersatzists about worlds, building them out of propositions or object-mode-property triples or what have you. We can grant, then, that the *truthmakers* for modal

the theoretical hypothesis accordingly. So, we won't be able to challenge McGinn by running a kind of indispensability argument: according to our best account of theories, theories have modal content; since we're justified in believing some theories to be true, we're stuck with some modal commitments; therefore, we must make room for modal knowledge in our epistemology for scientific theories, which means making room for the view that some modal claims are *not* empirically conservative. McGinn's challenge requires a different response.

So let's grant that McGinn is right: you can take theory T , strip out its modal content, and end up with a non-modal theory, $T\text{-minus}$, that has the same empirical consequences as T . Even if correct, we need a further assumption to show that empirical considerations can't select between T and $T\text{-minus}$. More precisely, given that we choose between empirical theories using IBE, we need two assumptions so that IBE will always recommend $T\text{-minus}$ over T . First, if we think of IBE as an argument, then it has this form:

1. Facts $f_1 - f_n$ obtain.
2. If true, hypothesis h would offer a better explanation of $f_1 - f_n$ than would any competing hypothesis.
3. So, probably, h is true.

For present purposes, the key thing to note is that IBE recommends $T\text{-minus}$ only if $T\text{-minus}$ explains $f_1 - f_n$ better than any competing hypothesis. And surely $T\text{-minus}$ explains better only if $T\text{-minus}$ explains. In other words, to make McGinn's argument work, we need to assume that $T\text{-minus}$ has some explanatory power. Second, we need to make an assumption about how IBE selects among competing hypotheses. Here is one such proposal (where 'P+S' stands for 'predictive power and simplicity'):

[P+S] Where T is any theory, IBE recommends T only if (a) T 's known⁴² competitors

claims are McGinn's primitive modes while still maintaining that modal claims (and their truth values) can be mapped systematically to our set-theoretic structures, and thereby to the worlds. So, if the models correspond appropriately to these ersatz worlds, then they thereby correspond appropriately to the truthmakers for modal claims.

⁴² Where a theory is a known competitor just in case it is a competitor of which you are aware – not just in case it is

are either less than or equal to T in terms of predictive power and (b) for any known competitor T^* that is equal in predictive power to T , T is simpler than T^* .

If [P+S] is the right model of IBE, then McGinn's conclusion follows: *ex hypothesi*, T and T -minus have the same empirical consequences and T is more complex than T -minus; hence, given [P+S], if we are justified in believing any theory in this context, it's going to be T -minus rather than T . So, if T -minus has some explanatory power, and if IBE boils down to [P+S], then McGinn's argument goes through.

I think that both assumptions are false. First off, prediction (and retrodiction) is not explanation, and T -minus only offers predictions (and retrodictions). Suppose that an oracle perfectly predicts the evolution of a system (e.g., if the oracle predicts that object o will be at location l at time t , then o is at l at t). Here is a theory: what the oracle says is true. This theory has marvelously accurate empirical consequences, but it sheds no light whatever on why the system behaves as it does, and hence provides no way to judge how the system would evolve in different circumstances. Our 'theory' does not explain anything.⁴³

Second, even I am wrong about the explanatory power of T -minus, it remains the case that [P+S] is an impoverished account of IBE. The standard account recognizes three other explanatory virtues besides predictive power and simplicity: namely, fecundity (or generality), modesty, and conservatism. So, if McGinn means to reduce IBE to predictive power and simplicity, then he owes us one of two arguments. First, he could try to convince us that we should not recognize the other (purported) explanatory virtues – and this despite their apparent role in explaining why we make the inferences we do. Second, he could try to show that, even if

a competitor that you know to be true.

⁴³ My own view is that explanations articulate dependence relations among various phenomena (in particular see (Kim 1994); also (Woodward 2003)). But dependence is a modal notion: x depends on y just in case, if y were not to obtain, then x would not obtain (suitably Chisholmed to handle finkish dependencies). So, I see no hope for the view that T -minus explains.

we recognize the other virtues, *T*'s greater complexity can't be outweighed by scoring well enough on those virtues. In view of the prominence of modal notions in our everyday and scientific reasoning, I doubt that it will be easy to make either case.

So McGinn hasn't shown that IBE cannot recommend a theory with modal content. But McGinn has brought into focus some of IBE's non-trivial commitments regarding explanation and the explanatory virtues. The disagreement about the nature of explanation runs deep; it is, perhaps, among the most fundamental disagreements in philosophy. Many of us, though, will be glad to adopt a more substantive view of explanation than those who would insist that explanation is just prediction and retrodiction – a view taken by, e.g., the early proponents of the deductive-nomological theory of explanation; e.g., (Hempel and Oppenheim 1948).⁴⁴ Still, this move does not put IBE in the clear. McGinn's objection forces us to be clear about both the explanatory virtues that IBE involves and the way that IBE deploys them to make its recommendations. Once we do, it isn't obvious that IBE can provide the justification that it promises. Is simplicity – or conservatism or modesty – really a guide to truth?

3.3 A Deeper Challenge

What, exactly, is the problem? The most common complaint against IBE is that some of the explanatory virtues either (a) amount to substantive – and probably false – assumptions about the world or (b) provide merely pragmatic reasons to accept a hypothesis as true, not epistemic reasons to believe it.⁴⁵ Either way, IBE cannot provide the justification that its proponents want from it. D. L. Gunner seems to be expressing the first view when he says that “a question should

⁴⁴ For the standard catalogue of problems with the view that explanation is prediction and retrodiction, see (Salmon 1989).

⁴⁵ Both accusations are usually leveled against modesty and simplicity, although they can be directed toward other explanatory virtues. (a), for example, may be the upshot of van Fraassen's criticism of 'privilege' as way to defend IBE, which amounts to an attack on conservatism. See (Van Fraassen 1989, 143-144).

be raised as to whether the principles of parsimony and simplicity have not become restrictive principles of stinginess and over-simplification.” For all we know, he contends, “[n]ature is lush, prodigal, messy, wasteful, sexy, etc.” (Gunner 1967, 4-5). Timothy Day and Harold Kincaid take the same line. They maintain that “appeals to the best explanation are really implicit appeals to substantive empirical assumptions, not to some privileged form of inference. It is the substantive assumptions that do the real work” (Day and Kincaid 1994, 282). We also get this challenge from Paul Benacerraf and Hilary Putnam:

But why should the simplest and most conservative system (or rather, the system that best balances simplicity and conservatism, by our lights) have any tendency to be *true*? Quine, good pragmatist that he is, tends to pooh-pooh this sort of question; but more realistically minded philosophers are sure to be bothered. It is hard enough to believe that the natural world is so nicely arranged that what is simplest, etc., by *our* lights is always the same as what is *true* (or, at least, *generally* the same as what is true); why should one believe that the universe of sets (or the totality of modal truths) is so nicely arranged that there is a preestablished harmony between *our* feelings of simplicity, etc., and *truth* (Benacerraf and Putnam 1983, 34).⁴⁶

The alternative is to write off some of the explanatory virtues as wholly pragmatic – i.e., not concerned with truth at all. Here is Bas van Fraassen’s way of putting this objection:

Judgments of simplicity and explanatory power are the intuitive and natural vehicle for expressing our epistemic appraisal. [But these] are specifically human concerns, a function of our interests and pleasures, which make some theories more valuable or appealing to us than others. Values of this sort [...] provide reasons for using a theory, or contemplating it, whether or not we think it true, and cannot rationally guide our epistemic attitudes and decisions. For example, if it matters more to us to have one sort of question answered rather than another, that is no reason to think that a theory which answers more of the first sort of questions is more likely to be true (not even with the proviso ‘everything else being equal’). It is merely a reason to prefer that theory in another respect (Van Fraassen 1980, 87).

Let’s call these ‘the Substantive Assumptions Objection’ and ‘the Pragmatic Reasons Objection’.

I submit that both objections misfire: the explanatory virtues are not working assumptions about

⁴⁶ See, too, Richard Fumerton: “A theory which is more consilient and simple than alternatives [...] is certainly more desirable than its competitors in the sense that it would be *nice* if it turned out to be true. But this not being the best of all possible worlds (some theologians aside) what would be nice is not always so” (Fumerton 1980, 596).

the world, nor do they offer merely pragmatic reasons. Rather, the explanatory are properties of hypothesis that have epistemic value, and so generate epistemic norms. Therefore, they are genuinely epistemic in nature. The purpose of this chapter is to offer an account of IBE that substantiates this claim, immunizing IBE to the objections above. Of course, these objections are not the only ones that IBE faces. Still, they are perhaps the most significant ones, and success here would provide cause for optimism about tackling the others.

Before diving in, we should ask about the reasoning behind these objections. I think that both objections arise in the same way. We begin by considering simplicity (say) in isolation from the other explanatory virtues. When we do so, we immediately note that it would be foolish to adhere to the following principle:

[S] Believe the simplest hypothesis.

Plainly, [S] is not truth-conducive: Thales' hypothesis ('All is water') is certainly simple, but it is not therefore what we ought to believe. But if IBE is truth-conducive, then surely each of the explanatory virtues is truth-conducive. And so we have two options. First, we can try to reconstruct what someone would need to believe in order to prize simplicity as a truth-conducive virtue of theories. Our hypothesis: "If IBE's proponents thought that the *world* were simple, then they might also believe that simpler theories are more likely to be true than their complex competitors."⁴⁷ To go this route is to say that IBE's epistemic credentials depend on the merits of the hypothesis that the world is simple – i.e., to say that IBE provides epistemic justification, but not *independent* epistemic justification. Second, we can flatly reject the claim that all the

⁴⁷ This line of reasoning has some historical precedents. Mill, for example, ends up in the same place about induction: "We must first observe that there is a principle implied in the very statement of what induction is; an assumption with regard to the course of nature and the order of the universe; namely, that there are such things in nature as parallel cases; that what happens once will, under a sufficient degree of similarity of circumstances, happen again, and not only again, but as often as the same circumstances recur. [...] Whatever be the most proper mode of expressing it, the proposition that the course of nature is uniform is the fundamental principle, or general axiom, of induction" (*A System of Logic*, Book III, Chapter III, §1).

explanatory virtues are *epistemic* virtues, arguing that they are pragmatic instead. Either way, IBE epistemic credentials are wanting.

The argument behind this reasoning seems to go like this:

1. If IBE can provide independent epistemic justification, then each explanatory virtue must be truth-conducive.
2. It is not the case that each explanatory virtue is truth-conducive.
3. So, IBE cannot provide independent epistemic justification.⁴⁸

Call this ‘the Main Argument Against IBE’ (or ‘the Main Argument’, for short). Some try to block the Main Argument by contending that each explanatory virtue *is* truth-conducive – i.e., they deny Premise 2.⁴⁹ But I happen to think that Premise 2 is true. My claim is that we should not accept Premise 1. The most serious problem with Premise 1 is that it appears to rely on epistemic value monism – the view that there is only one property with intrinsic epistemic value – to justify Premise 1’s austere stance on what counts as an epistemic reason. But epistemic value monism is false. A second problem with Premise 1 is that it depends on an unjustified claim about how IBE works. So, even if I am wrong about the falsity of epistemic value monism, the Main Argument is still in trouble. Now, even if both of these objections are successful, they do not show that IBE can provide epistemic justification; they simply show that the Main Argument fails. I close, then, by sketching a positive account of IBE’s epistemic credentials.

⁴⁸ I will suppress references to independence from here on out. Also, I should note that IBE also might fail to be an independent source of justification because, say, we trust it only insofar as we can offer a track-record argument for its reliability (via simple enumerative induction). But I have yet to see an argument for this view that doesn’t rely on either the Substantive Assumptions or Pragmatic Reasons Objection.

⁴⁹ Swinburne, for example, thinks that “it is a fundamental *a priori* principle” that simpler theories are more likely to be true than are more complex ones (Swinburne 2001, 102).

3.4 Epistemic Value Monism

Here is a summary of the argument of this section. *If Premise 1 is true, then epistemic value monism is true. Epistemic value monism is false. So, Premise 1 is false. If Premise 1 is false, then the Main Argument fails. So, the Main Argument fails.* The first two moves are the ones that need defending. I take up the first in the next subsection.

3.4.1 If Premise 1 Is True, Then Epistemic Value Monism Is True

In the passages quoted above, Benacerraf, Putnam, and van Fraassen⁵⁰ place great weight on the distinction between epistemic and pragmatic reasons, and they cash out the distinction in terms of truth-conduciveness: i.e., epistemic reasons are the reasons that are directed toward truth, or aim to secure the truth; pragmatic reasons, by contrast, aim to secure other, non-alethic goods. This does not imply, of course, that pragmatic reasons can never lead us to truth. Rather, the lesson is that it is something of an accident when they do. With this in mind, recall Premise 1:

1. If IBE can provide independent epistemic justification, then each explanatory virtue must be truth-conducive.

Why should we believe this premise? Well, it's all but analytic that only epistemic reasons can provide epistemic justification. If we accept the view that epistemic reasons are distinguished by their truth-conduciveness, then it follows that only truth-conducive reasons can provide epistemic justification. Of course, just like truth, the explanatory virtues are not themselves reasons; in the case of the virtues, they appear to be properties of hypotheses. However, of each property – explanatory virtue or not – we can consider why it might be worth pursuing. By van Fraassen's lights, the crucial question is: "Is this property worth pursuing because it is correlated

⁵⁰ For ease of exposition, I'll just refer to van Fraassen from here on out.

with truth?” If the answer is affirmative, then we can say that the relevant property generates an epistemic reason; if the answer is negative, then the relevant property does not generate an epistemic reason. And, when we put this question on the table, we expose the assumption that truth is what’s worth pursuing in itself; other properties are, at best, instrumentally valuable insofar as they point toward truth. So, if truth is all that’s of intrinsic epistemic value – i.e., if epistemic value monism is true – then Premise 1 is true. The upshot here is that, if we ask why we ought to believe Premise 1, then we will probably be given an argument that invokes epistemic value monism. But if epistemic monism is false, then this argument will fail, and so we will be left without a reason to believe Premise 1. And, without a reason to believe Premise 1, the Main Argument is in trouble.

3.4.2 Epistemic Value Monism Is False

Michael DePaul offers a simple argument against epistemic value monism (DePaul 2001). He points out that the fundamental epistemic aim is not truth, but *knowledge*. And, unless you think that knowledge is merely true belief, you must grant that at least one other property is of epistemic value: entitlement, or justification, or warrant, or whatever it is that you take to close the gap between true belief and knowledge.

I find this argument persuasive. However, the argument leaves it open for someone to try to reduce the value of justification (for example) to its truth-conduciveness. This is because DePaul assumes that knowledge is of greater epistemic value than mere true belief; he then seeks an account that explains this fact. He is not concerned to demonstrate that epistemic value monism is unsustainable (or, more carefully, sustainable only at great, great cost). I take up this more ambitious project here.

3.4.2.1 Utilitarianism in Ethics and Epistemology

It will be helpful to consider an analogy, so let's take a detour. Utilitarianism, at least in its simple, hedonistic form, is committed to both *value monism* and *proceduralism*. An ethical theory is committed to the former if it maintains that all situations have only one morally relevant feature; according to utilitarianism, that feature is well-being. An ethical theory is committed to the latter just in case it says that there is a decision procedure for determining whether an action is obligatory, permissible, or wrong; for utilitarians, this is the principle that you should maximize well-being.⁵¹ With these two points in mind, and simplifying greatly, we can represent utilitarianism as a function: it takes a set of action / outcome pairs as inputs, selects the one with the greatest overall well-being, and gives the action that leads to that situation as the output; that action, of course, is the one that utilitarians judge to be obligatory.

The function just outlined represents act utilitarianism. How would we need to modify it in order to represent rule utilitarianism? Easy. We replace the set of action / outcome pairs with a set of slightly more complex pairs, the first member of which is a candidate moral rule, the second of which is the outcome that would result from universal adherence to that rule. The function still selects the one with the greatest well-being. However, instead of giving an obligatory action as an output, it gives a moral rule; we then apply the rule to our situation in order to determine what's obligatory.

Rule utilitarianism is an ethical theory, but its structure is shared by an approach to epistemology. Instead of candidate moral rules, the first member of each pair is a candidate

⁵¹ Utilitarians, of course, eschew the permissible. Also, you might think that utilitarianism says that you should maximize *expected* well-being, not well-being *simpliciter*. Utilitarians diverge here. Some take the line just mentioned. Others, however, will say that praise or blame may be due to you only insofar as you do or do not successfully maximize expected well-being, but the rightness or wrongness of your action is dependent solely on the net well-being that actually results. I take the latter camp to have the more plausible position. If the former camp is right, then either utilitarianism is not a form of value monism or *expected* well-being is what's of value, not well-being itself (since their proposed maxim follows directly from their value monism; see below). But surely the former option is implausible, and the latter does not sit well with the most basic motivations for utilitarianism.

epistemic policy; instead of global outcomes, the second member of each pair is the number of truths that would be believed if that policy were followed. Instead of selecting the outcome with the greatest well-being, our new function selects the one with the greatest number of resultant true beliefs, giving that epistemic policy as an output. As before, the output is not itself the obligatory action; rather, the output is the principle that determines what you ought to believe in a given circumstance – equivalently, it determines the belief that you would be justified in holding in those circumstances. Let's call this view *truth-maximizing rule utilitarianism in epistemology* (TRUE). Like its cousin, TRUE is a form of value monism: it takes truth to be the only feature of a belief that is of worth. Also like its cousin, TRUE is a form of proceduralism: it takes there to be a straightforward decision procedure that settles which of the many possible epistemic policies is correct. Why does it recommend maximizing true beliefs? As in ethics, your theory of value drives your theory of the right: if you think that only well-being is of moral worth, then it is hard to see what you would recommend *other* than maximizing well-being. After all, if well-being is of moral value, then surely more is better, at least if all other things are equal. And if value monism is true, then all other things always *are* equal – there is never anything else with which well-being competes. So, you should maximize it. The same argument applies, *mutatis mutandis*, to truth given TRUE.

TRUE is probably not just a form of epistemic value monism: it is probably the only epistemological position that is plausible if epistemic value monism is correct. As I suggested in the preceding paragraph, it's likely that epistemic value monists are committed to an epistemology that is structurally analogous to utilitarianism. But in epistemology, the analog of act utilitarianism is hopeless: *that* view would say that a belief is justified iff it's true, since (a) such a view would only take into consideration the local features of the belief and (b) such a

view would take the truth of that belief to be the only feature that matters. But, of course, it is not the case that a belief is justified iff it's true. The analog of rule utilitarianism, TRUE, avoids this problem by introducing the epistemic policies: they are designed to take non-local factors into account – namely, the number of true beliefs that would be achieved given universal adherence to the epistemic policy – thereby preventing TRUE from having the awkward consequence that sinks the epistemic analog of act utilitarianism.

3.4.2.2 The Problem with TRUE

However, we should reject TRUE. According to TRUE, the right epistemic policy is the one that, if followed, maximizes the number of true beliefs. But there is no viable interpretation of this recommendation. So, we should reject TRUE.

Here is the argument. I suggested that we can represent TRUE as a function: the inputs are policy / success rate pairs, the output is the most truth-conducive epistemic policy. I also intimated that 'being the most truth-conducive epistemic policy' means 'being the policy that would produce the greatest number of true beliefs if it were followed'. But this can't be right. The policy that will do best here is the one that tells us to believe *everything*. If truth is the *only* valuable doxastic feature, then there is no value to avoiding falsehood. So, if we were to believe every proposition and its negation, then we wouldn't miss out on any truths, thereby maximizing what's of epistemic value.⁵² But this is ridiculous.

To avoid this problem, perhaps we should make a friendly amendment to TRUE. We'll still say that truth is still the only valuable doxastic feature, but we'll add a principle called

⁵² *Objection:* We can't believe contradictions, epistemic policies create epistemic obligations, and we aren't obligated to do the impossible; so, we can't be obligated to believe every proposition and its negation, which means that this policy is not in the running. *Reply:* It's not at all clear to me that we can't believe explicit contradictions. But even if that's right, then we certainly *can* believe implicit contradictions. In other words, even if we can't believe p & $\sim p$, it's surely the case that we can believe p and we can believe $\sim p$.

‘NOFALSITY’, according to which believing falsely has epistemic *disvalue*. Call our revised version of TRUE – i.e., the conjunction of TRUE and NOFALSITY – ‘T&~F’. T&~F preserves the spirit of TRUE, if not the letter. Problem solved?

No. Now, the most straightforward interpretation of ‘being the most truth-conducive epistemic policy’ is something like ‘being the policy that maximizes the ratio of truths to falsehoods believed’. This looks like a recipe for radical epistemic caution: if you take this policy seriously, then the best move is to believe only self-evident truths. If you believe even one falsehood, then it doesn’t matter how many truths you believe, since your ratio of true to false beliefs will invariably be lower than it would have been had you believed no falsehoods at all. But as long as you find at least one self-evident truth (the *cogito* or your favorite tautology) and you believe no falsehoods whatever, your ratio will be as high as it possibly can be.⁵³

So we seem to be torn between two extremes: either radical epistemic abandon (believe everything) or radical epistemic caution (believe only the self-evident). You might object I’m assuming both more and less control over our beliefs than is plausible, ignoring (a) that you can’t believe whatever you want, so you can’t believe everything (which is supposed to undermine my objection to TRUE) and (b) that so many of our beliefs form spontaneously, so we can’t limit ourselves to believing a single self-evident truth (which is supposed to undermine my objection to T&~F). I grant both (a) and (b), but they make no trouble for my argument. Concerning (a),

⁵³ *Objection:* Any number over zero isn’t a ratio (it’s ill-defined); so, you would have to believe at least one falsehood to achieve the goal of maximizing the *ratio* of true to false beliefs. *Reply:* First, the ‘maximize the ratio’ formulation isn’t mine; it’s common enough in the literature: see, e.g., (Nozick 1993, 69). Second, it’s easy enough to recast the conversation in terms of maximizing the percentage of truths believed, in which case my argument stands. And third, you can still make the ratio version work. Suppose that you believe one self-evident truth and believe its negation; you then believe as many propositions as you can that are logically equivalent to the self-evident truth. Since there are infinitely many of them, you can make the ratio as high as your mental capacities permit (and this with minimal epistemic risk). *Objection:* Logically equivalent propositions are equivalent, period; so, this solution puts your ratio at .5. *Reply:* Logically equivalent propositions are not equivalent, period. If they were, then ‘red is a color’ and ‘ $2 + 2 = 4$ ’ would express the same proposition, since they both express necessary truths. And that’s absurd.

is it really just your *inability* to believe everything that makes it a terrible epistemic policy? If TRUE is correct, then this seems to be the case. Surely it isn't, though. Even if it *were* psychologically possible to believe indiscriminately, it would be a poor plan. And the same point applies to (b): even if it *were* psychologically possible to believe only the self-evident, would this be a good epistemic policy? If T&~F is correct, then the answer is 'Yes'. But surely this would be a mistake.

Here is a further consideration. Perhaps some beliefs are inescapable: even if we judge them to be false, we cannot abandon them. If there are such beliefs, though, and we indeed judge them to be false, then surely we can still recognize the epistemic tension that this creates. I suspect that something similar is the case when we judge the risk of error to be unacceptably high: whether or not we can actually abandon the beliefs in question, if we judge the risk of error to be too great, then surely we can judge them to be epistemically subpar. But when is the risk excessive? If TRUE is correct, then our only advice is to believe as many truths as possible; it follows that the risk is *never* excessive. If a belief is epistemically subpar just in case the risk of being wrong crosses some threshold, TRUE seems to suggest that we should *never* judge a belief to be epistemically subpar. Alternately, if T&~F is correct, then our only advice is to maximize the ratio of truths to falsehoods believed; now, the risk is excessive whenever there is a threat that we might *not* maximize that ratio, which is to say that it's excessive whenever we believe what isn't self-evident. T&~F seems to suggest, then, that we should almost *always* judge our beliefs to be epistemically subpar. So, whether supplemented with NOFALSITY or not, TRUE is in trouble.

Someone might object that it's uncharitable to articulate either TRUE or T&~F in terms of truth *simpliciter*. Rather, they should be cashed out in terms of *significant* truths (i.e.,

“maximize the number of significant truths believed’ or ‘maximize the ratio of significant truths to falsehoods believed’). I agree that it should be, but the proponents of TRUE and T&~F cannot. What makes some truths significant while others are not? Whatever it is, it’s something other than their mere truth – perhaps their usefulness, or their explanatory power, or their fit with what we believed pre-theoretically, or what have you. And, crucially, the significance of a belief either is or is not explicable solely in terms of its truth. If it is, then significance won’t save either TRUE or T&~F from the problems that I’ve been detailing, since there will never be a case in which significance trumps truth, and so gives you a reason to take an epistemic risk. But if significance *isn’t* explicable solely in terms of its truth, then to set significance alongside truth is to reject epistemic value monism, and hence to reject TRUE and T&~F.

3.4.3 Back to van Fraassen and IBE

Granted, I may have overlooked a perfectly good policy that’s based on the assumption that truth is the only thing of epistemic worth; if so, then TRUE’s devotees should provide it. Suppose they can’t. How should we diagnose the problem? Well, as I’ve indicated, rationally increasing your stock of beliefs beyond the self-evident requires a policy about the management of epistemic risk. Whatever policy you adopt, it will need to give advice having the following form: *risk error only if...*, where the ellipses stand for something *else* of epistemic worth. Your policy might be, for example, that you should risk error only if the proposition would increase the coherence of your belief system. Alternately, you may maintain that you should risk error only if the proposition in question seems to be true, absent any defeaters – this is the way taken by those who defend ‘phenomenal conservatism’ (e.g., (Huemer 2001)). If you go the first route, then you’re assigning epistemic value to coherence; if you go the second, then you’re assigning it

to conservatism. There are no doubt plenty of other options, but they'll all lead you to assign intrinsic epistemic value to something other than truth. In other words, they'll lead you to deny epistemic value monism.

Now, the Main Argument has it that, if IBE can provide independent epistemic justification, then each explanatory virtue must be truth-conducive: this is Premise 1. But if epistemic value monism is false, then there is no reason to believe Premise 1. And we've just seen that epistemic value monism *is* false. So, there is no reason to believe Premise 1, which means that the Main Argument fails.

With this objection in place, here is an analysis of the dialectical situation. Our critics of IBE – i.e., proponents of either the Substantive Assumptions or Pragmatic Reasons Objection – suppose that epistemic value monism is true. As a result, they take themselves to be within their rights to deny that simplicity or conservatism can generate epistemic reasons. But epistemic value monism is false. So, there is a class of properties that the critics judged to be pragmatic, and they should now admit that at least one of those properties can generate epistemic reasons. If they continue to disapprove of IBE, then although they *may* still have some principled objection, it begins to look as if they are simply disposed to a greater degree of epistemic caution than are the defenders of IBE. But a difference in temperament is not an argument, and I do not see why the proponents of IBE should be troubled by those who are more risk-averse than they are.

3.5 Some Models of IBE

To my mind, epistemic value monism is very implausible. But there are those who are loath to abandon it. Should they be moved by the Main Argument? No. Even if epistemic value

monism is true, the Main Argument is still in trouble: Premise 1 makes an unjustified – and implausible – assumption about how IBE works. So it may be the case that, if IBE can provide independent epistemic justification, then IBE is truth-conducive. However, it does not follow from this that *all the individual virtues* are truth-conducive.

IBE is normally presented as an argument with this form:

1. Facts $f_1 - f_n$ obtain.
2. If true, hypothesis h would offer a better explanation of $f_1 - f_n$ than would any competing hypothesis.
3. So, probably, h is true.

This formulation is useful for many purposes, but it obscures some of the issues on which I want to focus. Here is an alternative way of presenting IBE:

The virtues are properties of hypotheses. Suppose, just for simplicity's sake, that the virtues come in fixed units: hypothesis h_1 might have three units of simplicity, four of conservatism, ten of predictive power, etc. Moreover, there is a score corresponding to h_1 's three units of simplicity (' h_1 's v_s -score'), another score corresponding to h_1 's four units of conservatism (' h_1 's v_c -score'), and so on down the line. These scores are assigned by a function. A second function takes the individual v -scores as inputs and gives an overall score as the output. The overall score represents the hypothesis's explanatory power. IBE, then, recommends believing a hypothesis if it has a better overall score than any of its competitors.

With this sketch in hand, we can generate lots of different models of IBE. Here is one of the simplest ones:

The Very Simple Model. For all h and v , h 's score with respect to v is equivalent to the number of units of v that h has. And, for all h and v , h 's overall score is equivalent to the sum of h 's individual v -scores.

It follows from the Very Simple Model that IBE is truth-conducive only if each explanatory virtue is truth-conducive. After all, if even one of the virtues *isn't* truth-conducive, then there is nothing to prevent non-truth-conducive considerations from outweighing truth-conducive considerations: a false hypothesis can have greater explanatory power than a true one just by scoring well enough on a non-truth-conducive virtue (i.e., by having enough units of a non-truth-

conducive virtue).

But the Very Simple Model is not the only way to understand how IBE works. Rather than present concrete alternatives, though, let me offer two model schemas that should suffice to illustrate some of the available alternatives:

The Non-Additive Model Schema. For all h and v , h 's v -score is equivalent to the number of units of v that h has. However, for all h and v , h 's overall score is *not* equivalent to the sum of h 's individual v -scores: there is some other, non-additive function that determines h 's overall score, and therefore its explanatory power.⁵⁴

The Context-Sensitive Model Schema. For all h and v , h 's overall score is equivalent to the sum of h 's individual v -scores. However, for all h and v , h 's v -score is *not* determined solely by the number of units of v that h has; in addition, h 's v -score depends on contextual factors.⁵⁵

The Non-Additive Schema takes no stand on the function that determines h 's overall score other than that it is not additive; the Context-Sensitive Schema takes no stand on what counts as a context, nor on how much the context influences the score. Obviously, different views on these matters will affect IBE's recommendations.⁵⁶ Crucially, though, there are going to be lots of ways to fill in these schemas on which it is *not* the case that IBE is truth-conducive only if each

⁵⁴ This schema is inspired by (Kagan 1988).

⁵⁵ At first blush, Timothy Day and Harold Kincaid defend a view of IBE that sounds quite a bit like the Context-Sensitive Model; see (Day and Kincaid 1994). But there are two major differences: (1) they have a very different picture of the structure of IBE; (2) they have a different picture of how contexts figure into IBE. On (1). As I've explained IBE, every candidate hypothesis must be such that, if it were true, it would explain the data (i.e., it must be a 'potential explanation'). We then go on to evaluate the candidates based on their scores on the virtues. Day and Kincaid don't think about IBE this way: on their view, there are various reasons that you can offer in favor a view, of which being explanatory is only one. So, they see being explanatory as a reason that *competes* with the reasons generated by the explanatory virtues. On their view, the best explanation might *not* be recommended by our total evidence; on the view that I've sketched, such a claim makes no sense. On (2). According to the Context-Sensitive Schema, contexts affect the v -scores, but they do not affect the natures of the virtues – i.e., although simplicity in context c_1 may receive a different score than it does in c_2 , simplicity is nevertheless the *same property* in c_1 and c_2 . On Day and Kincaid's view, by contrast, the virtues themselves are only determinable properties; the virtues-in-contexts are determinate properties. So, for any two contexts c_1 and c_2 , although there is bound to be some family resemblance between simplicity-in- c_1 and simplicity-in- c_2 , they need not be identical properties. And at any rate, given the assumption that simplicity-in- c_1 and simplicity-in- c_2 are identical properties, Day and Kincaid never question whether they should have equal weights – i.e., that simplicity-in- c_1 and simplicity-in- c_2 should receive the same v -scores. Therefore, they are offering a version of the Very Simple Model (or one of its close cousins).

⁵⁶ In addition, different views on these matters will reflect different views about which recommendations IBE ought to make – more on this below.

explanatory virtue is truth-conducive: e.g., we can generate models of IBE that *never* allow the non-truth-conducive virtues to outweigh the truth-conducive ones.⁵⁷

Of course, the critics of IBE shouldn't worry just yet. The critics aren't talking about just any 'version' of IBE; they are talking about *IBE*, that variety of ampliative inference that people actually use. And, of course, any serious model for this form of inference should not trivialize the role of the non-truth-conducive virtues.

True enough. But that said, whatever the correct model of IBE, it is probably going to be either (a) more like the Very Simple Model or (b) more like some instance of either the Non-Additive Schema or the Context-Sensitive Schema. If (a), then it's plausible that IBE is not truth-conducive, and so cannot provide epistemic justification. If (b), though, then it is far from obvious that IBE is not truth-conducive. As you might guess, I think that IBE is more like some version of either the Non-Additive Schema or the Context-Sensitive Schema. Let me offer two cases that demonstrate why.

First, recall Einstein's famous resistance to quantum indeterminacy, often summed up with the slogan, "God does not play dice."⁵⁸ Einstein has often been criticized for being too conservative in this respect. However, he wasn't wrong that, at the time, deterministic models really did score better on conservatism and modesty than did indeterministic models: the deterministic models certainly fit better (and probably *still* fit better) with many of our antecedently held beliefs, and many people have found (and do find) indeterminism to be a bit

⁵⁷ Consider the *Trivially Truth-Conducive Non-Additive Model*. For all h and v , h 's v -score is equivalent to the number of units of v that h has. Now separate the v -scores for the truth-conducive and non-truth-conducive virtues. Let the sum of the v -scores for the truth-conducive virtues be called 'TruthSum'; let the sum of the v -scores for the non-truth-conducive virtues be called 'NonTruthSum'. Now, for all h and v , h 's overall score equals: $\text{TruthSum} + (\text{NonTruthSum} \times 0.000\dots 1)$.

⁵⁸ These are not actually Einstein's words. The original phrasing is to be found in a letter that Einstein wrote to Max Born in 1926: "Quantum mechanics is certainly imposing, but an inner voice tells me that it is not yet the real thing. The theory says a lot, but does not really bring us any closer to the secret of the 'old one'. I, at any rate, am convinced that *He* is not playing at dice" (Einstein, Born, and Born 1971, 91).

weird. Now, the Very Simple Model fits well with the view that Einstein reasoned correctly, but because IBE is untrustworthy, he was led astray. In other words, the proponent of the Very Simple Model is at risk of having to challenge the common assessment that Einstein did not reason correctly.⁵⁹ But an instance of the Context-Sensitive Schema accounts for our judgment quite naturally. When a basic assumption like determinism is being questioned, the importance of conservatism and modesty diminishes; these two virtues may not even deserve any weight at all. If the context is allowed to affect the *v*-scores, then we get the right result: Einstein did not pick the best explanation because he was too concerned to preserve determinism – i.e., he was too conservative.

Second, note that there may well be cases in which the weight of one of the non-truth-conducive virtues is *amplified*, not diminished. Conspiracy theories, for example, score extremely well on generality: they offer explanations of many more facts than do the theories we actually accept – e.g., why the man on Central Avenue was carrying an umbrella, why there was exactly \$4.73 in the tip jar at the coffee shop, why a limousine was circling the block, etc. Moreover, they come with a ready-made explanation for their lack of predictive power: if there *is* a conspiracy behind the relevant facts, the individuals involved would no doubt prefer that this not be known. So, since there is no expectation that the hypothesis should make confirmable predictions, that it doesn't is no evidence against it. Now, the whole point of a conspiracy theory is that things are not as we ordinarily take them to be. So, it begs the question against the conspiracy theory to rule it out by way of either conservatism or modesty. But simplicity remains, and conspiracy theories are almost always more complex than the hypothesis that we

⁵⁹ I am not saying that there is *no* way for the proponent of the Very Simple Model to avoid saying that Einstein reasoned correctly: you never *have* to say anything in philosophy. My claim is just that our judgment about Einstein fits better with an instance of the Context-Sensitive Schema than it does with the Very Simple Model. That much, I think, is fairly plausible.

take to be correct. Conspiracy theories posit more complex sequences of events, or layers of deception where the other hypotheses posit none, or what have you. Now, if we limit ourselves to the Very Simple Model (and its close cousins), then because of how well conspiracy theories score on generality, this greater complexity may not lower the overall scores enough so that they lose to more plausible hypotheses. In other words, if we limit ourselves to the Very Simple Model, then we may have to conclude that IBE supports conspiracy theories. But this is implausible: it's by looking to an explanatory consideration (simplicity) that we *rule out* conspiracy theories. An instance of the Context-Sensitive Schema can account for this: it just needs to specify that, when we are seriously engaging with conspiracy theories, simplicity receives a higher v -score. The odds are that an instance of the Non-Additive Schema can handle this too: it might have it that the relevant non-truth-conducive v -score gets double or triple its normal weight when the other non-truth-conducive v -scores are negligible (as they would be in these circumstances; recall the point about not begging the question against the conspiracy theory.) Again, it appears plausible that IBE is modeled more accurately by either an instance of the Context-Schema or the Non-Additive Schema.

What's the upshot? For starters, the critics of IBE have given us no argument to the effect that we should opt for the Very Simple Model of IBE (or one of its close cousins). In the absence of such an argument, they cannot undermine IBE's ability to provide epistemic justification just by pointing out that some of the virtues are not obviously truth-conducive. They may be right about some of the virtues, but that is not enough to secure their conclusion: they also need to defend a particular model of IBE (or at least a model schema). But more importantly, there is some reason to think that IBE is best modeled by an instance of either the Non-Additive Schema or the Context-Sensitive Schema. If that's right, then it is not as if the

Main Argument has a missing premise that can be easily supplied. Rather, the Main Argument is based on an implausible model of IBE, and so reaches its skeptical conclusion. Once the implausible model is exposed, Premise 1 is in bad shape. So, the Main Argument fails.

3.6 Particularism and IBE

Admittedly, I have not offered a model of IBE; I have only gestured toward the sort of model that we should expect. And this is no place for conceptual legislation: we should look to clear cases for guidance, and after seeing how things pan out there, we can frame hypotheses about how the virtues interact with one another and / or with various contexts. Similarly, I have not yet argued that IBE provides justification; I have simply argued that the Main Argument does not show otherwise. However, if we should develop a model of IBE by considering paradigm cases, then we can receive some inspiration for a story about IBE's epistemic credentials: we should look to paradigm cases to determine whether IBE provides justification.

Roderick Chisholm taught us to be *particularists*: i.e., we should suppose, first, that we can identify particular instances of knowledge without the aid of a criterion of knowledge and, second, that we can then justify a criterion by arguing from those instances (Chisholm 1973).⁶⁰ There are (at least) two related reasons to take this lesson to heart. The first is that we are likely to opt for an unduly skeptical criterion of knowledge if we do not begin with what we know; this is the lesson of Descartes' *Meditations*. The second is that, given any of the paradigm cases of knowledge, we're probably entitled to make the so-called 'G. E. Moore shift' if presented with an argument that undermines it.⁶¹ In other words, if presented with an argument like this:

1. We don't know that p unless we are in conditions $c_1 \dots c_n$.
2. We aren't in $c_1 \dots c_n$.

⁶⁰ For further considerations in favor of particularism, see (Lemos 2004).

⁶¹ The term is due to William Rowe; see (Rowe 1979), which is based on (Moore 1963, 222).

3. So, we don't know that p .

...we are probably entitled to turn it on its head:

1. We know that p .
2. We aren't in $c_1 \dots c_n$.
3. So, it's not the case that we don't know that p unless we are in conditions $c_1 \dots c_n$.

We are often more confident about cases than about principles. If so, then the G. E. Moore shift seems like a reasonable move to make. After all, if given a choice between two propositions, there would be little sense in believing the one of whose truth you are *less* confident.⁶²

In the present context, the upshot of Chisholm's particularism is that, if you want to establish a criterion of justification, then you should argue for it from a set of paradigm cases. So, our critics need to show that IBE either (a) fails to recommend those beliefs that we take be justified or (b) recommends beliefs that we *don't* take to be justified. In other words, our critics need to show that IBE fails to match our considered judgments about those beliefs that are and aren't justified. This will be an uphill battle, though, since they are likely to disagree with IBE's friends about what the right results are.

An example should make the point clear. Consider the old debate between constructive empiricists and scientific realists. If someone doubts that we are justified in believing that there

⁶² Someone might object that the G. E. Moore shift allows people to preserve beliefs that are clearly unreasonable. Suppose that Miss Cleo (a psychic) tells Tom that he will be lucky in love this week. Tom believes her and runs to tell his friends the good news. They are less excited than he had hoped. They suggest that his belief is reasonable only if (a) Miss Cleo offers falsifiable predictions and (b) those predictions are reliable. Since she fails on both counts, Tom should not hold out hope for a romantic encounter. Tom replies by making the G. E. Moore shift: he reasonably believes that Miss Cleo is right, but his belief does not meet his friends' conditions; so, those conditions are not necessary for reasonable belief. Now, I agree that Tom has made a mistake in this situation. However, it's not clear to me that the mistake is making the G. E. Moore shift. Suppose that it *really is* more plausible to Tom that Miss Cleo is correct than that his friends have hit on the right criteria for reasonable belief in her testimony. In such circumstances, what else is he supposed to do? Confidence is usually correlated with high subjective probability; so, by Tom's lights, it is more probable that Miss Cleo is on target. And if he has that belief, then surely he would be irrational to opt instead for claims that he judges to be less probable. Of course, Tom is wrong to assign such a high subjective probability to Miss Cleo's testimony. But I have yet to see a defense of the view that IBE arguments in metaphysics are on par with the testimony of self-proclaimed psychics, nor even that they are in the same neighborhood. Arguably, there is *no* context in which psychics (*qua* psychics) provide you with good reasons to believe that p ; in metaphysics, good reasons are not wholly absent, they are just fairly weak.

are atoms, then it will be easy to argue that the explanatory virtues are deficient: the belief that there are atoms is typically defended by an IBE argument, and if it is not included in the set of paradigms, then it will be appropriate to conclude that IBE gives the ‘wrong’ result in this case. However, the rest of us need not have qualms about IBE on these grounds. Indeed, we may take the stance that Quine called “unregenerate realism, the robust state of mind of the natural scientist who has never felt any qualms beyond the negotiable uncertainties internal to science” (Quine 1981b, 72). If so, then we will be perfectly happy to take our best science at face value, regarding as justified all sorts of claims about the visible and invisible constituents of the world, and taking such claims to be among the paradigms for which our theory of justification ought to account. And, since scientists often seem to offer IBE arguments in defense of these claims, we have little reason to worry that our data will fail to vindicate IBE.

If you find yourself in this dialectical situation, then the challenge amounts to the complaint that you are not adhering to your critic’s more exacting epistemic standards. But why should this bother you? This is *not* a standard case of disagreement in which (a) you and an interlocutor agree about the relevant epistemic standards, (b) you and your interlocutor have the same epistemic vantage point, and (c) you and your interlocutor disagree about p . In such a case you *do* seem to have a defeater for your belief that p (or $\sim p$, as the case may be). In the present context, however, (a) is not satisfied. Now, perhaps you have an epistemic obligation to ask for arguments in favor of the more exacting epistemic standards. But in the absence of such arguments, it’s hard to see why your critic’s skepticism should faze you. (I will consider one such argument in the next chapter.) From where you sit, you have every reason to maintain that IBE can provide justification.

This concludes my defense of IBE. I grant that there are other issues to explore: e.g., we

could ask for more detail about each virtue, or the nature of explanation, or the relationship between IBE and other forms of ampliative inference. Still, this chapter has shown why the most pressing worries about IBE are unfounded, and it has sketched an account of IBE's credentials on which it's quite plausible that we can justifiably believe theories based on IBE arguments. And at least in the present context, that's all that TT needs.

4. CONDITION C

4.1 Introduction

Thus far, I've set out TT and argued that we have some reason to think that we can satisfy TT's Condition A: you justifiably believe a theory according to which p is true. This chapter addresses a challenge to satisfying Condition C: you have no defeaters for the belief that p . Recall that there are two types of defeaters: rebutting and undercutting. A rebutting defeater for the belief that p is a reason to believe that $\sim p$; an undercutting defeater for the belief that p is a reason to think that the belief was formed improperly. I am interested in an (alleged) undercutting defeater: namely, that we are not appropriately related to the truthmakers for modal claims, and so can't have justified beliefs about them. I maintain that this challenge fails; our isolation from the truthmakers for modal claims does not constitute an undercutting defeater for our beliefs about them.

(N.B., in this chapter the distinction between knowledge and justified belief is more important. I'll ultimately suggest that it can be ignored, but until then I'll be careful not to slide between the two notions.)

4.2 Benacerraf's Dilemma

There is no consensus on the metaphysics of modality; there is no standard account of the truthmakers for modal truths. (There are many, of course, who think that the story should be told in terms of possible worlds. But the unity here is illusory, since there is no agreement about what the worlds are. Are they concrete entities, as Lewis argues? Are they Plantinga's maximal states of affairs? Stalnaker's maximal properties? Adams' sets of structured propositions?)

Still, we can say this: on most accounts, we bear no causal relations to the truthmakers for modal truths.⁶³ Let's assume that those in the majority have gotten this much right. If they have, then we seem to be in trouble: how can we know anything about entities from which we are causally isolated?

This problem is best known in the philosophy of mathematics. The contemporary discussion begins with Paul Benacerraf, we must either (a) make "it unintelligible how we can have any mathematical knowledge whatsoever" or (b) deny that mathematical claims are, strictly speaking, *true* (Benacerraf 1973, 662).⁶⁴ Very roughly, the argument for (b) is that, if we don't posit a domain of abstract objects (as the platonist does), then we won't be able to extend the standard (Tarskian) semantics to mathematical claims, and hence will have to revamp our understanding of truth in mathematics (e.g., by equating truth with provability). Since this is a move of last resort, platonism looks rather appealing. However, platonism comes with its own troubles.

In broad outline, the problem is this. Gettier cases show that knowledge is incompatible with (a certain sort of) luck. An attractive solution to the luck problem requires some kind of causal relation between the knower and the known. But the platonist postulates a realm of abstract objects, and – *ex hypothesi* – abstract objects are causally inert. So, platonism entails that we bear no causal relations to the truthmakers for mathematical claims. Therefore, if the platonist can't provide an alternate solution to the luck problem, then – according to Benacerraf – she makes it unintelligible how we can have any mathematical knowledge whatsoever.

This argument has a radical conclusion, so we should demand good arguments for its premises. We are not questioning whether mathematical objects are causally inert. So, the

⁶³ Most, not all. Armstrong's combinatorial theory is (purported to be) an exception. See (Armstrong 1989).

⁶⁴ Christopher Peacocke thinks of Benacerraf's Dilemma as a particular instance of the 'Integration Problem' – i.e., the problem of balancing cosmology and epistemology. See (Peacocke 1999). No doubt he's right about this.

argument stands or falls on the case that can be made for the thesis that there should be some kind of causal relation between the knower and the known. Call this ‘the Causal Requirement’ (CR). Benacerraf appears to offer two arguments for it.

The first argument tries to establish that CR is a conceptual truth. Benacerraf reasons this way. When challenging X’s claim to know that a proposition is true, we often assume CR: we try to show “that X’s four-dimensional space-time worm does not make the necessary (causal) contact with the grounds of the truth of the proposition for X to be in possession of evidence adequate to support [knowledge of the proposition]” (Benacerraf 1973, 671). We grant that Smith has a *justified true belief* that Jones owns a Ford or Brown is in Barcelona, but we deny that Smith *knows* that Jones owns a Ford or Brown is in Barcelona. Why not? Smith only has evidence for the claim that Jones owns a Ford, but since it turns out that Jones does not own a Ford, the truthmaker for the proposition is Brown’s being in Barcelona. But Smith didn’t recently see Brown in Barcelona, or receive a letter from Brown describing Brown’s adventures in Barcelona, or what have you. In Benacerraf’s terms, Smith did not make the necessary with contact with the grounds of the truth of the proposition. Benacerraf’s suggestion is that this provides some evidence that our concept of knowledge requires a causal connection between knower and known: presumably, we wouldn’t challenge knowledge claims this way if causal connections were irrelevant to whether the concept of knowledge applies to the case.

I don’t know whether CR forms part of our concept of knowledge (supposing that there is such a thing), but even if so, that concept surely isn’t sacrosanct; if need be, we can always replace it. And the platonist is within her rights to insist that we do just this. First, there are some well-known objections to causal theories of knowledge: e.g., they have a hard time dealing with our knowledge of general claims (‘All dogs are mammals’) and claims about the future.

But let's suppose that these sorts of hurdles can be overcome. Even if we have a workable theory of knowledge for concrete objects that includes CR, should we use it to cast doubt on our knowledge of abstract objects? Not obviously. It's practically undeniable that we have some mathematical knowledge, and if there are good reasons to believe that mathematical objects are abstract, then there are good reason to believe that we have some knowledge of abstract objects. So, we need to compare the merits of the argument for CR with the merits of the argument for taking mathematical objects to be abstract. Benacerraf and the platonist will judge this case differently, but it's far from clear that the platonist is irrational in the judgment she makes (especially if we bracket considerations about CR, the inclusion of which would only beg the question against the platonist). So Benacerraf's first argument amounts to the nominalist's report that he judges CR to be better supported than is the thesis that mathematical objects are abstract. Apart from worries about peer disagreement, this probably shouldn't unsettle the platonist, since she no doubt disagrees about the relative merits of the two theses. We have a draw. Therefore, Benacerraf's first argument for CR is no threat to platonism.

Let's turn to Benacerraf's second argument for CR. He says this:

It must be possible to establish an appropriate sort of connection between the truth conditions of p (as given by an adequate truth definition for the language in which p is expressed) and the grounds on which p is said to be known, at least for propositions that one must *come to know* – that are not innate. In the absence of this, no connection has been established between *having those grounds* and *believing a proposition which is true*. Having those grounds cannot be fitted into an explanation of *knowing p* . The link between p and justifying a belief in p *on those grounds* cannot be made. But for that knowledge which is properly regarded as some form of justified true belief, then the link *must* be made (Benacerraf 1973, 672-673, emphasis in original).

I don't know why Benacerraf exempts innate knowledge from CR, but let's set this issue aside, since it doesn't appear to be pertinent to the mathematical claims in question. (After all, it's implausible that set theory is innate.) The interesting thing to note is that there is a shift here

from talk about concepts to talk about explanation. According to this argument, we ought to be able to explain why our reasons lead to epistemic successes – i.e., having true beliefs. The allure of CR is that it allows to offer the relevant explanations: our reasons to believe p lead us to truth because those reasons are somehow prompted by p 's subject matter. The platonist has no analogous resources to which she can appeal.

We can formulate the challenge as follows:

1. If you know that p , then p 's truthmaker must figure appropriately in an explanation of your belief that p .
2. But if p is a claim of pure mathematics, then p 's truthmaker cannot figure appropriately in a *causal* explanation of your belief that p (because, *ex hypothesi*, p 's truthmaker is abstract, and so causally inert).
3. Where p is a claim of pure mathematics, there is no other plausible way that p 's truthmaker might explain your belief that p .
4. So, if p is a claim of pure mathematics, then you do not know that p .

If we assume that this argument is wholly distinct from the prior (conceptual analysis) argument, then it isn't obvious why Benacerraf thinks that we should accept the first premise. But if nothing else, we can say this. Even if the principle of sufficient reason is false, the default position should be that all facts are explicable. If persistent attempts to explain some fact are unsuccessful, and if the lack of success is not plausibly attributed to a shortcoming on our part, then we may conclude that the fact in question is brute. However, brute facts are like theoretical primitives: you can always postulate more of them, but unless you can make a very strong case for their necessity, you have every reason to keep them to a minimum. In this sense, taking a fact to be brute is a move of last resort. At any rate, if the first premise is true, then the argument looks fairly strong. If p is a claim of pure mathematics, and if p 's truthmaker is indeed causally inert, then it's very hard to see how p 's truthmaker could figure appropriately in a causal explanation of your belief that p . And once we set aside the possibility of providing a causal

explanation, what other sort of explanation is available? No answer to this question seems to be forthcoming.

As dire as the situation may appear, there is – at first blush – an easy way out. Benacerraf appears to concede that the platonist can *justifiably believe* the claims of pure mathematics; what he seems to deny is that she can *know* those claims. That is, Benacerraf appears to concede that the platonist can justifiably believe the claims of pure mathematics even if she does not posit a link between our reasons to believe that p and p 's truth conditions. And justified belief may be enough for the platonist; perhaps she can live without knowledge. After all, if we give up on platonism, then we have to deny that mathematical claims are really true (or so Benacerraf argues, and I tend to think). By all accounts, though, you cannot know what is not true. So, whether or not we give up on platonism, it looks as if we must deny that we know mathematical claims. But if we all have to be satisfied with something less than knowledge, then it is not clear that the platonist is in a worse position than her interlocutor.

This reply is fine as far as it goes, but it trades on the particulars of the objection and does not address the deeper complaint. Ultimately, I think that Benacerraf wants to know why our reasons for believing p – frail, thin reeds that they are – should be guides to p 's truth. (Perhaps this explains why, in the passage quoted above, he waffles between talk about knowledge and talk about justification. Right after saying that the link is needed to explain how it is that we could *know* that p , he complains that, if we are platonists, “[t]he link between p and *justifying* a belief in p on those grounds cannot be made”⁶⁵ – suggesting that the causal requirement is not just a patch on the justified-true-belief account of knowledge.) John Burgess and Gideon Rosen suggest that Benacerraf’s “question [...] is essentially just a demand for a philosophical ‘foundation’ for common sense and science – one that would show it [i.e., common sense and

⁶⁵ I’ve switched the emphasis from “on those grounds” to “justifying.”

science] to be something more than just a convenient way for creatures with capacities like ours to organize their experience” (Burgess and Rosen 1997, 48). I think that this is basically right. After all, Benacerraf never questions whether mathematical claims are justified by the standards that mathematicians employ: he never doubts that, by some standard internal to mathematical practice, Euclid offered a legitimate proof that there are infinitely many primes. Rather, the demand is for an explanation of the (alleged) *reliability* of this practice – i.e., we want to know why it is not merely a happy accident that our reasons lead us to truth.⁶⁶ To be clear, the request is for an explanation of reliability on the *presumption* that our theories are true. This is not the quixotic insistence on certainty, but the more reasonable demand that our epistemology should mesh with our cosmology. In most domains of inquiry, our causal connections with the world (partially) explain the reliability of our theorizing. If we have some causal commerce with p ’s truth conditions, then we can say that our reasons to believe p are somehow shaped by our interaction(s) with p ’s truth conditions, and hence do not lead us wildly astray. The platonist rejects this line, and so one wonders whether there is anything that she can offer in its stead. Benacerraf, of course, is doubtful. We can think of this as a perfectly general undercutting defeater for beliefs about mathematics, not just telling against claims to knowledge, but also against claims to have justification. So, while I’ll continue to discuss CR as a threat to mathematical (and modal) knowledge, I think that the worry applies, *mutatis mutandis*, to having justified beliefs of mathematics (and modality).

⁶⁶ If this is the right this is the correct reading of Benacerraf – and I think that it is – then “Mathematical Truth” already contains the heart of the argument that is usually attributed to (Field 1980).

4.3 Replies Not Pursued

Let's return to the epistemology of modality. In this context, Benacerraf's challenge is that we should be able to explain the reliability of the means by which we acquire modal beliefs. If we are realists about modality, there are only two moves available here. On the one hand, we can try to offer the explanation that Benacerraf demands. On the other, we can reject his demand, offering instead some story about why the requirement is inappropriate.

Many would like to offer the explanation. A fashionable candidate appeals to evolutionary considerations. Very roughly, the story goes like this. *If it's true that bears can hide in caves, then it's good for your health to believe as much: if you have this belief, then you can be duly cautious when trying to decide where to seek shelter. So, at least in some cases, it's fitness-enhancing to have true beliefs about what's possible. But evolutionary forces select belief-producing mechanisms, not individual beliefs. Hence, evolutionary forces probably selected a belief-producing mechanism that is generally reliable about what's possible.*

I am skeptical of this reasoning. Evolutionary forces (a) are constrained by time and the available raw materials and (b) are not concerned, first and foremost, with producing true beliefs (even if true beliefs they produce). These points lead to three serious problems. First, given that evolutionary forces will be happy to trade truth for another benefit if the opportunity presents itself, we need a reason to think that they've operated long enough in conditions favorable to selecting reliable mental mechanisms. (Granted, we have reason to suppose that our mental mechanisms are superior to those had by our competitors, but for all we know, that's a backhanded compliment.) What argument can we offer for this view? Second, given that evolutionary forces are not directed toward truth, it would appear to be an *accident* that our faculties are reliable (if reliable they are). But if it is an accident that our faculties are reliable,

we should ask why *this* accident is an epistemic improvement over the seemingly-accidental correlation with which we began – namely, the correlation between modal matters and our beliefs about modal matters. What answer can we offer? Third and finally, even if evolutionary forces have been operating in the right conditions for a sufficient amount of time, and even if our reliability is not an accident (or at least not an epistemically-problematic accident), there is still a difficulty about the *scope* of the mechanisms that evolutionary forces are likely to select. The initial argument gives some reason to think that these mechanisms are reliable when it comes to matters of survival. But are they reliable when it comes to matters removed from the business of everyday life? Are they likely to produce true beliefs about highly theoretical issues? It is one thing to suggest that evolutionary forces are likely to produce a mechanism that tracks the truth about mundane matters; it is quite another to suggest that they will produce a mechanism that rarely veers into falsehood when faced with questions outside its original purview. Indeed, I would expect that a generally *unreliable* system will be reliable if it's corrected often enough, as the human cognitive system is. Where matters of survival are concerned, we receive an enormous amount of input from the world; we are constantly assaulted with information relevant to our continued existence. Hence, evolutionary forces can afford to produce a cognitive system that cuts corners, since they can bank on the world to take up the slack. But where the world offers less in the way of correction – and why shouldn't modal matters be a prime example? – we have much less reason for optimism about the reliability of our faculties.⁶⁷

The evolutionary explanation does not appear to be very promising. If we continue to accept the demand for explanation, what other options are there? As far as I can see, only one:

⁶⁷ For more objections to this kind of evolutionary argument, see (Stich 1990) and Plantinga's contributions to (Plantinga and Tooley 2008).

you can try to articulate a view on which we are causally related to modal facts.⁶⁸ Robert Koons goes this route (Koons 2000). He maintains that there are primitive modal facts – neither reducible to nor supervenient on the nonmodal facts – and that we glean information about them by interacting causally with them. Timothy O’Connor also thinks that we are causally related to modal facts, but he gets this result via a curious kind of occasionalism (O’Connor 2008). O’Connor says that God’s power is the truthmaker for modal claims. He takes God’s modal knowledge to be self-knowledge (i.e., God’s knowledge of the scope of God’s power), and then he argues that our faculties are generally reliable about modal matters because God designed them to have this feature. It would take me too far afield to engage seriously with these proposals, so I’ll just say this: Koons and O’Connor may have landed on viable replies to Benacerraf’s challenge, but most of us will have a hard time suppressing incredulous stares. That said, if we rule out their replies, then the realist is going to struggle to meet the terms of Benacerraf’s challenge. So, it looks as though the realist needs to reject Benacerraf’s demand for an explanation of our reliability.

4.4 Rejecting the Demand for Explanation

David Lewis leads the way here. Recall the initial set up. Gettier cases show that knowledge is incompatible with (a certain sort of) luck. An attractive solution to the luck

⁶⁸ If we steal another idea from the literature on the epistemology of mathematics, we could try to argue that we can perceive possibilities without being causally related to them. This is the line taken by the early Penelope Maddy on our knowledge of sets (Maddy 1990, 50-67). In short, she maintains that we rely on theories to fix the contents of our perceptions; so, if our best theories commit us to sets, and if we believe that sets are located where their members are, then we can perceive them. (To my mind, if Maddy’s story makes sense, it’s because she tacitly assumes that nearly *all* perceptible objects are abstract. Given this assumption, we’ve got to say that light can bounce off sets; if we don’t, then we’ll have to deny that it can bounce off your cup of coffee.) I don’t know whether her account works for sets, but it faces a major hurdle if told about possibilities. There is an interesting case to be made for the view that sets are where their members are. This matters because, with the exception of a few wild-eyed Platonist (and a few theists; see (Alston 1991)), everyone thinks that perceptible objects must be located. But where is the fact that I could have had tea rather than coffee with breakfast this morning? I doubt that there is any good answer to this question.

problem requires some kind of causal relation between the knower and the known. Platonism ensures that mathematical claims are true, but it does this by postulating a realm of abstract objects, all of which are – *ex hypothesi* – causally inert. So, platonism entails that we bear no causal relations to the truthmakers for mathematical claims. Therefore, if the platonist can't provide an alternate solution to the luck problem, then – according to Benacerraf – she makes it unintelligible how we can have any mathematical knowledge whatsoever.

This argument obviously applies, *mutatis mutandis*, to modal realism (not because Lewis takes possibilia to be causally inert, but because he thinks that we are causally isolated from them). Lewis is well aware of the challenge, and he begins by replying to the mathematical version. First, he claims that – at least at present – platonism is the only viable account of mathematical truth; on his view, nominalist interpretations of mathematics *reform* mathematics, and philosophers have no right to challenge the deliverances of mathematicians. Second, he contends that we are more sure of our mathematical beliefs than we are of the epistemology that is supposed to cast doubt on them. (Presumably, he also thinks that we are more sure of platonism than we are of CR. If we weren't, then we might hang on to our modal knowledge and CR, admitting that – at least for the time being – we have no plausible account of mathematical truth.) Given all this, CR should be axed. And if it is, then we can extend this reasoning to the modal case:

So mathematics will do as a precedent: if we are prepared to expand our existential beliefs for the sake of theoretical unity, and if thereby we come to believe the truth, then we attain knowledge. In this way, we can even attain knowledge like that of the mathematicians: we can know that there exist countless objects causally isolated from us and unavailable to our inspection. Causal accounts of knowledge are all very well in their place, but if they are put forward as general theories, then mathematics refutes them (Lewis 1986b, 109-110).⁶⁹

⁶⁹ Three comments. First, Hawthorne doesn't question the assumption that our mathematics is justified by its theoretical utility, so I won't either. Second, there is no obvious reason for Lewis to regard theoretical unity as the only explanatory virtue for which it is worth expanding our beliefs; we might also expand our beliefs for the sake of,

Thus far in the reply, Lewis hasn't explained *why* it's wrong to use CR against claims to mathematical and modal knowledge; he's just argued that we shouldn't, since doing so would require us to give up other more plausible commitments. But he does have an explanation for the wrongness, and this is the heart of his reply to Benacerraf. Lewis has no truck with CR as a requirement on knowledge of contingent truths, as it serves to eliminate (the pernicious variety of) luck. However, where necessary truths are concerned, he denies that our beliefs can be accidentally true in the way that worried Gettier, and hence denies that our epistemic successes require the same sort of explanation:

If I know by seeing, for instance, my visual experience depends on the scene before my eyes; if the scene had been different, within limits, my experience and my subsequent belief would have been correspondingly different. [...] But nothing can depend counterfactually on what mathematical objects there are, or on what possibilities there are. Nothing sensible can be said about how our opinions would be different if there were no number seventeen, or if there were no possibility for dragons and unicorns to coexist in a single world (Lewis 1986b, 111).

In other words, because all counterpossibles are vacuously true, it makes no sense for an epistemology to impose conditions like these:

- You don't know that p is possible unless, if p weren't possible, you would believe that p is impossible.
- You don't know that p is possible unless, if p weren't possible, your evidence for p 's possibility would be other than it (actually) is.

If all counterpossibles are vacuously true, then these conditions are trivial; you can't avoid satisfying them. So, we have an explanation of why it's wrong to impose CR on our knowledge of necessary truths: such a requirement only makes sense where our beliefs non-trivially depend

say, generality. I don't think that Lewis would disagree, although given his target here, he emphasizes one virtue over the others. Third, there is no obvious reason for Lewis to limit himself to *existential* beliefs, at least if that notion is construed so that it only applies to the existence of individuals. On the face of it, there is no barrier to using the same method to justify beliefs about properties, modes, structure, relations, etc. Adjusting for these caveats, he might say: if we expand our stock of beliefs to secure greater explanatory power, and if we thereby come to believe the truth, then we attain knowledge.

on their truthmakers, and this is not the case where our beliefs concern necessary truths. Lewis concludes, therefore, that we can have both mathematical and modal knowledge.

4.5 Hawthorne's Objection

I don't want to work through the thorny issues raised by counterpossibles here, so let's grant Lewis his stance.⁷⁰ Still, John Hawthorne balks at Lewis' solution, maintaining that Lewis' proposal would leave us without the full-blooded justification that's needed for knowledge. He tries to retool the original luck problem so that it can be stated without appealing to counterpossibles. Here is his first pass:

Even if serviceability⁷¹ confers some degree of justification upon a theory and even if that theory is true, it just does not seem that this is sufficient for knowledge. Suppose, for example, that theory A is a little less elegant than B, or a little harder for us to derive predictions from than B due to the computational idiosyncrasies of human beings. Supposing one adopted B, it just does not seem that one now knows that B is true. We know from Gettier, after all, that justification plus truth does not suffice for knowledge. Where the serviceability of a modal theory is not connected in the appropriate way to the truth of a modal theory – and in particular where the serviceability is not explained by modal reality – we seem have the sort of gap that Gettier alerted us to (O'Leary-Hawthorne 1996, 190-191).

This first formulation is unsuccessful. The explanandum here is an event – namely, someone's believing a modal claim based on its serviceability.⁷² But Lewis thinks that we explain an event by providing some information about its causal history (Lewis 1986a). So, either (a) Hawthorne is reiterating a causal requirement on knowledge, in which case Lewis has a ready reply, or (b) Hawthorne owes us a non-causal account of explanation – and more generally, an account of explanation that makes no appeal to counterfactuals, since any species of counterfactual

⁷⁰ For Lewis' defense of his view about counterpossibles, see (Lewis 1973, 21-26); for arguments in the same vein, see (Williamson 2007), particularly Chapter 5. For counterarguments, see (Nolan 1997) and (Yagisawa 1988, 2010).

⁷¹ Hawthorne means this term to be a synonym for Lewis' 'theoretical utility'.

⁷² ...or perhaps a hideously complicated event, constituted by every instance of someone's believing a modal claim based on its serviceability. The difference won't affect the point in the main text.

dependence will lead to the problem that Lewis details. I'm going to assume here that Hawthorne does not intend that his objection depend on a particular theory of explanation; so, we can set this version aside.

Shortly after the above passage, Hawthorne offers a better formulation of the objection. Where mathematics and modality are concerned, Lewis is right that the *facts* can't be different, but he overlooks the simple point that *knowers* can be:

Suppose that there are two tribes with different computational architectures that render different theories more manageable to the respective tribes. One modal theory is more manageable for A, another more manageable for B. A is true. Is one now tempted to say that one tribe knows that A is true? Only someone in the grip of a theory would even feel the temptation, let alone succumb to it (O'Leary-Hawthorne 1996, 191).

The luck described here does not appear to be exactly what we see in Gettier cases, but I don't see a straightforward way to characterize the difference. However, this version of the luck problem clearly avoids counterpossibles, so it can't be dodged using the strategy above. Therefore, if this variety is incompatible with knowledge, then Lewis is in trouble. And I think that many will be inclined to think that it *is* incompatible with knowledge.

4.6 Handling Hawthorne's Objection

What can we say in Lewis' defense? To begin, let's concede that the A-tribers are luckier than the B-tribers: it just so happens that, through no merit of their own, the A-tribers have computational architectures that are better suited to discovering modal truth. To say this is to highlight the fit between the A-tribers and a certain aspect of their environment. But is this convenient fit incompatible with knowledge? Not obviously. We too are lucky that the mechanisms behind our beliefs are operating in a congenial environment. The possibility of global deception suffices to prove this point, but we needn't appeal to anything so dramatic.

Instead, we can look back at many of our ancestors and note that their scientific beliefs were mostly false. We are lucky that we are not in their unfortunate position, where the theory that was most manageable was one that postulated tree spirits and water gods, or the four classical elements. But it would be perverse to let this luck lead us to doubt that we have any scientific knowledge.⁷³ Not all luck is epistemically problematic, and it is far from clear that Hawthorne has identified the relevant variety.

What's more, we can explain the initial plausibility of the two-tribes case. To believe that p is to take a stand on p 's truth, and whatever the merits of your reasons for believing p , there is always some chance – however slight – that p may be false.⁷⁴ Let's say that our threshold of epistemic risk is the point at which we are unwilling to venture belief based on the evidence; it is the point at which we judge that the evidence is insufficient to risk error. If you have a very low threshold of epistemic risk, then you lean toward skepticism: you tend to think that the evidence rarely lowers the risk of error sufficiently. If you have a very high threshold, then you are willing to be a sucker: you tend to think that it is usually worth risking error for the sake of gaining truth. Now let's suppose – rather plausibly – that our epistemic desiderata are not equally reliable. Specifically, let's suppose that predictive power is a very reliable indicator of truth while theoretical utility is not. So, our epistemic risk is lower when we limit ourselves to beliefs that have great predictive power, though the potential scope of our knowledge is correspondingly limited. Our epistemic risk is higher when we allow ourselves to believe based on theoretical utility; but in exchange for this increased risk, we gain access to a greater range of

⁷³ Or, at the very least, it would be perverse for a scientific realist to let this luck lead her to doubt that we have any scientific knowledge. After all, proponents of the pessimistic meta-induction will surely be glad to judge this sort of luck to be incompatible with knowledge.

⁷⁴ I don't mind the implication that we have no incorrigible beliefs, but if you do, then just suppose that those beliefs are being bracketed – the difference won't matter here, since none of the beliefs in question has any claim to being incorrigible.

truths. The two-tribes case brings into focus the risk that we take when we believe based on theoretical utility, and it's effective because it's natural to recoil at the prospect of error: we are not often forced to take a good look at the contingency of our epistemic successes. However, the case does *not* show that the A-tribers don't know the modal theory. All the case shows is that risk-taking doesn't always pay off: things worked out for the A-tribers, though not for the B-tribers. But we already knew that. For the case to show that the A-tribers don't know the modal theory, there would have to be some level of epistemic risk that's incompatible with knowledge. Let's grant as much. Even so, we still need an argument to show that the A-tribers have crossed that line.

At this juncture, it may be tempting to think that there is a way to retool CR. Someone might contend that (*ceteris paribus*) causal interaction with p 's subject matter lowers the risk of being wrong about p 's truth value, while there is much greater risk of being wrong about p 's truth value if p subject matter is causally inert. Presumably, if this relationship between risk and causality holds, it has something to do with the fact that causal relations provide a mechanism by which information can be transmitted from known to knower. And I presume that *that* matters because we want it to be the case that we are responsive to the facts – i.e., that we would not believe p in just any circumstances whatever, but only when p is (at least approximately) true. We are now dangerously close to reinstating the *verboten* counterfactual requirement on knowledge, but I won't pursue this worry. Instead, I simply want to point out that Lewis' view on counterpossibles is not the only consideration that tells against demanding responsiveness to the modal facts.

I know that I am not being deceived by an evil demon, but I would believe this even if it were false. I know that the world did not pop into existence five minutes ago, complete with all

the details that make it *look* billions of years old; and again, I would believe this even if it were false. I know that I have not been bombarded with some wacky form of radiation in virtue of which I universally mistake red for green, but also universally use ‘red’ to mean green and ‘green’ to mean red. And I would believe this even if it were false. In all these cases, the mechanisms that produce my belief are not responsive to the relevant facts. The moral here is that if we know *any* skeptical hypothesis to be false, then we can have knowledge even where our belief-producing mechanisms are in no way responsive to the facts. So, if we condemn our modal (and mathematical, and perhaps moral) knowledge on the grounds that the mechanisms that produce these beliefs are not responsive to the facts, then we will have to make the same pronouncement about our knowledge of the falsity of various skeptical hypotheses. But since we shouldn’t condemn the latter, we shouldn’t condemn the former.

(N.B., someone can always deny that we know that various skeptical hypotheses are false. I think that this is a quixotic move. Remember that philosophy is a plausibility game: if an honest soul has to choose between (a) saying that our belief-producing mechanisms must be responsive to the facts if we are to have knowledge and (b) saying that we know that we are not brains in vats, it’s all but certain that she’ll choose the latter. And if she doesn’t – by some strange quirk of her psychology – then surely we are allowed to say that she ‘is in the grip of a theory’.⁷⁵)

⁷⁵ I should make a methodological point here. It’s unfair to the skeptic to demand that her epistemology not lead to skepticism: she’s entitled to the view that she finds most plausible. It may even be unfair to the skeptic to demand that her epistemology not *necessarily* lead to skepticism: perhaps the only viable epistemology, in her eyes, is one that entails that knowledge is scarce. What we *can* demand of the skeptic is that she not try to convict us of irrationality by way of premises that we reasonably reject. So, the complaint against Benacerraf is not that he’s jiggered his epistemology so that skepticism is the only possible result: he’s within his rights to do this (if that is indeed what he’s done). The complaint, rather, is that skepticism can’t be foisted on us if we reasonably reject the machinery that would take us there. And thus far, Benacerraf has provided no argument to the effect that, if the machinery is rejected, it is rejected unreasonably. In other words, we’ve been given no reason to accept his preferred threshold for epistemic risk.

4.7 Taking Stock

So Hawthorne has not shown that the luck in the two-tribes case is incompatible with knowledge. Moreover, we can explain the allure of the two-tribes case with some considerations about epistemic risk. And finally, it is hard to see how to show that the risks taken by the A-tribers preclude them from having knowledge.

All that said, I concede that the two-tribes case may lead us to question a risk threshold that permits the use of theoretical utility as an epistemic desiderata. But if it does, the right question is not, ‘What level of epistemic risk is compatible with knowledge?’, but rather ‘Are we getting our money’s worth for our current level of epistemic risk?’ After all, the appropriateness of risk is relative to the expected net return. We adopt quite low thresholds when the stakes are high; in contexts where we have a lot to lose if we err, we demand more evidence. Likewise, we quite reasonably adopt high thresholds when the cost of error is insignificant; there is no reason *not* to say, ‘Nothing ventured, nothing gained’, when what’s gambled is of little worth. So, what we ought to ask ourselves is, ‘Does our modal theory indeed have great theoretical utility?’ and ‘What do we gain by believing it?’ and ‘What do we stand to lose if we’re wrong?’ The answers to these questions should guide our risk threshold, and not *a priori* considerations about that with which knowledge is compatible.

Plainly, Lewis will answer these questions differently than will Hawthorne. Accordingly, Lewis will be at peace with a higher threshold of epistemic risk, and he will be well within his rights to deny that Hawthorne’s objection undermines his case for modal realism. If pressed to choose between the two perspectives, I side with Lewis. However, it is *not* my contention that it is irrational or unreasonable to be moved by Benacerraf’s challenge. Thresholds of epistemic risk – like thresholds of other kinds of risk – vary from person to person, and I suspect that there

is a range of acceptable ones. Wayne Riggs offers an example that lends some support to this view:

Consider two children, Sam and Pat, approaching a local park that is crowded with the other neighborhood children. To get to the park, they must get across a small creek. There is a bridge half a block away, but there is also a place where the more daring older kids just jump across right where the two children are. Neither Sam nor Pat has ever jumped the creek before, judging it to be at the very limit of their capabilities to make it across. However, this time, Sam runs towards the creek and takes the daring leap, landing successfully on the other side amidst the cheers of the other children. Pat, on the other hand, considers it for a moment, and then walks over to the bridge to cross (Riggs 2008, 2).

The creek is neither an inch nor a mile wide, and the benefit of clearing it is neither negligible nor remarkable; so, it is not as if there is an obvious way to argue that one of the boys made the better decision. And in this case, we feel no pressure to pass judgment one way or the other: we can say that neither Sam nor Pat acts ‘rightly’; rather, they both act reasonably given their respective thresholds of (physical) risk. But why should the epistemic case be any different? Why should we insist that, where belief is concerned, there is no analogous flexibility? I see no plausible answer to this question.

Of course, it is perfectly legitimate to dispute that a *given* threshold falls outside the bounds of rationality: this, essentially, is the judgment that we pass on skeptics. But on what grounds might we pass this judgment on the platonist? The complaint, of course, is that where we cannot explain our reliability, we cannot justifiably believe. But in the present context, this boils down to the assertion that we ought to adopt a low threshold of epistemic risk – not an argument for doing so.

4.8 Conclusion

Let me sum up. As I read Benacerraf, he wants a foundation for knowledge; he wants a story that explains why we get things right when we do. And in general, it's reasonable to demand such an explanation. However, on closer inspection, it seems unlikely that we can meet this demand without giving up some very plausible anti-skeptical knowledge claims. If we insist that we do not have knowledge where we aren't responsive to the facts, then we will have to concede that we do not know that we aren't brains-in-vats.

I'd guess that Benacerraf and Hawthorne have relatively low thresholds of epistemic risk. But it isn't clear to me that this is a question of right or wrong. Risk thresholds are relative to values, and I suspect that – within rough limits – there is no fact of the matter about how much you ought to value truth or understanding or the avoidance of error. Hence, I find it plausible that there is a range of rational thresholds for human beings to employ. If we can accept a greater level of epistemic risk, then Benacerraf's challenge does not provide a defeater for our beliefs about matters from which we are causally isolated.

5. MODAL SKEPTICISM

5.1 Introduction

I've spent the last few chapters defending TT against objections. By and large, the challenge has been to show that TT is compatible with various claims to knowledge. But what does TT suggest about our epistemic limits – current or absolute? Very roughly, TT says that, where p is any modal claim, we could be justified in believing p if and only if we are justified in believing a theory that says that p is true. It follows that the *current* potential scope of our modal knowledge is determined by our current theories: if they pronounce on p 's modal status, then we could have justified belief about it; if not, then we couldn't.⁷⁶ The *absolute* potential scope of our modal knowledge is determined by those theories that we could come to justifiably believe: if we could come to justifiably believe a theory that pronounces on p 's modal status, then we could justifiably believe that p has the relevant status; if not, then we couldn't.

In this chapter, I argue that these features of TT make it a version of Peter van Inwagen's 'modal skepticism'; what's more, I argue that TT is a more compelling version of the view than the one that van Inwagen develops. What is van Inwagen's modal skepticism? Let's say that a proposition is 'extraordinary' if it meets these conditions: (1) it concerns matters "remote from the practical business of everyday life," (2) its truth-value "cannot be determined by logic and reflection on the meanings of words or by the application of mathematical reasoning," and (3) its truth-value is "unknown to us or [is] known to be false" (van Inwagen 1998, 74, 84). Modal skepticism is the view that we do not know the modal status of any extraordinary proposition.⁷⁷

⁷⁶ I am suppressing concerns about conflicts between theories and undercutting defeaters (like Benacerraf's Dilemma).

⁷⁷ Van Inwagen admits: "This name was perhaps ill-chosen, since, as I have said, I think that we do know a lot of modal propositions, and in these post-Cartesian days, 'skeptical' suggests someone who contends that we know

Van Inwagen thinks that this view rules out knowing the modal statuses of propositions like these:

- There is a being that has all perfections essentially (where necessary existence is taken to be a perfection).
- I exist and nothing material exists.
- There exists vast amounts of suffering for which there is no explanation (van Inwagen 1998, 67-68).⁷⁸

So if this version of modal skepticism is true – and if van Inwagen is right to think that these propositions satisfy the conditions above – then we don’t know (for example) either that it’s possible that I exist and nothing material exists or that it’s *impossible* that I exist and nothing material exists; we are simply ignorant about the modal status of this proposition. (To be clear, I don’t think that anything turns on whether van Inwagen is right about these particular propositions; the conditions are what matter.) So modal skepticism is *not* the view that we do not have any modal knowledge whatever; it is perfectly compatible with our having modal knowledge about matters (a) that are *not* removed from the business of everyday life, (b) that *can* be determined by logic and reflection on the meanings of words or by the application of mathematical reasoning, or (c) that are known on the basis of their actual truth (i.e., inferences from actuality to possibility). For van Inwagen, anyway, modal skepticism is really modal *modesty* – and in particular, modesty about many of the modal claims that professional philosophers are happy to deploy in arguments.

I do not intend to defend modal skepticism generally; at the beginning of Chapter 1, I

nothing or almost nothing. It should be remembered, however, that there has been another sort of skeptic: someone who contends that the world contains a great deal of institutionalized pretense to knowledge of remote matters concerning which knowledge is in fact not possible. (Montaigne was a skeptic in this sense, as were, perhaps, Sextus and Cicero.) It is in this sense of the word that I am a modal ‘skeptic’” (van Inwagen 1998, 69).

⁷⁸ But note that van Inwagen does *not* think that a proposition is extraordinary just in case it is philosophically contentious. There are, for example, those who deny that all identity claims are necessary if true, but van Inwagen is open to the view that we know the detractors to be in error (van Inwagen, 1998, p. 82, n. 11).

said everything that I have to say in favor of epistemic humility regarding modal matters. Nor do I intend to defend van Inwagen's version of modal skepticism: his conditions for being an extraordinary proposition point in the right direction, but they are simultaneously too permissive and too restrictive (more on this later). My claim here is conditional: *if* we are to be modal skeptics, then we should opt for my version over van Inwagen's. So here is the plan of this chapter. First, I set out van Inwagen's case for modal skepticism. Most of its elements are present in (van Inwagen 1998), but he has long been advocating for a more cautious theory of modal knowledge, so I will be drawing from a number of his earlier papers too. Then, I consider some problems with van Inwagen's argument. After offering the best solution I can, I contend that van Inwagen's variety of modal skepticism rests on an unfortunate analogy between modal and perceptual knowledge. Finally, I show how TT avoids these problems and makes for a more attractive version of the view.

5.2 Van Inwagen's Argument

I understand van Inwagen's overarching argument as follows.⁷⁹ If you believe that p is possible, then you are committed to there being a possible world at which p is true. If there is such a world, then there are various historical and nomic facts with which p 's truth is compossible. So, if you believe that p is possible, you are committed to there being a history and set of laws with which p 's truth is compossible. Suppose you recognize this (as you should, if you are a philosopher). Then, you justifiably believe that p is possible only if you justifiably believe that there is a history and set of laws with which p 's truth is compossible. Now let's suppose that p is an extraordinary proposition (i.e., one (1) that concerns matters "remote from

⁷⁹ As I said, I'm extracting this line of reasoning from (van Inwagen 1998). However, he defends a similar line in a number of other places. See, e.g.: (van Inwagen 1977, 1979, 1997, 2006).

the practical business of everyday life,” (2) whose truth-value “cannot be determined by logic and reflection on the meanings of words or by the application of mathematical reasoning,” and (3) whose truth-value is “unknown to us or [is] known to be false” (van Inwagen 1998, 74, 84)). You should be agnostic about whether there is a history and set of laws with which p ’s truth is compossible. Hence, you should be agnostic about p ’s possibility.

Let’s say that there is a ‘ p -friendly world’ just in case there is a history and set of laws with which p ’s truth is compossible. The key claim in this argument is that you should be agnostic about whether there is a p -friendly world. Why believe it?

There are a number of pieces to van Inwagen’s answer, but the most important one is that possible worlds are *coherent* entities – their parts ‘hang together’ in some important sense of that phrase. He puts this as follows:

To assert that p is possible [...] is to commit oneself, willy-nilly, to the thesis that there is a whole, coherent reality – a possible *world* – in which p is true, of which the truth of p is an integral part. To assert that it is possible for the moon (or a thing in the moon’s actual orbit that looks like the actual moon when observed from the surface of the earth) to be made of green cheese is to commit oneself, willy-nilly, to the thesis that a physical universe in which a moon-like thing made of cheese came into existence and continues to exist is possible, that there are possible laws of nature and possible initial conditions that permit such a thing. (Or, if the object is supposed not to have arisen in the natural course of events, but to be literally miraculous, to the thesis that there could be a supernatural agency that was capable of creating and sustaining it and which either had a good reason to create and sustain a huge ball of cheese in orbit around the earth, or which might create and sustain such a thing without having any good reason to do so) (van Inwagen 1998, 77-78).⁸⁰

⁸⁰ I don’t think that van Inwagen regards all this as contentious, even though the passage may commit him to denying various Humean theses – e.g., that the laws are just the axioms of the true deductive system that best balances simplicity and strength; that there are no necessary connections between distinct existences; that (some version of) the principle of sufficient reason is false. But since (a) I agree with everything that van Inwagen says in this passage and (b) I don’t have space to explore the merits of Humeanism, I’d like not to pursue this issue further. Perhaps it will suffice to say the following. First, if it turns out that van Inwagen’s argument turns on various non-Humean claims, then all the better for me: I regard it as a virtue of TT that it ties our modal knowledge to our best theories, and I am glad to take the relevant non-Humean claims to be (part of) a theory. (This shifts the debate to the metaphysical theories themselves, and if nothing can be said in their favor that doesn’t presuppose contentious modal claims, then so be it: we will have to accept silence about the relevant modal matters.) Second, it is hard to provide a non-circular argument for *any* metaphysical thesis, and insofar as our metaphysical beliefs are justified, it seems to me that that justification is quite weak. (And again, it’s worth pointing out that many arguments for large-scale metaphysical views – like Humeanism – are bound to rely on modal claims that are in question here.) So, an

Given this, then it becomes much, much harder to see how we could come to know that an extraordinary proposition is possibly true. Take, for example, the suggestion that we can know that the moon could be made of green cheese by imagining this state of affairs:

...anyone who thinks he can imagine that the moon is made of cheese has a very sluggish imagination: the active imagination demands a pasture for the antecedently necessary thousands of thousands of millions of cows, demands a way to preserve a piece of cheese in broiling heat, freezing cold, and vacuum for thousands of millions of years, demands some off-stage *machina* to protect a piece of cheese thousands of miles across from gravitational compression into non-cheese, demands... But any *serious* attempt to imagine the moon being made of green cheese – and what besides a serious attempt could prove “coherence”? – must, like the unimaginable object itself, soon collapse under its own weight (van Inwagen 1979, 671-672).

The upshot is this: van Inwagen thinks that we’re justified in believing that *p* is possible if we can provide good reason to think that there is a world *p*-friendly world, but he despairs of our ability to do this for extraordinary propositions.

There are two natural reactions at this juncture. First, you can doubt whether the task before us is as difficult as van Inwagen suggests: is it *really* that hard to imagine a *p*-friendly world? Second, you can grant that it’s hard to imagine a *p*-friendly world, but doubt that we never succeed: is it plausible that our imaginings are *never* sufficiently detailed to provide evidence that a *p*-friendly world exists?

Van Inwagen thinks that a confused idea stands behind the first reaction: namely, that consistency is a reliable test for possibility. He points out that logical ‘possibility’ is not a species of absolute possibility (i.e., ‘metaphysical possibility’ or ‘possibility *simpliciter*’). We know this because we know that there are necessarily false propositions that are not contradictory: e.g., some things are simultaneously red all over and green all over; there are

epistemology of modality should be very cautious about making metaphysical assumptions – including Humean ones. This point also cuts against van Inwagen, of course, but if it’s true, then it’s enough to get the essence of his conclusion: since we should not be so confident that the metaphysical facts are congenial to our epistemic practices, we should not be so quick to say that *p* is possible, at least where *p* is an extraordinary proposition.

male vixens; I am a donut.⁸¹ It follows that the space of logical ‘possibilities’ includes some absolute *impossibilities*. (Contrast this with the space of physical possibilities – those absolute possibilities consistent with the actual laws of nature. Not all absolute possibilities are physical possibilities, but every physical possibility is absolutely possible; hence, physical possibility is a species of absolute possibility.) Now, if a proposition does not appear to be contradictory, then you are probably entitled to infer that it is not logically impossible. But this is to make an epistemological observation, namely, that the proposition in question cannot be shown to be absolutely impossible by logic alone. If logical possibility *were* a species of absolute possibility, then you could infer from its (probably) not being logically *impossible* to its (probably) being absolutely possible. But since it isn’t, you can’t. (*Objection:* There are very few absolute impossibilities that aren’t logical impossibilities; hence, logical possibility is an imperfect but nevertheless genuine guide to absolute possibility. In other words, while it’s true that the space of logical possibility is larger than the space of absolute possibility, the former isn’t *that much* larger than the latter. *Reply:* Both ‘spaces’ contain indenumerably many worlds; finding out that one isn’t *that much* larger is cold comfort. Moreover, if you know that there are very few absolute impossibilities that aren’t logical impossibilities, then you must know what is and is not absolutely impossible; otherwise, you’d be in no position to judge the ratio of absolute to logical impossibilities. How did you come by this knowledge?) The upshot: you don’t know that there is a history and set of laws with which *p*’s truth is compossible – i.e., you don’t know that there is a *p*-friendly world – just because you detect no contradiction when you search for one.

The second reaction is to doubt that our imaginings are *never* sufficiently detailed to

⁸¹ Van Inwagen doesn’t actually give examples, though he clearly has propositions like these in mind. Nor does he explain how we know that claims like these are necessarily false. I tend to agree that they are necessarily false, but I wouldn’t use them to object to the thesis that consistency is a guide to (absolute) possibility. As I suggested in Chapter 1, I think *a posteriori* necessities are more useful for this purpose.

provide evidence that a *p*-friendly world exists. But van Inwagen simply denies that we have the resources necessary to assess extraordinary claims: in some cases, the task is just too complex for us; in others, it requires information that we don't have. He uses Stephen Yablo's conceivability-based modal epistemology to illustrate this point (Yablo 1993). (He regards his view as correct in outline, if not in detail, but nothing turns on the particular features of Yablo's framework. The argument is supposed to apply to *any* view on which conceiving or imagining is supposed to be evidence of possibility – i.e., where telling a story or visualizing a state of affairs is supposed to be evidence that there is a world of which that story is true, or in which that state of affairs obtains.) Very roughly, Yablo's theory goes as follows. If *p* is conceivable for me, then I can imagine a world that I take to verify *p*. In the absence of defeaters, I'm justified in believing that things are as they seem. So, if I can imagine a world that I take to verify *p*, then I am justified in believing that there *is* a world that verifies *p* – which, of course, is to be justified in believing that *p* is possible. Therefore, if *p* is conceivable for me, then I am justified in believing that *p* is possible. Van Inwagen latches onto Yablo's first step, arguing that, for any extraordinary proposition, we have no reason to suppose that we've imagined a world that verifies it. He asks us to consider, for example, whether we can imagine a world in which there is transparent iron. His response is that we can't:

...unless our imaginings take place at a level of structural detail comparable to that of the imaginings of condensed-matter physicists who are trying to explain, say, the phenomenon of superconductivity. If we simply imagine a Nobel Prize acceptance speech in which the new Nobel laureate [...] displays to a cheering crowd something that looks [...] like a chunk of glass, we shall indeed have imagined a world, but it will not be a world in which there is transparent iron. [...] This sort of effort of imagination will [...] show that a certain proposition has the modal status "possible," but the proposition will be a disjunctive one. Here are some of its disjuncts:

- Transparent iron exists
- The scientific community has somehow been deceived into thinking that transparent iron exists

- A crackpot physicist who thinks he has created transparent iron is the butt of a cruel and very elaborate practical joke
- A group of fun-loving scientists have got together to enact a burlesque of a Nobel Awards Ceremony.

[...] No doubt, by working our imaginations a bit harder, we could imagine a world in which some of the “unwanted” disjuncts failed. We might, for example, add to what we have already imagined a codicil to the effect that all the scientists in the cheering audience are sincere. But this would not rule out the second of the above disjuncts (“mass deception”). To rule that out, our imaginations would have to descend to “a level of structural detail comparable to that of the imaginings of condensed-matter physicists who are trying to explain superconductivity” (van Inwagen 1998, 79-80).

So van Inwagen is *not* contesting the following conditional: if you could indeed imagine transparent iron, you would have some evidence that transparent iron is possible. What he contests is the antecedent of this conditional. His suggestion is that we only imagine transparent iron if our imaginations descend to a level of structural detail comparable to that of the imaginings of condensed-matter physicists who are trying to explain superconductivity. And, of course, that never happens: either the task is too complex or it would require knowledge that we lack. So, we aren’t justified in believing that there could be transparent iron.

5.3 **Problems for the Disjunctive Argument**

Let’s call the argument in the last block quote, ‘the Disjunctive Argument’ (DA). Heimir Geirsson reads it as insisting that, for any obscure proposition p , you don’t know that p unless you can imagine “all the relevant detail in the world that makes $[p]$ true” (Geirsson 2005, 285-288).⁸² Peter Hawke offers a similar interpretation; he claims that DA is based on a principle like this: you haven’t imagined a world that verifies p unless you’ve imagined a world in sufficient detail so as to rule out $\sim p$ (Hawke 2011). (Geirsson’s ‘relevant’ and Hawke’s ‘rule out’ can be given stronger and weaker readings, and presumably ones that make the two

⁸² For a virtually-identical interpretation, see (Cohnitz 2002).

interpretations extensionally equivalent.) Both worry that DA leads to a more pervasive skepticism than van Inwagen intends. For example, Geirsson considers whether DA is compatible with justifiably believing that an LP could play without pops and clicks (a modal claim that he regards as being ordinary, and so one that van Inwagen would say that we know).⁸³ If being able to do so depends on our being able to imagine all the relevant details of a world that verifies this proposition (or, more carefully, its non-modal counterpart), then I know that I'm sunk: *I* certainly can't imagine in complete detail the technology that would allow you to play an LP without background noise – not least because I can't imagine in complete detail the technology that allows you to play an LP *with* background noise. And, Geirsson reasons, if van Inwagen's argument undercuts our beliefs about ordinary modal matters, then we have cause to reject it.

This sort of objection misses the mark, and it's important to see why. Van Inwagen never says that our everyday modal beliefs are justified in the same way as are our beliefs about extraordinary propositions – not least because he denies that we have justified beliefs about extraordinary propositions. On his view, our ordinary modal beliefs are either properly basic, inferred from beliefs that are properly basic, or justified by inferences from actuality. There is no mystery in inferences from actuality to possibility. However, if asked to provide an account of properly basic beliefs, he declines: his official position is that, while we know that we have (properly basic) ordinary modal knowledge, we have no idea *how* we have it:

Although I do not doubt that we have some modal knowledge, I regard much of this knowledge as mysterious. Some modal statements, I have said, we know by reasoning from what I have called “basic” modal knowledge – simple, obvious modal statements

⁸³ I don't know whether Geirsson is right to say that this claim is ordinary. Here is an argument to the effect that it isn't obviously extraordinary. The following is an ordinary modal claim: an LP could play with *fewer* pops and clicks. And the following is not obviously an extraordinary modal claim: where *n* is greater than zero, there is no *n* such that an LP could not play with *n*-1 pops and clicks. So, that an LP could play without any pops and clicks is not obviously extraordinary.

whose truth we are somehow in a position to know – together with some facts about how the world is constructed. But how do we get started in this reasoning? How do we know the “simple, obvious” modal statements to be true? What is the ground of “basic” modal knowledge? I do not know how to answer these questions (van Inwagen 1998, 73-74).⁸⁴

Now, I don’t know whether van Inwagen would grant that our ordinary modal knowledge does not meet the standard to which extraordinary modal claims are held in DA. He could maintain – or at least I see no reason why he couldn’t maintain – that our acquaintance with a range of ordinary objects in ordinary circumstances explains why our imaginings or conceivings meet the relevant standard. (I don’t know how the explanation would go, but it would be an interesting line to pursue.⁸⁵) Alternately, he could dig in his heels, conceding that our ordinary modal beliefs do not meet DA’s standard, but insisting that this just shows them to have some *other* foundation – not conceiving or imagining.

Even if Geirsson’s objection fails, it does manage to reveal that we are without a clear argument for the high standard that DA employs, and therefore without a clear answer to doubts about the claim that our imaginings are *never* sufficiently detailed to provide evidence that a *p*-friendly world exists. What’s more, I think that there are deeper reasons to suspect that van Inwagen cannot provide the needed argument.

Let’s suppose that you need to be able to imagine all the details relevant to an

⁸⁴ Perhaps this is a way of saying that our everyday modal knowledge is Moorean. Or perhaps this ignorance is an implication of our everyday knowledge being Moorean. I’m not sure. Whatever the details, you might wonder whether this view actually works in his favor. If we know that we have modal knowledge, then even if we can’t explain it, wouldn’t it be natural to suppose that our actual practices are trustworthy (absent defeaters) not just for ordinary propositions, but for extraordinary ones too? I am not so sure. If van Inwagen is right that we can’t explain our modal knowledge, then surely we’re in a poor position to discriminate genuine cases of modal knowledge from spurious ones. And if that’s right, then we should tread very carefully once we move beyond those claims about which we are all quite confident. This is just a burden-shifting move: instead of being able to rely on the presumption that a belief is innocent until proven guilty, van Inwagen’s interlocutor is now under some pressure to offer a positive story about how she knows that there is a history and set of laws with which a given proposition is compossible; if she can’t, then her claim to know the modal status of a given extraordinary proposition becomes suspect. The upshot is that, if you don’t know how you know, then you shouldn’t trust yourself when you shift from a familiar to an unfamiliar context (i.e., from the ordinary to the extraordinary); and if you don’t know how you know, then you should be much more bothered by disagreement that you might be otherwise.

⁸⁵ I think that this is supposed to be Felipe Leon’s strategy in (Leon 2009). However, if Leon offers an explanation of how ordinary objects ‘close the epistemic gap’ (as it were), I do not understand it.

extraordinary proposition's truth in order to know its modal status. What might motivate this requirement? Recall the transparent iron example. On the assumption that you aren't simply *stipulating* that you're imagining a hunk of transparent iron, on what basis are you *identifying* it as such? Well, if you only identify *actual* samples of iron on the basis of expert testimony, then you are in no position to take your imagining as evidence of possibility.⁸⁶ Trusting your imagining would be like trusting the testimony of a hapless prospector: he can't tell the difference between fool's gold and the real thing; so, his claim to have struck gold should be eyed with suspicion. (And in fact, trusting your imagining would be *much, much* worse than trusting the testimony of the hapless prospector: he's got a fifty-fifty shot at being right, whereas there are many, many ways for you to have a false positive.) The 'all-the-relevant-details' requirement is supposed to ensure that your imagining has better epistemic credentials: the requisite detail makes it more plausible that you have the ability to identify the thing that you're imagining, and hence makes it more plausible that what *seems* possible *is* possible.⁸⁷ So far so good. But what constitutes imagining all the relevant details? I understand this as the requirement that you be able to imagine a scenario *s* such that, necessarily, if *s* were the whole world, then extraordinary proposition *p* would be true. In other words, you need you need to be able to imagine a *model* of *p*. This explains why the Nobel Prize acceptance speech scenario is insufficient, since the description of it is compatible with there being no transparent iron. It also explains why the right scenario would probably be extraordinarily detailed (e.g., as detailed as

⁸⁶ Moreover, it's no good to stipulate that the expert testimony is correct, since your imagining provides no reason to believe any such thing.

⁸⁷ Incidentally, Kripke is often cited as the authority on the legitimacy of stipulation when it comes to matters modal: see (Kripke 1980, 44-47ff.). But Kripke is not really interested in the epistemology of modality in the relevant section of *Naming & Necessity*. Rather, he's making a semantic point: we certainly seem to know that *this very man* could have been different in various ways, and the best semantic theory allows us to say as much (which, according to Kripke, you can't say if you adopt counterpart theory). If we *can* say as much, and if we know that Nixon could have lost the election, then we needn't characterize a world purely qualitatively. Instead, we can stipulate that we're talking about the world in which Nixon loses the election. Even if correct, this observation is no comfort to the defender of stipulation as part of a modal epistemology. More on this in the next chapter.

the model constructed by condensed-matter physicists who are trying to explain superconductivity). However, once we put the condition this straightforwardly, it quickly falls apart. Whatever a possible world *is*, it is somehow *maximal*, which is to say that (assuming bivalence for simplicity) every proposition is either true or false at that world. But there is no similar restriction on the construction of models: they *always* leave unspecified all sorts of details. So, our scenario *s* may well be consistent, but that it is does not entail (or even suggest) that there is some world of which *s* is a part: as van Inwagen argues, consistency is no guide to possibility. It follows that *p*'s being true of *s* is no evidence that *p* is possible, since it is no evidence that there is some world at which *p* is true. So, even if we could meet the standard in DA, the gap between consistency and possibility should lead van Inwagen to deny that we've done enough. (What *would* be enough? Well, I presume that if you could imagine *every* detail – not just every relevant one – then you could have extraordinary modal knowledge. In other words, you would need to imagine *a whole world* that verifies *p*. But of course this is impossible: no one can entertain that many propositions (either simultaneously or sequentially), and to demand as much is unreasonable. Granted, if you are already inclined to regard extraordinary modal knowledge as impossible, then van Inwagen may have a way to explain why it has this unfortunate status. But this is not much of a basis for modal skepticism.)

These problems make me suspect that we have not been charitable enough in our interpretation of van Inwagen's position. In conversation, though, van Inwagen assured me that he indeed holds a view like the one that Geirsson and Hawke attribute to him. I have no reason to doubt van Inwagen on the matter of van Inwagen interpretation. So, I conclude that van Inwagen's version of modal skepticism faces some serious problems, and that if the spirit of the view is correct, then we need to flesh it out in a better way.

5.4 The ‘Direct Evidence’ Version of Modal Skepticism

Here is an alternate route to modal skepticism that aims to preserve the basic structure of DA. Although I think that it fails, it’s failure is instructive.

When considering a scenario that’s *supposed* to verify the proposition, ‘there is transparent iron’, the natural worry is that the imagination doesn’t maintain a sufficiently tight grip on the substance in question: perhaps the scenario fails to depict a substance at all; or perhaps it depicts some familiar substance that we confuse for transparent iron (e.g., glass); or perhaps there is no fact of the matter about which of many transparent substances it depicts. Geirsson and Hawke are trying to make sense of how van Inwagen thinks that the imagination can ‘hang on’ to its object under various transformations. And this is, I think, the right goal, even if we have yet to find a way to reach it.

Perhaps we should say that you can’t successfully imagine p by imagining circumstances that would, if actual, provide just any old evidence of p ’s truth. Why not? Suppose that I know nothing about iron’s nature and some expert tells me that iron is transparent under such-and-such conditions. No doubt her testimony constitutes evidence for me that there are (or could be) samples of transparent iron. However, the value of her testimony is parasitic on a certain kind of non-testimonial evidence – I’ll call it ‘direct’ – that comes in two main forms: first, observation by someone with the appropriate discriminatory abilities; second, suitable confirmation of the theory that predicts that there are such samples. If we want to provide an account of the source of our justification about modal matters, then we should focus on direct evidence. So, if your aim is to imagine transparent iron, then you need either (a) to be able to discern genuine instances of transparent iron from imposters or (b) confirm a theory – based on a thought experiment – that predicts the reality of transparent iron. Good luck.

But the situation may be less dire than it looks. Consider what it would take to convince you that x is a sample of transparent iron. If the transparency is obvious, then the thing to do would be to check for iron's characteristic properties. (If you don't know what those properties are, no matter: just look them up. Isn't that what you'd do normally?) Suppose you find that your sample has all the relevant properties. No doubt you would conclude – or at least *could* reasonably conclude – that you were dealing with transparent iron. Of course, given your lack of expertise, perhaps you shouldn't be wildly confident about your conclusion, and wisdom recommends deference to an expert should one disagree with you. Still, if your evidence points to p 's truth, and you have no defeaters, then where's the error in believing?

Yablo's epistemology is based on the parallel between perceiving and conceiving; the presumption, therefore, should be that you can use the same epistemic tools in the imagination that are available in perception. In particular, you should be able to believe whatever would best explain your imaginative experience. So, if you can imagine a transparent sample having the properties characteristic of iron, and if its being a sample of transparent iron would best explain that experience, then the analogy seems to give you license to conclude that you're dealing with transparent iron. Note, moreover, that the standard for justified belief is sensitive to context. So, before you've been challenged, it may take very little imaginative evidence to justify the belief that there could be transparent iron. You'll have to provide more once your belief has been called into question, but why think that you can't? Surely you can imagine a bubbling reaction when you put a few drops of the right chemical on the sample, or imagine seeing the solution in the test tube change colors when you add some filings to it.

You might worry that this doesn't address one of van Inwagen's central concerns, namely, that if you believe that p is possible, then you are committed to there being a history and

set of laws with which p 's truth is compossible. Again, the analogy with perception is instructive. Consider this argument pertaining to beliefs about actuality:

If you believe that p , then you are committed to p 's being consistent with our world's history and laws (since we don't live in an inconsistent world). If you recognize this, then your entitlement to believe that p stands or falls with your entitlement to believe that p is compossible with our world's history and laws. Now consider some proposition p that meets the following conditions: (1) it concerns matters remote from the practical business of everyday life, (2) its truth-value cannot be determined by logic and reflection on the meanings of words or by the application of mathematical reasoning, and (3) its truth-value is unknown to us. You should be agnostic about whether p 's truth is compossible with our world's history and laws. Hence, you should be agnostic about p .

This argument is ridiculous. If the best explanation of your experience is that p obtains, then you are *thereby* justified in believing that the actual world's history and laws are compossible with p . If we model modal knowledge on perceptual knowledge, then the beliefs produced by the imagination will be innocent until proven guilty, and they will be aided and abetted by our methods of ampliative inference. Hence, if the best explanation of your imaginative experience is that p , then you are justified in believing that there is a history and set of laws with which p 's truth is compossible.

In this light, we might compare DA to the cartoon-version of Hume's critique of miracles: both are based on an implausibly stringent account of when it is reasonable to form a belief based on appearances. Suppose that a miracle is a violation of a law of nature, that you believe L to be a law of nature, and that it seems to you that you witnessed a violation of L . Hume (or cartoon-Hume) seems to think that you always have better reason to affirm L than to believe that the miracle has occurred: laws, after all, are supported by uniform experience, and we know that the senses sometime deceive us. One of the standard objections to this argument is that, if correct, we would never be in a position to discover anomalies of *any* kind – miraculous or otherwise. Since we surely can, the argument must be flawed. The problem is that this way

of weighing evidence is unduly conservative: it puts too much emphasis on preserving current theories, presumably in the interest of not disrupting the web of belief. But conservatism and consistency are not the only epistemic desiderata: it's also important to follow the evidence wherever it leads, whether or not doing so disrupts the web. The same is true in modal epistemology. If it *seems* to you that there are p -friendly worlds, and you aren't aware of any reasons to be suspicious of this seeming, then it would just be dogmatism to resist the belief that p is possible. Of course, it may not seem to van Inwagen that there are any p -friendly worlds, or he may think that he has a defeater for the seeming. But if the defeater is just that neither he nor anyone else can rule out the possibility of error, then his argument is on poor footing indeed. We are all fallibilists now. If van Inwagen is demanding certainty, then while we should admit that modal skepticism is true, we shouldn't care.

The long and the short of it is this: if we take seriously the analogy between modal and perceptual knowledge, then we seem to be left with an epistemology on which we have (or can have) *lots* of knowledge about extraordinary modal matters; so, van Inwagen doesn't actually provide a basis for modal skepticism. Therefore, if we want to preserve modal skepticism, then we will have to find an epistemology that is up to the task.⁸⁸

5.5 An Alternative Proposal

Summarizing his objection, Geirsson says that “[t]he problem with van Inwagen’s account is that it ties justification of p , where p is the relevant possibility statement, too closely

⁸⁸ My criticisms in this section suppose that there is indeed a useful parallel between modal and perceptual knowledge. I’ll question this assumption in the next chapter, arguing that the perceptual analogy should lead us to deny that imagining provides *any* evidence of possibility. But there are two reasons why this is no consolation to van Inwagen. First, he doesn’t want to deny the conditional: *if* we imagine that p , then we have evidence for p ’s possibility. He wants to deny only the antecedent. Second, it remains the case that the considerations that motivate his modal skepticism fit better with TT than with an imagination-based modal epistemology, as I’ll suggest below.

to both our present state of knowledge and our ability to imagine the details that make p possible” (Geirsson 2005, 288). In my view, these features aren’t flaws of van Inwagen’s account, they’re virtues. However, given the analogy between modal and perceptual knowledge, I doubt that van Inwagen’s account has them. So, if you want to preserve modal skepticism, then you ought to reject the parallel between perceiving and imagining or conceiving. TT does just this.

According to TT, our justification for believing any modal claim is closely tied to our present state of knowledge: for any modal proposition p , we are justified in believing p only if we justifiably believe a theory that says that p ; moreover, our justification for believing p is no better than is our justification for believing the theory that says that p .⁸⁹ And according to TT, our justification for believing any modal claim is tied to our ability to imagine the details that make p possible – or more accurately, our ability to *understand* the details that make p possible, at least on some level. TT doesn’t require that we (sensuously) imagine one of the models specified by the theoretical definition, but it does require that we justifiably believe a theory about p , and that theory will (or should) shed some light on why p is true in the world(s) in question. After all, our best theories are the best – at least in part – because they explain why propositions about actuality are true; so, it’s plausible that they can also explain why propositions about other worlds are true, and hence provide at least some details about why those propositions are possible or necessary.

Here is an example that puts these two features in perspective. Let’s suppose that you are competing in the something akin to the Pinewood Derby. The inclined portion of the track is one

⁸⁹ ...and it’s potentially much worse. Suppose that we have great evidence for a theory, but we aren’t sure whether this theory says that p ; or we think that it does, but we suspect that another theory that we justifiably believe says that $\sim p$; or... In such cases, our justification for believing p would be worse than our justification for believing the theory that says that p .

meter long and it is at a 30° angle to the flat portion. Now, if you were to set your car at the top of the track, it is possible that your car reach a velocity of 20m/s before it hits the bottom? If TT is correct, then it's natural to begin answering this question by revisiting our high school physics. As you reduce resistance, what is the limit that the velocity approaches? The details of the answer matter. We begin by recalling that the force of gravity, F_{grav} , equals $(m \cdot g)$. That force is perpendicular to the ground, not the angled part of the track; so, we have to decompose F_{grav} to find the car's acceleration: one part parallel to the inclined part of the track (F_{parallel}); the other part perpendicular to it ($F_{\text{perpendicular}}$). If the angle of inclination is 30° , then a little geometry will reveal that the angle between F_{grav} and $F_{\text{perpendicular}}$ is also 30° . This means that F_{parallel} is $(F_{\text{grav}} \cdot \sin 30)$, which is equal to the net force on the car along the track (because we're ignoring friction). By Newton's second law, F_{parallel} is equal to the mass times the acceleration of the car. So, $(F_{\text{grav}} \cdot \sin 30) = (m \cdot a)$. Solving for a (and remembering that $F_{\text{grav}} = (m \cdot g)$), we get $a = (g \cdot \sin 30)$. Next, we take a formula from Newtonian kinematics ($v_{\text{final}}^2 = v_{\text{initial}}^2 + (2 \cdot a \cdot d)$) and plug in (1) the distance traveled (one meter), (2) the initial velocity of the car (zero, since you simply place the car at the top of the track) and (3) the standard value for gravity (9.80665 m/s^2). The result is that the maximum final velocity of the car is (roughly) 3.13m/s. In other words, barring any other considerations, TT would have us say that the car *could not* reach a velocity greater than 3.13m/s; it is necessary that its velocity be less than (or perhaps equal to) 3.13m/s. So, it is *not* possible that the car reach a velocity of 20m/s before it hits the flat portion of the track.

Our theory gives us some insight into why the car couldn't reach the relevant velocity, and this confirms Geirsson's worry about our being required to imagine the details that make p possible. In the present case, once we recognize the relationships between the various factors

that influence the final velocity – the force of gravity, the angle of the track, the initial velocity, etc. – we see why some final velocities are possible and other final velocities are not. Now, someone is bound to complain that the car could have a much higher velocity than the one that I’ve calculated. Why? Because gravity could be stronger, or because the conservation law (which is assumed but not mentioned above) could be replaced by the *multiplication* law, according to which potential energy converts to kinetic energy at double (or triple, or quadruple) the actual rate. But given the truth of TT, the details above *reveal what it would take to justifiably believe that the final velocity could be higher than the one that I derived.*⁹⁰ If you want to say that it is merely *physically* necessary that the car’s final velocity be less than (or perhaps equal to) 3.13m/s, but that it is possible *simpliciter* that the velocity be higher, then you need to justifiably believe a theory on which a proposition like one of the following is possible:

- F_{grav} takes some value other than 9.80665 m/s^2 ; or
- Newton’s law postulates a different relationship between force, mass, and acceleration; or
- the principles of kinematics suggest some other relationship between initial and final velocities, acceleration, and the distance travelled.⁹¹

So if, for example, you justifiably believe a theory according to which gravity could be stronger than it is, then you are free to use that theory to recalculate the highest possible velocity. (But what is that theory, exactly? It will not do here to contend that we could move our track to Jupiter – whose mass is greater, and so where gravity is stronger, and so where a car could reach a higher velocity. The question is not whether the car could reach 20m/s *anywhere*, but whether it could reach that velocity *here*, on earth, and relatively close to the earth’s surface.) It is a consequence of TT, however, that if you don’t justifiably believe a theory that says that gravity could be stronger (or that something else could be different that would affect the velocity of the

⁹⁰ This is not exactly right, since (for example) you could get the same result using the equations for kinetic and potential energy. But the general point remains. Thanks to Brian Stachowiak for pointing this out to me.

⁹¹ There are not the only options: you could also look for alternate geometries, or mathematics, or logics, or what have you. But lists have to end somewhere.

car), then you aren't justified in believing that the velocity could be any higher than is suggested by the calculations above, and this for any sense of 'could' you like.

Two clarifications. First, TT *doesn't* imply that you need to know the physics just sketched in order to justifiably believe other modal propositions about our derby: e.g., that car could be moving at a velocity greater than one meter per *month*, or that the car could be moving fast enough to reach the end of the flat portion of the track. Other, less precise theories (folk physics, or folk heuristics for actual physics) probably suffice to justify these modal claims. (Likewise, TT doesn't say that you need to know the physics just sketched to *rule out* the possibility of the car's velocity being 20m/s; again, less precise theories may do the trick.) However, TT *does* say that if (1) you justifiably believe the physics above and (2) you do not justifiably believe a theory on which it's possible for the relevant values and relationships to be different than they are, then you are not justified in believing that the car's velocity could reach 20m/s. And this consequence, I take it, is enough to qualify TT as a version of modal skepticism.

I think that this version of modal skepticism improves on van Inwagen's in at least three ways. First, it provides a better account of what it is to be an extraordinary proposition. Recall van Inwagen's conditions:

1. it concerns matters "remote from the practical business of everyday life,"
2. its truth-value "cannot be determined by logic and reflection on the meanings of words or by the application of mathematical reasoning," and
3. its truth-value is "unknown to us or [is] known to be false" (van Inwagen 1998, 74, 84).

Van Inwagen puts Condition (1) more strongly than he needs to (and perhaps more strongly than he even intends to): modal skepticism needn't be tied to the business of everyday *life*, if this refers to the kinds of activities and interests that we have in common; rather, it can be tied to the business of everyday *lives*, so that specialists can have modal knowledge that is unavailable to

the rest of us. It would be nice to be able to grant, for example, that physicists and day traders have the modal knowledge appropriate to their respective fields, but not the knowledge appropriate to each other's fields (hobbyists and the bivocational aside). TT's version of modal skepticism can manage this, since it is plain that we do not all believe the same theories, and so do not all have the same modal knowledge available to us. TT offers an account of extraordinary propositions relative to persons and times: p is an extraordinary proposition for an individual i at time t if p is not true according to any of the theories that i justifiably believes at t .

Second, TT does not require us to adopt a bifurcated epistemology, with different stories about the requirements for ordinary and extraordinary modal knowledge.⁹² One and the same story about the justification of our theories – whatever it may be – explains both why we are justified in believing the one sort of proposition and why we are *not* justified in believing the other. Moreover, TT assimilates our modal knowledge to our theoretical knowledge, so there is a sense in which it doesn't even require us to accept a distinctly *modal* epistemology. We all need some account of how our theories are justified, and there is widespread sympathy for a model-based approach to theorizing amongst philosophers of science. TT simply puts these two theses together. So, TT does not require us to multiply sources of justification; it does not even require us to extend or retool epistemological principles to which we are already committed. TT simply points out that a plausible epistemology for ordinary and scientific knowledge *includes* a modal epistemology: if most of our ordinary and scientific beliefs are justified, and if most of those beliefs can be understood as being theoretical, then we have much of what we need to

⁹² Or, more carefully, *if* there is any reason to acknowledge properly basic modal beliefs (say), that reason has nothing to do with the ordinary/extraordinary distinction. And that is a virtue too. Suppose that we were to encounter a theory of perception on which we cannot see microphysical particles with the naked eye, *and* on which we can't see medium-sized dry goods with the naked eye. If we were to follow van Inwagen's lead, then we wouldn't reject the theory outright; rather, we'd accept it, and then insist that our perceptual beliefs must be justified in some other way. Now, perhaps there is some reason to think that our perceptual beliefs have some other basis than the one that they appear to have, but it is awfully implausible that this reason has anything to do with the difference between microphysical particles and medium-sized dry goods.

make TT work.⁹³

Third, TT can handle an objection to modal skepticism that van Inwagen cannot. The backbone of van Inwagen's argument for modal skepticism is the following claim: you justifiably believe that p is possible only if you justifiably believe that there is a p -friendly world – i.e., that there is a history and set of laws with which p 's truth is compossible. But how do we know that *every world* has a history and set of laws? In other words, how do we know that it is *necessary* that there be a history and set of laws? This claim isn't an extraordinary proposition *per se*, since its non-modal counterpart is not unknown to us or is known to be false – i.e., we know that our world has a history and set of laws (or, at least, I'm glad to grant as much). However, the claim is *essentially* an extraordinary proposition, since (a) it concerns matters remote from the business of everyday life and (b) its truth value cannot be determined by logical or mathematical reasoning or from the meanings of words. So, it is not one that we would expect van Inwagen to claim to know. But claim to know it he does, and without it his argument falls apart. What can he say in his own defense?

Nothing comes to mind. However, TT is not in the same position. According to TT, we are justified in believing that there must be a history or set of laws only if we justifiably believe a theory that says as much. *Perhaps* we justifiably believe a theory according to which there are worlds without the passage of time, and hence without histories. (We might call these 'solid-state' worlds.) I doubt that this is so, but I don't want to argue about it here. What we do *not* believe is a theory according to which there are worlds without laws. If TT is correct, then the argument for this is simple. If we are justified in believing that there is a world without laws, then we justifiably believe a theory that has a lawless model. By definition, a theory's models are just those models that satisfy its theoretical definition. And the theoretical definition

⁹³ I expand on this point in the next chapter.

represents (at least some of) the laws that hold in the models that satisfy it.⁹⁴ Hence, if TT is correct, then you are not justified in believing that there is a world without laws, since there is no model of any theory that corresponds to this alleged possibility. It is, moreover, quite plausible that the real work in van Inwagen's argument is being done by the laws, and not the history. So, it's plausible that TT can salvage van Inwagen's argument for modal skepticism, whereas his own epistemology leaves it in rough shape.

5.6 Some Cases

Before wrapping up this chapter, I want to show that my version of modal skepticism agrees with van Inwagen about cases, even if not with the considerations that stand behind his judgments. I think it's plain, for example, that very few of us – if any – justifiably believe a theory according to which there could be transparent iron. According to TT, then, we aren't justified in believing that there could be such a substance. But to show that this isn't a special case, I'll present two other examples.

5.6.1 The Problem of Evil

First, let's consider van Inwagen's objection to Richard Gale's version of the evidential problem of evil (as it appears in (Gale 1996)). Here is Gale's argument. Let g be the proposition that God exists. Let e be the proposition that details every apparently gratuitous evil of which we're aware, and suppose further that e includes all we know about each evil's causal antecedents and consequences, at least insofar as that information is relevant to whether the evil's occurrence secured some greater good or was necessary to prevent some greater evil.

⁹⁴ Or, if you prefer, the theoretical definition represents (at least some of) the invariant features of those worlds, and hence (at least some of) the features that determine what's possible in those worlds.

Gale's contention is this: you can infer from e that, probably, there is at least one gratuitous evil; and, relative to our background knowledge k , the probable existence of at least one gratuitous evil lowers g 's probability. In short: $P(g/(k \ \& \ e)) < P(g/k)$.

Van Inwagen objects as follows. Even we're justified in believing that, probably, there *are* some gratuitous evils, we should therefore assign a lower probability to God's existence only if our background knowledge includes the proposition that God could have prevented them.⁹⁵ And, he argues, we have no reason to suppose that God could have done any such thing; we just don't know what probability to assign to the proposition, 'God could have prevented there from being any gratuitous evils'. Now, I'm not going to discuss van Inwagen's argument for this conclusion, since (a) it is very similar to his argument against our knowing that there could be transparent iron and (b) I've already said what I want to say about that general approach. My aim here is to show that TT reaches the same conclusion.

The crucial question is this: do you justifiably believe a theory that entails that God could have prevented there from being any gratuitous evils? If you don't, or if you do and don't realize that the theory has this implication, then you aren't justified in believing as much, at least by TT's lights. Well, what theory could do the trick? Presumably, the right theory about God's power would suffice. Given a suitably generous account of the limits of omnipotence, it might follow that God could have prevented there from being any gratuitous evils. But how does this theory work, and whence your justification for believing it?

There are two points to make here. First, as is well known, it won't do to have a theory that implies that God can do absolutely *anything*, since that leads to a contradiction: if God can do absolutely anything, then he can create an unliftable rock; if he can create an unliftable rock, then there is something that he can't do (i.e., lift the unliftable rock), and hence he can't do

⁹⁵ ...and would have. For simplicity's sake, I'll suppress this extra condition in what follows.

absolutely anything. So, at bare minimum, the theory needs to be restricted enough to avoid this nasty implication. Second, consider what might count as evidence for a theory of God's power. Presumably, it's going to be the data that drives some version of either the cosmological or teleological argument: e.g., the existence of contingent objects, or the causal powers of contingent objects, or the fine-tuning of the universe for life, or what have you. (Forget about the *success* of any version of the cosmological or teleological argument: justification is relative to your epistemic situation, so someone might justifiably believe a theory about the scope of God's power on the basis of an utterly hopeless argument, assuming that she doesn't see its flaws.) But it's hard to see how any of this data will provide – or even give you reason for an opinion about – the answer to the question at hand. Suppose you justifiably believe that God is responsible for the existence and arrangement of all contingent objects. What should you infer, on this basis, about the contingent objects that God could have created, or about the arrangements in which he might have placed them? Not much, or so it seems to me. Assuming that the data supports the existence of God, it clearly provides a reason to think that God is very powerful. But we know that the theory shouldn't attribute completely unrestricted power to God. So, any attempt to flesh out 'very powerful' has to obey this stricture. But what other guidance does the data provide? Remembering the problem with appeals to 'logical possibility', you can't say that God can do whatever's logically possible: that lets in too much. You might say, then, that God can do whatever's metaphysically possible (possible *simpliciter*). This is, no doubt, one of the more elegant hypotheses available, and might be preferable in part on those grounds. But this theory is *epistemologically* useful only insofar as you have justified beliefs about what's metaphysically possible. On what grounds do you believe it to be possible that there could be a world *sans* gratuitous evils? What theory has this consequence, and on what basis do you

believe it?⁹⁶ Without answers to these questions, you won't be able to make any inferences about God's ability or inability to create a world *sans* gratuitous evil: at least with respect to this issue, the scope of his power will be inscrutable to you.

(N.B., it is no small feat to answer these questions. To choose just one of the many ways to illustrate this point, recall that natural disasters are commonly cited as instances of gratuitous evils (the 1755 Lisbon earthquake being a favorite of several authors). So, a world without gratuitous evils will be a world in which no natural disaster adversely affects human beings. What theory do you justifiably believe that implies that there is a world in which humans never suffer from the baleful effects of hurricanes, earthquakes, floods, and volcanic eruptions? None, of course. It follows that we are in no position to claim that there could be a world *sans* gratuitous evil, and hence in no position to claim that God could have actualized such a world. And this, of course, is just the conclusion that van Inwagen recommends.⁹⁷)

A few reflections. You might suppose that the above supports the view our theory about God's power is vague or in some way incomplete. I'm inclined to say, rather, that we don't have a theory of omnipotence at all. One reason for saying this is due to the fact that it's very hard to articulate a theory of omnipotence without explicitly using modal notions, as if being-able-to-do-what's-metaphysically-possible is a property that has some serious explanatory power. In my view, however, this claim is doubtful.⁹⁸ Suppose, then, that we have what counts as a sketchy theory of the scope of God's power. Now, ampliative inference strategies are only required

⁹⁶ This problem doesn't threaten the viability of the theory as an explanation of any facts about actuality. Since the contents of the actual world are actual, you can infer that they're possible. You know, therefore, that it follows from the theory that God could create the actual world.

⁹⁷ TT puts some teeth on one of van Inwagen's more memorable comments. He writes, "if you think that it would be possible to design a planet, and a universe to contain it, that was both capable of supporting human life and contained no earthquakes or tornadoes, I can only point out that you have never tried" (van Inwagen 1995, 106).

⁹⁸ If I'm correct, then this fact probably diminishes, but does not destroy, the explanatory power of the God hypothesis. We needn't understand everything about the molecules of a gas in order to know that those molecules best explain Brownian motion – though greater understanding would probably make deeper (and hence better) explanation possible.

when you've got underdetermination of theory by evidence; so, it's not underdetermination *per se* that's problematic here. Rather, it's that the above theory of God's power (God-can-do-anything-that's-metaphysically-possible) doesn't specify, all on its own, more or less determinate scenarios as possibilities-that-God-could-have-actualized. Instead, it relies on our having an independent grip on the notion of metaphysical possibility. Hence, the *epistemological* payoff of the theory about omnipotence is limited; it only tells you that, for any state of affairs *s* that we know to be possible, God could actualize *s*. And this is hardly enough for Gale's argument. So, TT sides with van Inwagen against Gale's version of the evidential problem of evil.

5.6.2 The Ontological Argument

Van Inwagen also discusses the modal ontological argument. He summarizes it this way:

1. It is possible for there to be a perfect being (that is, a being that has all perfections essentially).
2. Necessary existence is a perfection.
3. So, there is a perfect being (van Inwagen 1998, 67).

As I indicated at the outset of the chapter, van Inwagen denies that we are in a position to form a justified belief about (1)'s truth or falsity. He contends, first, that there is no good argument for (1), and second there is no reason to suppose that (1) is the sort of claim that we would be justified in believing without argument. In defense of the first point, he rehashes concerns about inferences from 'logical possibility'; in defense of the second, he draws an interesting analogy.

...suppose we call a real number *septiquaternary* if '7777' occurs in its decimal expansion; and let us call a real number *perimetric* if it measures the circumference of a circle whose diameter measures 1. Then

Possibly, something is septiquaternary and perimetric

(or, alternatively,

Septiquaternarity and perimetricity are compatible)

is obviously such that no philosopher has a “right to his opinion” about its truth-value (van Inwagen 1977, 391).

This analogy made for a better argument in 1977. Now, all of ten seconds on Google reveal that it is indeed possible to have something that is septiquaternary and perimetric. (Two fun facts: the first such string begins at the 1,589th digit in π 's decimal expansion, and there are 84 such strings in the first million digits.) Still, it's certainly the case that *before* you perform a Google search, you have no right to an opinion about whether it is indeed possible to have something that is septiquaternary and perimetric. But (1), he argues, is just like ‘Possibly, something is septiquaternary and perimetric’, and since we have no right to an opinion about the one, we have no right to an opinion about the other. Van Inwagen doesn't explain what the similarity consists in, but it's plausible (a) that it isn't up to us whether either proposition is true, (b) that neither proposition matters much for the business of everyday life, and (c) that both propositions concern matters that are fairly complex. Hence, it's plausible that neither proposition is of the sort on which we should venture an opinion without a tolerably good argument.

I think that van Inwagen's case against (1) missed the mark. Suppose that you justifiably believe the conclusions of Aquinas' Third and Fourth Ways (which are supposed to demonstrate the existence of a necessary being and a maximally perfect being, respectively), and you justifiably believe that these conclusions are true of one and the same being. (Again, it doesn't matter whether the reasoning is actually valid; what matters is whether someone could justifiably believe that it is, and so justifiably believe the conclusion. And surely someone could: Aquinas, for example.) Now, if you justifiably believe that there is a maximally perfect necessary being, then no doubt you could be justified in believing that it is *possible* there be a maximally perfect necessary being; even a very weak modal logic will sanction that inference. The problem with the ontological argument is not, therefore, that no one is in a position to take a stand on (1)'s

truth. Rather, the problem is that the people who *are* in that position don't need the modal ontological argument, and those who aren't in that position won't be convinced by it. In other words, (1) is the sort of claim that only a theist would believe (and not all theists at that), and so the argument has no force. If an atheist were ever disposed to concede (1), that generosity would surely (and rightly) vanish as soon as she saw that necessary existence is being taken to be a perfection. Since the atheist doesn't think that a maximally perfect necessary is actual, she therefore won't think that one is possible.

TT can account for this straightforward problem with the modal ontological argument. No atheist should affirm that it is possible for there to be a perfect being, since no atheist should accept a theory according to which that proposition is true. And no theist will affirm that it is possible for there to be a perfect being if he doesn't believe a theory according to which that proposition is true (Richard Swinburne and Keith Yandell, for example, are theists who appear to reject (1) for just this reason; see (Swinburne 2004; Yandell 1988)). Moreover, if a theist *does* justifiably believe such a theory, then she already justifiably believes a theory according to which there exists a maximally perfect necessary being; in this light, it's hard to see what the modal ontological argument adds to her justification for theism.

But let's return to the main line of argument. If van Inwagen offers the wrong diagnosis of the problem with the modal ontological argument, this does nothing to show that his skepticism about (1) is misplaced. The most plausible theory according to which (1) is true is probably one that tries to explain the existence of contingent, imperfect beings – a theory that Aquinas tries to motivate in his Third and Fourth Ways. But elsewhere, van Inwagen argues rather forcefully that the cosmological argument (of which the Third Way is one instance) does

not succeed.⁹⁹ First, the argument depends on some version of the Principle of Sufficient Reason (PSR), which faces a number of well-known challenges. Perhaps the most formidable is the threat of modal collapse. Let's say that 'C' is the conjunction of all true contingent propositions. C is true if its conjuncts are true, and C is contingent if even one of its conjuncts is contingent. So, C is a contingent truth.¹⁰⁰ By the PSR, there is an explanation of C's truth. If the explanans is a contingent truth, then it is either C or one of C's conjuncts. But no contingent truth explains itself, so either way, the explanans cannot be contingent. On the other hand, if the explanans is a necessary truth, then C is necessary, since the explanans entails the explanandum. Therefore, if the PSR is true, then C is necessary, and hence there are no contingent truths.

Second, even if the modal collapse problem can be addressed, the cosmological argument faces another formidable challenge. In the face of the modal collapse problem, the standard move is to weaken the PSR: you deny either (a) that no contingent truth explains itself or (b) that the explanans must entail the explanandum. I see no hope for (a). But if you opt for (b), then it becomes much, much harder to show that Explanation X is better than Explanation Y:

X: Contingent beings exist because a necessary being caused them to exist.

Y: Each world is as likely to be actual as the next. So, it's true that *there are some contingent beings or other* because one world had to be actual, ours is that world, and our world contains contingent beings.¹⁰¹ And it's true that *there are these contingent beings as opposed to some others* because one world had to be actual,

⁹⁹ See, e.g., (van Inwagen 2009).

¹⁰⁰ I recognize that this formulation leads to a version of Russell's paradox, but there is no reason to complicate the story to avoid it since the details would only obscure the main point.

¹⁰¹ A stronger non-theistic explanation is available for this proposition, though not for the next one. Contingent beings exist in every world but one; so, the probability of there being contingent beings is very high. (Indeed, if n is the number of worlds, then probability of there being contingent beings is $(n-1)/n$.) So, it's true that *there are some contingent beings or other* because it is highly likely that a contingent-being-world is actual. (Why think that there is only one world without contingent beings? Well, we are considering two pictures of how things are: one – X – according to which there is at least one necessary beings who is an agent (or very much like an agent), and another – Y – according to which there are no such necessary beings. If Y is true, then it's safe to say that the only necessary beings are inert objects – like abstracta – whose (non-Cambridge) properties do not change from world to world. So, all the worlds without contingent beings are exactly alike in every respect. But there is no reason to postulate multiple worlds that are exactly alike in every respect; one will do. Therefore, according to Y, there is only world without contingent beings.

ours is that world, and our world contains these contingent beings.

Plainly, Y's viability depends on the claim that each world is as likely to be actual as the next. But there is nothing suspect about this claim. First, we are considering two pictures of how things are: one – X – according to which there is at least one necessary beings who is an agent (or very much like an agent), and another – Y – according to which there are no such necessary beings.¹⁰² If Y is true, then it's safe to say that the only necessary beings are inert objects – like abstracta – and hence not the sorts of objects that can affect the probability of which world is actual.¹⁰³ That is, they are not like God, whose character might make it extraordinarily more likely that some worlds would be actualized over others. Second, it is no good to say that the relative complexity of some worlds makes them less likely than others: a random sequence of heads and tails is more complex than an equinumerous sequence of heads only, but the sequences are equally likely. Third, there is not, of course, a force 'outside' the worlds that pressures some toward actualization at the expense of others. So, it's plausible that each world is as likely to be actual as the next. And given this, I see no reason to prefer X over Y (indeed, I tend to think that Y is the more plausible story, due to its simplicity. It appears, therefore, that we can't motivate the 'necessary being theory' of the existence of contingent entities, and so there is no reason to think that it is possible for there to be a perfect being.¹⁰⁴

Given the arguments in this section and the previous one, it is quite plausible that TT

¹⁰² Note that Y does not rule out the possibility of supernatural agents *per se*; it simply rules out the possibility of *necessary* supernatural agents.

¹⁰³ ...except by ruling out some worlds as having zero probability of being actualized – namely, the impossible worlds.

¹⁰⁴ What's more, if Y is in fact a better explanation than X – and the Condorcet winner generally – then we should probably say that (1) is false. On another note entirely, a different theory that might underwrite (1) is the Medieval doctrine of the Transcendentals. On this view, being and goodness are identical; hence, having 'maximal being' (which entails, though is not the same as, necessary existence) involves having 'maximal goodness' (which I take to be equivalent to being maximal perfect). Insofar as this view requires the existence of God, any evidence for it would be evidence for (1) as well. But I don't know what that evidence could be – or, rather, I don't know what the evidence could be if the truth if theism is not part of our background knowledge.

agrees with many of van Inwagen's intuitions about which propositions are beyond our epistemic reach, even if it rejects the reasoning that is supposed to underwrite those judgments.

5.7 Conclusion

I have not argued that modal skepticism is true. It is, however, an attractive position. The upshot of this chapter is that TT offers a viable version of modal skepticism, and one that clearly improves on van Inwagen's. So, insofar as we have good reason to be modal skeptics, we have some reason to accept TT.

6. THE OPTIMIST'S OBJECTION

6.1 Introduction

We've all felt the pull of the skeptic's reasoning. ('But in the end, isn't certainty what we're after?') We rightly reject it, though, to avoid the conclusion it entails. Why 'rightly'? *Inter alia*, an epistemology ought to draw a plausible distinction between the known and the unknown, where 'plausible' should be read as 'more or less in accord with our pre-theoretic judgments about what's known and unknown'. The skeptic's standard isn't successful on this score; indeed, it fails miserably. And for this reason, we're within our rights to reject it.

You might make an analogous charge against TT. I've argued that modal skepticism and TT go hand in hand; hence, sympathy with the former ought to lead to sympathy with the latter. But many won't be sympathetic to the former. And 'won't be sympathetic' is surely an understatement. Here, for example, is Richard Gale's reaction to (van Inwagen's) modal skepticism:

I not only disagree with van Inwagen's [skeptical] modal intuitions but fear that he suffers from a serious modal affliction, in spite of his being an excellent philosopher from whom I have learned much. In *On the Nature and Existence of God*, I tried to resolve my modal disagreements with the likes of van Inwagen and Phil Quinn through my modal intuition bowl, but it proved a bust. The networks dropped us because there wasn't enough violence, just a bunch of out-of-shape guys, with the exception of Al Plantinga, who looks like he can jump tall mountains in a single bound, staring at each other and emphatically asserting back and forth, "It is possible that *p*," "No it isn't!"

I have decided to take a more radical, therapeutic approach and have founded the EMDS (Extreme Modal Deficiency Syndrome) Foundation. EMDS is no respecter of rank or philosophical orientation – even tenured analytic philosophers have come down with it. It is tragic to realize that there are people like Peter (my poster person) who cannot modalize as normal people do, or, as we at the Foundation prefer to say, are modally other-abled (Gale 1996, 213).¹⁰⁵

Pretty clearly, Gale takes himself to be within his rights to reject modal skepticism. We normal

¹⁰⁵ Gale's reply is not to (van Inwagen 1998), but to two earlier articles. Gale's rhetoric may explain, at least in part, the tone of van Inwagen's 1998 rejoinder.

folk know that there are cups of coffee; hence, Cartesian skepticism is false. We normal folk *also* know, apparently, that God could have created a world in which human beings suffer no gratuitous evils; hence, modal skepticism is false. Whatever entails a falsehood is false. Hence, if TT entails modal skepticism, then TT is false. Let's put a friendly spin on the aplomb embodied by this criticism, calling someone an 'Optimist' if he has the above view about the scope of, and appropriate degree of confidence in, our modal knowledge. His criticism of TT, then, I'll call 'the Optimist's Objection'.¹⁰⁶

I do not know how to argue directly for the thesis that our modal knowledge is as limited as TT suggests. I have already reported why *I* am not an Optimist: we don't interact with the truthmakers for modal claims, and there is no reason to suppose that the modal facts are organized in a way that's convenient for us; so, some measure of caution is prudent. But I have also conceded that, while I take these considerations to justify a particular threshold of epistemic risk, they may also be compatible with a range of other thresholds, and perhaps even ones that are incompatible with any form of modal skepticism. So, I am without a knockdown argument for the specific restrictions that TT recommends, and hence I am without a straightforward way to turn back the Optimist's Objection.

I do, however, have two more circuitous strategies available to me. First, if the argument of Chapter 4 goes through, then any reasons to accept TT are *ipso facto* reasons to accept modal skepticism. So, any argument for TT is thereby an argument against the Optimist. Call this 'the Indirect Argument Strategy'. However, because I'll sum up the case for TT in the Conclusion, I won't pursue this strategy now.

The second strategy is the one on which I'll focus. Let's begin by noting that the

¹⁰⁶ Robert Farley – optimist that he is – first leveled the Optimist's Objection against TT. I thank him for raising the worry that drives this chapter, and I'm sure that he won't be satisfied with my solution.

Optimist is vulnerable in a way that the Cartesian skeptic is not. The Cartesian skeptic denies that we know much of anything, and so we dismiss him outright. Of course, the Cartesian skeptic can complain that we have no better epistemology to offer – i.e., no epistemology that isn't either (a) equally conservative or (b) far too liberal even by our own lights – but this needn't bother us. Even when we take into consideration our inability to replace the Cartesian skeptic's epistemology with a better one, it's still more plausible that we have some knowledge than none. But it's much less clear that the modal skeptic is in the same weak position. The modal skeptic grants that we have quite a bit of modal knowledge, and the milder the form of skepticism, the less plausible its outright dismissal. So, let's suppose that the modal skeptic can make it plausible that there is no epistemology that isn't either (a) at least as conservative as the one that underwrites modal skepticism or (b) far too liberal by anyone's lights. Then, I think, the burden of proof shifts: the Optimist should either provide a suitably Optimistic epistemology or accept the one that underwrites modal skepticism. Or so it seems to me. Call this 'the Burden-Shifting Strategy'. The aim of this chapter is to mitigate the force of the Optimist's Objection via the Burden-Shifting Strategy: my task is to show that, in one way or another, three prominent and representative modal epistemologies fail to underwrite knowledge of extraordinary propositions.

To make matters more concrete, let's grant that the following proposition seems to be true:

(G) God could create a world in which human beings suffer no gratuitous evils.

What might justify being skeptical about (G)? Well, suppose that only an unscrupulous modal epistemology could underwrite our knowing (G). And suppose that the well-behaved modal epistemologies were either more or no less conservative than is TT. And suppose that we could explain the appeal of (G) without reference to (G)'s truth. In such circumstances, it would surely

be reasonable to doubt (G), as I do. (There is nothing special about (G); what I say here should generalize to any extraordinary proposition. I choose this example only because it's one that van Inwagen and I agree to be extraordinary.) I want to show that these are precisely the circumstances in which we find ourselves, and hence that the Optimist's Objection poses no threat to TT. I'll make my case by examining four illustrative modal epistemologies: Amie Thomasson's semantic approach, Stephen Yablo's defense of conceivability as a guide to possibility, and Timothy Williamson's counterfactual theory of modal knowledge. Hopefully, what I have to say against these views will make it clear what I'd say against their kin.

6.2 Thomasson

The aim here is to show that Thomasson's view is too permissive to plausibly underwrite knowledge of (G): i.e., the claim that God could create a world in which human beings suffer no gratuitous evils. If the argument goes through, and if I am right that there is nothing special about (G), then Thomasson's view doesn't plausibly underwrite knowledge of any extraordinary modal propositions.

To be clear, Thomasson rejects the realism about modality that I assume here. Her preferred brand of anti-realism is *normativism*. According to the normativist, claims about necessity are claims about how terms ought to be used; claims about possibility are claims about how terms can be used. I'm going to ignore this issue. I am really interested in the way that Thomasson's view could be *used*, not in the way that she uses it. Her view is of interest because she is particularly clear about the alleged relations between modal claims and semantic facts – and in particular, clearer than two-dimensionalists like (Chalmers 1996) and (Jackson 1998). So I want to ask this question: suppose that you could tell a story on which there is some non-trivial

correlation between paradigmatic semantic facts and objective modal facts; if so, could we reasonably extend that story so that our semantic knowledge could be used to acquire extraordinary modal knowledge? My answer is that we couldn't. Whatever we know about the meanings is unfit to provide us with knowledge of extraordinary modal matters. Meanings simply aren't fine-grained enough to underwrite knowledge of propositions like (G).¹⁰⁷

6.2.1 Thomasson's View *qua* Modal Epistemology

So let's make the false assumption that Thomasson wants to use her view to offer a modal epistemology. A philosopher who goes this route owes us a theory about the nature of truth and falsity in virtue of meaning. She also owes us a theory about how that truth or falsity is to be assessed. Thomasson takes up these challenges. She asks us to suppose that the meanings of singular and general terms are metalinguistic rules that characterize the correct uses of those terms within the object language (according to your favorite theory of the source of linguistic correctness, since nothing turns on it here). Among other things, a term's rule states its application and coapplication conditions, the former being those conditions under which it can be used at all, the latter being those conditions under which it can be used in expressions like '*x* is identical to *y*'. Restricting our attention to application conditions, a first pass at the rule for 'house' might go like this: 'house' is correctly applied just in case it is applied to a physical structure that was either constructed or modified for habitation. Of course, we need not have explicit knowledge of these rules in order to use the terms correctly; still, correctness will be defined in terms of consistency with the rules, themselves constructed, one presumes, by various psychological, social, and historical facts.

¹⁰⁷ I don't think that Thomasson would disagree. See her "Modal Normativism and the Methods of Metaphysics," *Philosophical Topics* (Forthcoming). However, my main source for what follows is (Thomasson 2007).

Descriptivists may be very hospitable to an account like this, but not so for those drawn to the theory of direct reference. In hopes of staving off objections from this quarter, Thomasson argues that even those in the latter camp need to avail themselves of something like it. Direct reference theorists need to provide some story that explains how terms are associated with the things that are their semantic values. Kripke's is that there are causal chains that connect speakers to an original baptism that establishes the connection, and something in this neighborhood is the popular view. But as many have observed, a *purely* causal theory of reference faces a deep challenge: given that there are an incredible number of candidate referents at any baptism, how can a purely causal process select from among them?¹⁰⁸ It seems that some descriptions must be invoked in order for the reference of a term to be fixed determinately. But if there are such descriptions, then competent use of the term will be in accord with them – i.e., they will be part of the application and coapplication conditions for the term. (This is *not* to suggest that, on Thomasson's version of direct reference, the meaning of the term is provided by the rules; that would be to give up direct reference altogether. Rather, the thesis is that *competent use* of the term will be in accord with the rules. Since direct reference is silent on the nature of competent use, this is surely a legitimate combination of views.) So, Thomasson claims that while the best account of the semantics of singular and general terms may not be a pure descriptivism, there should still be room for the metalinguistic rules that she needs for her account to get off the ground.

Thomasson then suggests that these rules determine what she calls 'the frame-level' application and coapplication conditions. In other words, and again focusing on application conditions, they specify that, if the term refers at all, it refers to an object (or class of objects) satisfying some more general description. (In my toy example, the frame-level application

¹⁰⁸ Michael Devitt calls this 'the qua-problem'; see (Devitt and Sterelny 1999, 65).

condition for ‘house’ is the bit requiring that it refer to a physical structure.) If terms have frame-level application and coapplication conditions, they can be used to ground ‘analytic entailments’. Analytic entailment is a metalinguistic relation that is, roughly, something like sufficiency for truth. Suppose that a sortal, ϕ , has been correctly applied to some object x , and the sortal’s associated rule states that it applies only if x also falls under some other sortal, ψ ; in that case, ‘ x is ϕ ’ analytically entails ‘ x is ψ ’ – i.e., ‘ x is ϕ ’ is sufficient for the truth of ‘ x is ψ ’. Similarly, suppose that a sortal, ϕ , applies just in case conditions $C_1 - C_n$ are satisfied; then, ‘ $C_1 - C_n$ are satisfied’ analytically entails ‘there is a ϕ ’ (or ‘ x is ϕ ’, depending on the sortal and the circumstances in question), and vice versa. Analytic entailments are also necessary entailments: So, if one knows that ‘ x is ϕ ’ analytically entails ‘ x is ψ ’, then one knows that, necessarily, if x is ϕ then x is ψ . And if one knows that ‘ x is ϕ ’ does *not* analytically entail ‘ x is ψ ’, then one knows that, possibly, x is ϕ but not ψ . Because analytic entailments are grounded in the rules that govern the uses of terms, Thomasson thinks that “competent speakers and reasoners” can know that analytic entailments hold simply by knowing the meanings of the terms involved (Thomasson 2007, 22). So, our justification for our modal beliefs consists in our linguistic competence.¹⁰⁹

As I said, analytic entailment is supposed to be a metalinguistic relationship, which is to say that it is supposed to be a relationship among propositions, and therefore only derivatively a relationship between sentences or utterances or inscriptions.¹¹⁰ Note, moreover, that if there are no determinate relationships among propositions, then there are no analytic entailments: for any two propositions Φ and Ψ , if it is indeterminate whether Φ analytically entails Ψ , then it isn’t

¹⁰⁹ Some degree of logical competence may also be required, but I won’t worry about this detail.

¹¹⁰ Not everyone would agree with this characterization of analytic entailment, since analytic entailment is a species of entailment, and not everyone thinks that entailment holds among propositions. However, most of the parties to this debate are happy to talk in terms of entailment among propositions, so I will not defend the assumption.

true that Φ analytically entails Ψ .

Of course, one's theory of propositions will determine how one spells out relationships among propositions. If one think that propositions are, say, functions from worlds to truth values – as Robert Stalnaker does – then one will have a very simple explanation of when one proposition, Φ , analytically entails another, Ψ : Φ analytically entails Ψ iff the set of worlds to which Φ assigns the value 'true' is a subset of the set of worlds to which Ψ assigns the value 'true'. But if one thinks that propositions are structured entities – as neo-Russellians think – then one will have to give some other account. The crucial point for our purposes is that different conceptions of propositions have different implications concerning the nature of relationships among propositions: On Stalnaker's view, it is easy to see how relationships among propositions might be determinate or indeterminate; the issue turns on whether every proposition is defined for every world. On neo-Russellian views, it's a bit harder to see how the details will go. If the constituents of propositions cannot be vague, then, presumably, all relationships among propositions will be determinate. However, if vague objects can be constituents, or if even senses themselves can be constituents, then perhaps some story could be told on which those relationships would not be determinate. So, if one were being careful, then one would say that skepticism about analytic entailments commits one to skepticism about determinate relationships among propositions, which commits one to certain views about the nature of propositions.

I mention all this to set it aside: I do not want to talk about analytic entailments *per se* because I do not want to get mired in issues that are, ultimately, irrelevant to my project. Instead, I want to focus on the fact that, regardless of one's view of propositions, one has to explain how sentences (at times in contexts) are correlated with the propositions that they express. It's only if those correlations are established that one can have derivative analytic

entailments. What I'll argue is that indeterminacy probably infects the meanings of the terms used to express claims about extraordinary modal matters. If so, then when a sentence contains one of those terms, it is indeterminate which proposition the sentence expresses. So, even if the relationships among propositions *are* determinate, there won't be any fact of the matter as to whether one sentence containing an indeterminate term derivatively analytically entails another.¹¹¹ And if that's right, then it will be a mug's game to base our modal knowledge on our linguistic competence, at least as Thomasson tries to do. (Note, moreover, since no one thinks that we have unmediated access to propositions, this problem will crop up for any view that bases our modal knowledge on the objects of our immediate semantic knowledge – i.e., the objects, whatever they are, that have meanings or express propositions, but are not themselves meanings or propositions.)

6.2.2 Problems for Thomasson's View *qua* Modal Epistemology

Now consider a simple case to which Thomasson appeals:

... that there is a man living in a certain house does not itself analytically entail that a bachelor lives there. But the minimal set Φ of claims, (ϕ_1) that there is a man living in that house and (ϕ_2) that that man is unmarried, does analytically entail Ψ – that there is a bachelor living in that house (Thomasson 2007, 23).

Before worrying about determinacy and indeterminacy, why might one question the idea that the truth of ϕ_1 and ϕ_2 is sufficient for the truth of Ψ ? Marcus Giaquinto offers two cases that seem to have precisely this implication:

1. Peter once married a refugee merely to save her from deportation to a country suffering civil war; after the marriage ceremony he never saw her again, but continued to live as young single men do, available for a long-term personal

¹¹¹ ... at least where the semantic value of the indeterminate term is supposed to be what explains the holding of the analytic entailment. So, for example, 'That bald man owns a triangle' analytically entails 'That bald man owns a polygon', but while 'bald man' occurs in both sentences, that term is not what explains the holding of the analytic entailment. To put the point analogically, the occurrences of 'bald man' are background conditions, not the cause.

partnership.

2. Mike [...] never married; he and his partner have been together for several decades, have two adult daughters and continue to live together as companions (Giaquinto 2008, 97).

I'm inclined to agree with Giaquinto that Peter, though married, is a bachelor; Mike, though unmarried, is not. But if Giaquinto is correct, then it's false that the truth of ϕ_1 and ϕ_2 is sufficient for the truth of Ψ ; at the very least, you need an additional proposition, ϕ_3 , to the effect that the man is not cohabiting with a woman with whom he is in a long-term, committed relationship.

First off, the obvious: just about everyone who is sympathetic to analyticity thinks that the sentence, 'Bachelors are unmarried males', is a paradigmatic instance of an analytic truth. If those philosophers are wrong about a paradigmatic and relatively simple case, as Giaquinto suggests, then we should have very little confidence that we can determine sufficient conditions for the more complex cases: e.g., 'baseball', 'water', 'pain', 'red', 'God', 'gratuitous evil'. So, even if there are analytic entailments among propositions, they may not be of much value for epistemic reasons, since we may be extraordinarily bad at ferreting them out. (The other option, of course, is that there may not *be* very many analytic entailments. If that's the case, then Thomasson's approach directs us to have an extraordinarily generous view of what's possible. I'll come back to this point.)

The way to mitigate the force of this challenge is to generate a list of the reasons why conceptual analysis is hard, the upshot being that, while it's no easy task to settle the meaning of a term, there is nothing other than the requisite effort standing between us and the knowledge analytic entailments. But this is implausible, and here is where the worry about indeterminacy comes to the fore. We can represent the two candidate meanings of 'bachelor' with two sets of

properties: $\langle \textit{maleness}, \textit{being unmarried} \rangle$ and $\langle \textit{maleness}, \textit{being unattached} \rangle$. (Assume that these are the only two candidates and that we're right about what the candidates are; this puts Thomasson's view in the best position possible. Also, note that I'm suppressing *adulthood* for ease of exposition.) Now our interlocutor must answer the following question: has the meaning of 'bachelor' changed, or has it remained constant? On the first view, one says that $\langle \textit{maleness}, \textit{being unmarried} \rangle$ once represented the meaning of 'bachelor', but now it is best represented by $\langle \textit{maleness}, \textit{being unattached} \rangle$, perhaps because the meaning of 'bachelor' was fixed by certain social structures and mores (we used to identify bachelors based on marital status, but now we do so based on lifestyle). On the second view, one says 'bachelor' was always best represented by $\langle \textit{maleness}, \textit{being unattached} \rangle$ and speakers were incorrect if they said otherwise. It was reasonable to assume that being an unmarried male was necessary and sufficient for being a bachelor when the class of unmarried males was coextensive with the class of bachelors, but after seeing that those classes are not necessarily coextensive, we revised our belief.

Our interlocutor needs to explain why one of these responses is preferable to the other. If she can't, then we're without a basis on which to choose one of the two candidate semantic values. (This underscores that it won't do to maintain that one of these responses *is* preferable to the other, though we aren't in a position to know which.) This is no easy task. Note that Thomasson appeals to apparent referential indeterminacy to motivate the idea that descriptive content is partially constitutive of the meanings of singular and general terms. It's unlikely that any of the descriptive content that might have been employed to coin 'bachelor' settled whether its meaning is best represented by $\langle \textit{maleness}, \textit{being unmarried} \rangle$ or $\langle \textit{maleness}, \textit{being unattached} \rangle$. After all, the odds are good that the distinction wasn't at the forefront of anyone's mind when the term was coined. Alternately, perhaps the issue wasn't settled then, but it's

settled now. But when did we precisify the meaning? And how? It certainly seems as if the ambiguity is still present, as evinced by the appeal of both candidate meanings of ‘bachelor’. It is far more plausible, then, that the meaning neither changed nor was precisified: the meaning was and still is indeterminate, which is to say that never was, and still is not, a fact of the matter concerning which of the two sets represents the ‘real’ meaning of the term, and which the imposter.

Why suppose that this problem afflicts the terms used to express claims about extraordinary modal matters? Well, one reason is that, if a term like ‘bachelor’ suffers from semantic indeterminacy, then surely it also afflicts terms like ‘mind’ and ‘God’. After all, one would have thought that ‘bachelor’ has determinate application conditions if any term does. That there is ambiguity in a paradigm case casts doubt on the rest.

A second reason is this. It’s plausible to think of vague predicates (‘bald’, etc.) as having indeterminate semantic values. As Thomasson points out, though:

...there seem to be cases in which there is no determinate answer as to whether or not [predicates] apply, and where this lack of an answer seems not a mere function of our lack of information, but rather seems to be ineliminable no matter how much other information we have about the case or how perfect our epistemic situation is. [Moreover,] *most if not all expressions of natural language are vague*, and the vagueness of natural language expressions also threatens to infect scientific or logical terms insofar as they are defined even in part on that basis (Thomasson 2007, 90; emphasis mine).

The problem crops up as a result of the italicized clause. Let’s suppose that natural languages are indeed shot through with vague predicates. As with ‘bachelor’, the candidate semantic values for vague predicates may not be disjoint.¹¹² So, if some property ϕ is shared by all the candidate semantic values for some predicate ψ , then ‘ x is ψ ’ analytically entails ‘ x is ϕ ’. I

¹¹² Some contest this, arguing that the candidate semantic values usually *are* disjoint (or, more carefully, that there usually are pairs of disjoint candidate semantic values for a given term); see the arguments in (LaPorte 2004). Those arguments aside, I tend to think that Giaquino’s cases make it easy enough to construct analogous ones about transgendered individuals; the candidate semantic values will, therefore, be disjoint (something like: < *maleness, being unmarried* > and < *masculinity, being unattached* >).

would wager, though, that the common properties (or common property, in some cases) are likely to be very general. And if it works out that ‘*x* is a dog’ and ‘*x* is building’ only analytically entail sentences with very general predicates – e.g., ‘*x* is a physical object’ – then this is bad news for Thomasson’s view, since analytic entailments will not be fine-grained enough to help us sort out extraordinary modal claims. And given how slippery meanings are, I see no reason to think that analytic entailments will underwrite inferences involving richer information.

The third reason, and probably the most important one, is that terms are inventions. Consequently, we have little reason to suppose that their meanings are significantly more precise than are required by the ends for which we use them. But extraordinary modal matters are, *ex hypothesi*, well beyond our ordinary interests; they concern questions that do not arise in the workaday world. It follows from this alone that there will be very few interesting analytic entailments about extraordinary matters. Should we conclude, therefore, that the answer to every question about an extraordinary modal matter is, ‘Yes, that’s possible!’? No. The explanation just given for the lack of analytic entailments is better than the hypothesis that they track the truth. So, an advocate of the semantic approach should simply abstain from belief when it comes to extraordinary matters: her theory is poorly suited to those topics, and so she should have no opinions about them. The advocate of the semantic approach can soften this blow by arguing that our ordinary interests are quite expansive, but this is a desperate move. We are short-sighted creatures with narrow focuses. Even when the topic is close to home, we can’t be bothered with much that goes beyond the practical. To take just one example, because we’d like to experience less pain, we care about its neurological basis. Suppose that we were to discover that, by calming those poor C-fibers, we could manage pain in all the circumstances that concern us.

With this information in hand, would our interest in neurology give us a reason to settle questions about the precise *formal* relationship between painful experiences and C-fibers firing? Would it matter whether they were identical, or merely stood in one or another supervenience relation? I doubt it. Of course, we philosophers might come up with a *theory* about their relationship, and it might even be one that we're justified in believing. However, it will then be the theory that informs our modal inquiries, and not the meanings of the phrases 'painful experiences' and 'C-fibers firing'.

So, while it may seem as if Thomasson's view can underwrite knowledge of (G), this appearance is deceiving. I'm not sure whether there are sufficiently determinate semantic facts to underwrite knowledge of ordinary modal matters, but we needn't settle this question. Even if so, there is every reason to suppose that indeterminacy infects the terms used to express claims about extraordinary modal matters, and that's all that's needed to show that a semantic approach can't secure (G). Therefore, we can grant that '*x* is God' doesn't analytically entail '*x* creates a world in which human beings suffer gratuitous evils'. However, it's probably also true that there are no analytic entailments between the following propositions (presented here as schemas):

- '*x* is an unmarried male' and '*x* is a bachelor';
- '*x* is a sample of iron' and '*x* is transparent';
- '*x* is a sample of water' and '*x* is H₂O';
- '*x* is a pain state' and '*x* is a physical state';
- '*x* is red all over' and '*x* is green all over'; and
- '*x* is Bob Fischer' and '*x* was born to John and Shelley Fischer'.

The semantic approach, then, is a poor foundation for extraordinary modal knowledge. Hence, the fact that it underwrites (G) should be cold comfort to the Optimist.

6.3 Yablo

The second proposal is that if we can conceive that p , then we are justified in believing that p is possible.¹¹³ (Call this ‘the conceivability thesis’.) Steven Yablo has the most influential account of this kind. He contends that conceivability is a guide to possibility because, on his view, genuine conceiving is analogous to genuine perceiving: “Just as to perceive that p is to be in a state that (i) is veridical only if p , and that (ii) moves you to believe that p , to find p conceivable is to be in a state which (i) is veridical only if possibly p , and (ii) moves you to believe that p is possible” (Yablo 1993, 7). The idea is this. Perception presents itself as a guide to its domain. It is reasonable to believe that things are as they seem in the absence of defeaters. Therefore, it is reasonable to believe that things are as one perceives them. But conceiving also presents itself as guide to its domain. So, it is reasonable to believe that things could be as one conceives them.

There are two oft-cited problems for the slogan, ‘conceivability is a guide to possibility’. The first is that there seem to be cases in which we can conceive that p and that $\sim p$, though p is necessary if possible and $\sim p$ is necessary if possible: e.g., we can conceive that Goldbach’s Conjecture is true and we can conceive that it’s false. Second, there seem to be cases in which we can conceive that p despite the fact that we take it to be impossible that p : e.g., we can conceive of Jones being born to different parents, though we might not think that this state of affairs is possible. No doubt there is some sense of ‘conceive’ such that, if the conceivability thesis were to employ it, these would be genuine problems. I can conceive that Goldbach’s conjecture is true in the sense that I can conceive of reading the headline, “Mathematicians say: Goldbach was right!” – and I can conceive of going to the conference and hearing the relevant

¹¹³ A close cousin of this view: if we can *sensuously imagine* that p , then we are justified in believing that p is possible. I’ll discuss this alternative shortly.

authorities agree that the proof is sound. This is one of the senses of ‘conceive’ that worries van Inwagen, and of course it isn’t of much epistemological value: in just the same sense of ‘conceive’, I can conceive that Goldbach’s conjecture is false. Indeed, Yablo identifies five epistemologically-worthless senses of conceiving, each of which would make the conceivability thesis subject to counterexample:

1. p is *conceivable_b* iff it is (not un)believable that p .
2. p is *conceivable_{bp}* iff it is (not un)believable that *possibly*, p .
3. p is *conceivable_{ijb}* iff one can imagine justifiably believing that p .
4. p is *conceivable_{itb}* iff one can imagine believing p truly.
5. p is *conceivable_{ep}* iff one can imagine believing something true with one’s actual p -thought (Yablo 1993, 26).

Yablo contends, though, that none of these varieties presents itself as a guide to possibility; consequently, each fails to be sufficiently like perception. However, he thinks there is another kind of conceiving that is appropriately correlated with judgments of possibility. He explains this variety of conceivability – and its foil, inconceivability – as follows:

(CON) [p is conceivable for me if] I can imagine a world that I take to verify p .

(INC) [p is inconceivable for me if] I cannot imagine any world that I don’t take to falsify p (i.e., for every world that I can imagine, I take that world not to verify p) (Yablo 1993, 29).

(CON) and (INC) are supposed to help us understand why conceiving and failing to conceive should be correlated with the appearance of possibility or impossibility. If the correlation holds, then the analogy with perception can be maintained. Let’s focus on (CON). According to (CON), if p is conceivable for me, then I can imagine a world that I take to verify p . Yablo points out that imagining can be either objectual or propositional. Propositional imagining is imagining *that*; it is *de dicto* imagining, in the sense that one imagines that some proposition is true. Objectual imagining, by contrast, is imagining *de re*; it is imagining some *thing*. Yablo contends that objectual imagining is typically correlated with the impression that the object could

exist. And in both (CON) and (INC), he intends ‘imagine’ be taken in the objectual sense. (Yablo denies that the difference between objectual and propositional imagining is to be cashed out in terms of the use of mental images. “Some philosophers use ‘imagine’ so that imagining a thing is *imaging* it, that is, conjuring up an appropriate sensory presentation. I do not require a sensory-like image for imagining, and certainly not a distinct such image for distinct imaginings” (Yablo 1993, 27 n.55; emphasis in original).)¹¹⁴ The upshot is this. If p is conceivable for me, then it seems to me that I can imagine a world that verifies p . If there exists a world that verifies p – i.e., if there is a world of which p is true – then p is possible. We can maintain, therefore, that objectual imagining – and hence conceiving – is correlated with the appearance of possibility; if you objectually imagine that p , then it seems to you that the truthmaker for ‘ p is possible’ exists. So, if it seems to me I can imagine a world that I take to verify p , then I am justified in believing that there is such a world, and hence that p is possible. Therefore, if I can conceive that p , then I am justified in believing that p is possible (in the absence of defeaters).

There are various objections that have been leveled against Yablo’s view. Some have argued that it relies on antecedent modal knowledge, such as some sort of Lewis-style principle of recombination (O’Connor 2008, 32-36). Others have claimed that it can’t handle negative existentials or temporally-indexed facts (Sosa 2000, 1-5; Geirsson 2005, 292-293). Still others deny that we can imagine a world in enough detail for the appearance of possibility to warrant belief (Oreste Fiocco 2007, 371-373; van Inwagen 1998, 76-81). Each of these objections has some merit, but none, in my view, is devastating to Yablo’s project. If one grants the initial parallel between conceiving and perceiving, then each can be addressed. (Or so it seems to me.

¹¹⁴ Yablo rejects imagining-as-imaging because he thinks that we can’t image all the propositions that strike us as being possible; he seems to be convinced by Descartes’ argument to this effect based on the difficulty of picturing a chilliagon. At any rate, it follows from this that although you can use images to objectually imagine a world, you needn’t; hence, images play no essential role in the justification of your modal beliefs.

I think that the argument in Chapter 4, §4 can probably be extended to handle these problems.) Nevertheless, I deny that Yablo's view is viable. Yablo's view relies on the principle of credulity (also called 'phenomenal conservatism'): i.e., the view that if it seems to you that p , then you are justified in believing p , at least in the absence of defeaters. I do not contest the claim that this principle applies in the case of perception, so the burden on Yablo is just to show that conceiving is sufficiently like perceiving to justify extending the principle. However, he also needs to tell a story about how we identify the things that are the objects of our imaginings, and I cannot see how he will do this without seriously damaging the analogy between imagining and perceiving. If that analogy is broken, though, then the account flounders.

6.3.1 Imagining without Images

Suppose – as Yablo clearly does – that imagining is not an essentially sensuous mental process (i.e., that you can imagine without 'imaging'). Now recall those famous debates over 'transworld identity' – the claim that individuals can exist in more than one possible world. Those who weren't sympathetic to transworld identity often raised the worry that if Nixon, say, were to exist in another possible world, we'd have no way of *knowing* that it was Nixon. Since we can only identify him by his properties, and, *ex hypothesi*, there is no one in the world in question with precisely the same set of properties,¹¹⁵ we lack the means with which to pick him out. Kripke denied that this is a real problem; coining a now-famous phrase, he maintained that we do not look at worlds through a telescope. Rather, we *stipulate* that the man about whom we're speaking is Nixon, the reference of that name already being fixed in the actual world, and

¹¹⁵ This is true in two senses. First, when considering counterfactuals about Nixon, it's true that no one in the world in question has the same properties as does the actual Nixon (for the simple reason that we are considering contrary-to-fact scenarios *about* Nixon). Second, if you countenance relational properties and reject duplicate worlds, then you'll take it to be a general truth that no one in any world has precisely the same properties as anyone in any other world.

the identity between Nixon-in- $w_@$ and Nixon-in- w_1 doing the rest of the work.

Kripke is surely right that, *if* there is transworld identity, then there is no problem about reference across worlds. But this is not of much interest for epistemological purposes. If we're in a position to *stipulate* that the man in the situation is Nixon, then surely that's because we already *know* that he is. And, if we don't know that he's Nixon, then it would be folly to claim as much without argument. After all, stipulation seems to be on an epistemic par with supposition, and supposition is clearly no guide to possibility. We often suppose the truth of necessarily false propositions, as we do when considering *reductio* arguments.¹¹⁶ And if we have no trouble supposing the truth of a necessary false proposition, then it's implausible that supposition carries with it the appearance of possibility; it seems much more likely that supposition is neutral with respect to the appearance of possibility, and is merely a way to investigate the implications of a given proposition. (It's also no good to say that someone is only allowed to run with a supposition until it encounters trouble, unless Yablo provides a richer theory of 'being in trouble' than 'is contradictory'. As discussed in the last chapter, consistency is no guide to possibility.)

So Yablo shouldn't say that we stipulate the contents of our imaginings. What can he say instead? To answer this question, let's consider Jones' attempt to imagine a world that verifies a proposition about the Queen Mother, and let's grant that he doesn't get to say, from the outset, that the world he's imagining contains the Queen Mother. Remember too that there needn't be any images involved in Jones' imagining, so the identification isn't (necessarily) based on the scene's being Queen-Mother-ish. To get what Yablo wants, we need an account that doesn't rely on the *sensuous* imagination. I don't know where to go from here. To my ear, Jones'

¹¹⁶ To my knowledge, this observation first appears in Thomas Reid's *Essays on the Intellectual Powers of Man*, Essay IV, Chapter 1, §IV.1.

‘imagining’ sounds a lot like intuiting: Jones considers a proposition about the Queen Mother and it simply strikes him as being true.¹¹⁷ But if this is the view, then it won’t do for Yablo’s purposes. What we want is an explanation of conceiving in terms of objectual imagining – i.e., in terms of a *de re* mental state. What we seem to have is an explanation of conceiving in terms of a *de dicto* mental state. In other words, what Yablo promised us was an account on which the (intentional) objects of conceiving are (concrete) objects and properties in certain configurations; what we have – it seems – is an account on which the (intentional) objects of conceiving are propositions. The former (arguably) preserves the parallel with perception, but the latter does not: whatever else may be true about perceiving, it’s very implausible that we perceive propositions.

Of course, I haven’t shown that intuiting cannot plausibly be construed as a *de re* mental state.¹¹⁸ And perhaps it can be so construed. But even if so, I doubt that the parallel with perceiving is tight enough for Yablo’s purposes. Let me try to explain why.

The view under consideration is one on which (a) we somehow attend to a certain configuration of objects and properties and (b) this attending justifies belief that the configuration is possible. The move from (a) to (b) is supposed to be warranted because attending to the configuration makes it seem to us that the configuration is possible – not just that it isn’t *impossible*. I’d like an explanation for the appearance of possibility. What, exactly, does Yablo think that these appearances track? To what feature of our imaginings are we reacting? To my mind, if Yablo wants to explain the appearance in a way that preserves the epistemic credentials of imagining, then the explanation should appeal to our understanding of – i.e., our available information about – the objects and properties in question. I am not suggesting that the

¹¹⁷ ...or possibly true, depending on whether the proposition is already ‘modalized’.

¹¹⁸ I think that Anand Vaidya means to defend a view on which it *is* a *de re* mental state, although I suspect that he would resist the language of ‘intuition’. See (Vaidya 2010).

explanation be in terms of antecedent modal knowledge; clearly, it's unfair to Yablo to demand that he explain the appearance of possibility in terms of, say, our knowledge of objects' essential properties. However, it does seem fair to ask for some explanation or other. (He might appeal to our recognition of some properties as characteristic of the object, or as necessary for all explanatory purposes, or kind-defining, or what have you.)

Still, when it comes to the issues raised by extraordinary propositions, I doubt that we have the information that will figure in this explanation. The case is hard to make absent a positive proposal to evaluate, but I think that I can finesse this issue. Let's inquire about the *source* of our knowledge of objects. Aristotle, for example, has a theory of perception on which the form of an object literally impresses itself on the perceiver:

That which can perceive is, as we have said, potentially such as the object of perception already is actually. It is not like the object, then, when it is being affected by it, but once it has been affected it becomes like it and is such as it is.¹¹⁹

On his view, the eye receives (is 'enformed' by) the form of the distal object – i.e., the eye actually instantiates the perceived properties.¹²⁰ And since the form of an object is what determines its modal profile, there is no mystery about how we can discern the modal properties of the objects that we have encountered. When we perceive, we take in more information than we consciously process (since we attend to only some aspects of the impression due to the form), and some of that information is already modal (or has clearly modal implications). However, I take it that no modern theory of perception works as Aristotle's does. If we have replaced Aristotle's view, we should wonder whether ours as useful as his – at least for the purposes of modal epistemology. If it isn't, then we should doubt whether we have the information needed

¹¹⁹ *De anima*, Book II, Chapter 5, 418a3-6.

¹²⁰ There is debate about this point; for details, see Christopher Shields excellent *Stanford Encyclopedia of Philosophy* article, "Aristotle's Psychology." However, the exegetical questions can be set aside here, since the morale that I want to draw from Aristotle's view doesn't require deciding between the various interpretations of his position.

to explain the appearance of possibility. And if not, then surely the view is in trouble, since it makes a mystery of appearances of possibility, and hence of the knowledge that these appearances (allegedly) underwrite.

The way to dodge this problem is to take *de re* intuition as epistemically primitive; a theory can't be faulted for failing to explain what it takes as fundamental. I think that this is a recipe for disaster. The reason for this is simple. If Yablo maintains that *de re* intuition is epistemically primitive – and thereby avoids having to explain what our appearances track – then we will be left without a way to choose between competing modal claims. Suppose that we disagree with Smith about whether there could be transparent iron. On what basis should trust our judgment over his? We could say, “Because we had our intuition, not Smith’s”. If so, then we will want to say either (a) why Smith is not an epistemic peer or (b) why, even though Smith *is* an epistemic peer, we should discount his opinion on the matter; then, we can properly maintain that his disagreement does not provide us with a defeater for our claim. But (a) is not an option, since (presumably) we *share* the faculty of intuition (or perhaps less contentiously, the ability to intuit), and because it is an epistemic primitive, we have no basis to say that his works less well than ours. Nor is (b) an option, since *de re* intuition is not plausibly construed as an analyzable process, some part of which can break down. After all, Smith didn't *reason* to his claim; there is no calculation error to which we can appeal. Smith – like us – was simply left with an impression upon considering a configuration of objects and properties. And, if we have no basis for preferring our intuition over Smith's, then it's hubris to trust our intuition over Smith's, a groundless bias if ever there were one. Call this ‘the Disagreement Problem’.

We could solve the Disagreement Problem by saying that we prefer our intuition because it coheres better with our beliefs, or has certain explanatory virtues, or whatever. But if the

justificatory work is being done by these sorts of considerations, then it's unclear why we should take *de re* intuition as an epistemic primitive, as it now seems like an unnecessary and dramatic solution to the problem of modal knowledge. So it seems to me that, if we take *de re* intuition as an epistemic primitive, then we are without a solution to the Disagreement Problem (or, more carefully, we are without a solution that does not bring into question the wisdom of thinking that *de re* intuition is indeed an epistemic primitive). Now, one could take this to be a general problem for a modal epistemology based on *de re* intuition. Perhaps that argument can be made to work, but I won't pursue it here. I will content myself with this point. There is plenty of disagreement about extraordinary modal claims, not the least of which concerns propositions like (G). So, whatever the merits of a modal epistemology based on *de re* intuition, it is implausible that it is a theory that can give us knowledge of extraordinary propositions. And that, of course, is enough for my purposes.

Let's take stock. At least where extraordinary modal claims are concerned, it does not appear as though conceiving can do the work that Yablo needs it to do. The problems stem from the difficulty of knowing what you've conceived. Understood one way, conceiving relies on stipulation; understood another, conceiving reduces to some form of intuition, and whatever the form, it's far from clear that it can underwrite extraordinary modal knowledge. So, the Optimist cannot plausibly rely on Yablo's modal epistemology.

6.3.2 Imagining with Images

However, we can give up one of the assumptions that leads to these difficulties – namely, that imagining is *not* an essentially sensuous mental process. What if we suppose instead that imagining *is* essentially sensuous? It's unclear how this will help underwrite knowledge of a

proposition like (G), since I'm not sure how to sensuously imagine God. But perhaps this strategy will help us with other extraordinary modal propositions – e.g., there could be a three-inch thick sheet of transparent iron – and that would be success enough.

The view that imagining is a guide to possibility, where imagining is construed as an essentially sensuous process – is now being called ‘modal empiricism’, and it’s been defended most recently by Rebecca Hanrahan.¹²¹ Those who endorse this view try to salvage the parallel with perception by claiming that it is your ability to *sensuously* imagine that *p* that justifies your belief that *p* is possible, not your ability to conceive that *p*. Modal empiricism avoids some of the worries about stipulation that plague Yablo’s theory. Instead of imagining a tiger (i.e., *stipulating* that you are imagining a tiger, or doing something on par with stipulating it), you imagine a suitably rich scene and try to determine the best explanation of your imagined experience. Could there be a tiger with green stripes? Imagine a scene and ask whether, if confronted with it, the best explanation would be that there is a tiger with green stripes. If that *is* the best explanation, then you’re justified in believing that there could be such a beast.

You might worry that this is no improvement on Yablo’s account: we inevitably make *some* assumptions about what we are imagining (or we do something that is no better, epistemically speaking), and hence the explanatory inference will always be corrupted. This is probably too quick. No doubt stipulation is involved in generating mental images, but this is compatible with being able to step back and regard those images as data. So, the recipe for calling up the image may look like it begs the question: “imagine a transparent bar of iron.” But this isn’t problematic if it’s used just as a recipe for producing the appropriate mental image. Then, you can consider your experience not *as of a bar of transparent iron*, but just *as of this*

¹²¹ See (Hanrahan 2007). W. D. Hart defends a similar view, although I hesitate to attribute this motivation to him. For the essentials of his version of modal empiricism, see the second and third chapters of (Hart 1988).

transparent and very heavy substance. The explanatory inference enters at this point to evaluate the imagined scene, thereby avoiding one of the difficulties that Yablo's account faces.¹²²

Still, even if modal empiricism improves on Yablo's view, it doesn't improve enough: it only makes you slightly more passive with respect to the way that things seem to you, and not nearly as passive as you are in perception. According to modal empiricism, your modal beliefs are justified because you're justified in believing that things could be as they seem in the imagination. True, an explanatory inference is crucial to determining how best to interpret a given seeming, and when making that inference, you treat the imagined scene as given. This gives the impression of passivity, but it's still the case that you're fully responsible for how the scene appears in the imagination. An expert on vision once told me that you can produce the appearance of bursts of light by putting pressure on certain parts of the eye. If you try this on yourself, then you certainly shouldn't believe, on the basis of your experience, that there is carefully-hidden someone pestering you with a flashlight – even if that would be the best explanation of bursts of light in the relevant region of your visual field. Why not? The obvious reason is *that you fabricated the experience*. And, because you fabricated it, you shouldn't suppose that its significance can be read off its apparent content. In other words, you shouldn't suppose that the way things *seem* is a good guide to way things *are*. You should take instead a more detached view, considering the best explanation of your *having* the experience, and not the best explanation of the content, since it doesn't enjoy the *prima facie* trustworthiness that attaches to normal experience. I see no reason to think that the imagination is on surer epistemic

¹²² Incidentally, this also makes it easier to see that modal empiricism is not a version of TT. TT bases your justification on the epistemic credentials of your theories; modal empiricism rests it on the principle of credulity. Essentially, modal empiricism says that the principle of credulity entitles you to believe that you could have perceptual evidence e_p , and IBE comes in to justify interpreting e_p in one way or another. However, the core of modal empiricism is the means of securing your justification to believe that you could have e_p , and there explanatory inference plays no role.

footing, and I see no reason to think that p 's possibility is likely to be the best explanation of your ability to imagine that p .¹²³ Hence, I doubt that modal empiricism has any better claim to preserving the analogy with perception than does Yablo's conceivability account, and, since it too stands or falls with that analogy, it's no better off.

6.4 Williamson

To begin, I should make it clear that it is unfair to talk about 'Williamson's modal epistemology', as if he intended to offer anything so general: the account that he develops is keyed to our knowledge of *metaphysical* modality, not modality *tout court*, and so he assumes that we have some modal knowledge in order to secure the rest. However, since he doesn't assume that we have knowledge of extraordinary modal propositions, but intends to show how we could, there's no barrier to evaluating his project on this score.

Williamson's story is straightforward. First, he notes that counterfactuals play an important role in our cognitive economy, and that we're fairly good at assessing their truth values. That this is so should be expected:

¹²³ The standard move here is to make an evolutionary argument. *True beliefs about what's possible would have enhanced the fitness of our ancestors. So, evolutionary forces probably selected a mechanism that produces a preponderance of true beliefs about modal matters. Our beliefs about what's possible are produced by the imagination. So, the imagination probably produces a preponderance of true beliefs about modal matters.* I am very skeptical of arguments that have this form. First, evolutionary forces are probably blind to belief contents, so we've got no reason to suppose that those forces will select mechanisms that produce mainly true beliefs. Second, even if I'm wrong about this, it remains the case that evolutionary forces (a) are constrained by time and the available raw materials and (b) have many concerns other than producing true beliefs, even if true beliefs they produce. So, given that evolutionary forces will be happy to trade truth for another benefit if the opportunity presents itself, what reason do we have to think that they've operated long enough in conditions favorable to selecting reliable mental mechanisms? (Granted, we have reason to suppose that our mental mechanisms are superior to those had by our competitors, but for all we know, that's a backhanded compliment.) Finally, if evolutionary forces select mechanisms that produce true beliefs, and if they've been operating in the right conditions for enough time, then we still ought to wonder about the *scope* of the mechanisms that evolutionary forces are likely to select. Are they reliable when it comes to matters removed from the business of everyday life? Are they likely to produce true beliefs about highly theoretical matters? If not – and why suppose they are? – then this reply won't help the Optimist, and so won't damage the larger argument in the main text. For more on these objections to this kind of evolutionary argument, see (Stich 1990) and Plantinga's contributions to (Plantinga and Tooley 2008).

Our overall capacity for somewhat reliable thought about counterfactual possibilities is hardly surprising, for we cannot know in advance exactly which possibilities are or will be actual. We need to make contingency plans. In practice, the only way for us to be cognitively equipped to deal with the actual is by being cognitively equipped to deal with a wide variety of contingencies, most of them counterfactual (Williamson 2007, 137).

Williamson then points out that you can define necessity and possibility in terms of the counterfactual conditional: a proposition is necessary if and only if, were its negation to be the case, then a contradiction would obtain; a proposition is possible if and only if it *isn't* the case that, if it were so, then a contradiction would be true. In short:

$$\Box A \equiv (\sim A \Box \rightarrow \bot)$$

$$\Diamond A \equiv \sim(A \Box \rightarrow \bot)$$

Williamson is then free to treat the epistemology of metaphysical modality as a special case of the epistemology of counterfactuals. Typically, we evaluate ordinary counterfactuals in the imagination. Suppose that I left a book in my son's room, where he is now sleeping. I try to determine whether, if I were to sneak in, I would be able to retrieve the book without waking him. I develop the antecedent of this counterfactual in my imagination: my mind turns to the squeaky floorboards, the locations of which I have never been good at remembering, and I recall past failures. The prospects of success look dim, and I judge that I would not retrieve the book without drama. Williamson's claim is that roughly the same process can provide insights about metaphysical modality. The aim there, however, is to develop the antecedent in the interest of ferreting out contradictions, if any there be. As he puts it,

...we assert $\Box A$ when our counterfactual development of the supposition $\sim A$ robustly yields a contradiction; we deny $\Box A$ when our counterfactual development of $\sim A$ does not robustly yield a contradiction (and we do not attribute the failure to a defect in our search). Similarly, [...] we assert $\Diamond A$ when our counterfactual development of the supposition A does not robustly yield a contradiction (and we do not attribute the failure to a defect in our search); we deny $\Diamond A$ when our counterfactual development of A robustly yields a contradiction. Thus our fallible imaginative evaluation of counterfactuals has a conceivability test for possibility and an inconceivability test for

impossibility built in as fallible special cases (Williamson 2007, 163).

Like Yablo, Williamson uses the imagination to underwrite conceivability as a test for possibility. Williamson, however, makes no appeal to the principle of credulity, and he doesn't seem to think that there is anything special about the imagination other than its being the means by which we assess counterfactuals. On his view, the heavy lifting seems to be done by our reliability as counterfactual reasoners. (It's unclear to me whether he takes belief in that reliability to be justified by evolutionary considerations, though he obviously hints in that direction.) The rest of the work is done by the logical equivalences – meant, of course, to show that the capacities that make us reliable about ordinary matters will also allow us to be reliable about matters further removed from the business of everyday life.

C. S. Jenkins complains that this is all too quick: perhaps we *could* come by our modal knowledge this way, but that doesn't show that we do; Williamson needs more than the logical equivalences (plus a general appeal to theoretical economy, heretofore unmentioned) to show that he's identified *our* source, and not just a potential one (Jenkins 2008). Yet Jenkins is confused: it may well be that we don't often perform the imaginative exercises to which Williamson alludes, but no matter, since his theory doesn't predict otherwise. Roughly, what Williamson's theory predicts is this: if we justifiably believe that *p* is necessary, then, were we to consider how things would be if *p* were false, we would detect a contradiction; analogously, if we justifiably believe that *p* is possible, then, were we to consider how things would be if *p* were true, we would detect no contradiction, nor would we see anything wrong with our search for one.¹²⁴ Williamson's theory is plausible insofar as (a) these predictions are accurate and (b) their accuracy explains the justification that attaches to the relevant beliefs. He need not offer a

¹²⁴ I'm not completely confident whether the language of 'justified belief' is most appropriate here; Williamson would probably prefer 'knowledge'. The important point for my purposes, however, is the structure of the prediction, not the precise epistemic vocabulary.

theory of the genesis of our modal beliefs in order to offer a theory of their justification.

I think that these predictions probably *are* accurate, but I deny that their accuracy explains the justification that attaches to the relevant beliefs. I think that they are accurate because Williamson helps himself to some unspoken assumptions. Moreover, I think that these assumptions do the heavy lifting in Williamson's epistemology – not the counterfactual framework that he develops.

We can see this by revisiting some familiar terrain. Kripke convinced most of us that there are necessary truths that can only be known *a posteriori*: necessarily, water is H₂O, heat is mean molecular kinetic energy, lightning is electrical discharge, and so on.¹²⁵ Williamson grants that there are such truths, but he contends that they pose no problem for his epistemology; they simply require that we be appropriately well-informed:

If we know enough chemistry, our counterfactual development of the supposition that gold is [not] the element with atomic number 79 will generate a contradiction. The reason is not simply that we know that gold is the element with atomic number 79, for we can and must vary some items of our knowledge under counterfactual suppositions. Rather, part of the general way we develop counterfactual suppositions is to hold such constitutive facts fixed (Williamson 2007, 164).

Williamson's reasoning here is confusing, but before explaining why, here is an observation. Essentially, Williamson is contending that *noseeum* inferences underwrite possibility claims: we don't see any contradictions, so they ain't there. Now, as he fully appreciates, the quality of a *noseeum* inference is only as good as the degree to which our abilities fit the subject matter at hand; hence, he stresses the importance of having the relevant chemical knowledge. But ours is a world with a serious division of cognitive labor; unless you're a specialist on *x*, then you

¹²⁵ You can resist Kripke's conclusion by denying that there are genuine proper names on both sides of the identity sign, maintaining instead that at least one of them is a definite description in disguise; then, you can make a scope distinction to avoid having the necessity attach to the properties ascribed by the definite description. This shows that you can resist the conclusion; it does not show that you should. Whether you should turns on how plausible it is that we use certain terms – e.g., 'H₂O' – as proper names. If we do, then we're committed to the necessity; if we don't, then we aren't. Most of us think that we do. You might ask whether we ought to use those terms as proper names. I don't know how to answer this question.

probably have little more than rudimentary knowledge of the ‘constitutive facts’ about x . One unfortunate thing about having rudimentary knowledge is that you often don’t know what inferences to make based on the knowledge you’ve got: I know, for example, that water is H_2O , but I know very little about the implications of this fact for water’s behavior in various circumstances. So, if Williamson’s account is right, then I’m probably not in a position to make many interesting modal claims about water, or about anything else beyond my expertise; I ought to limit myself to inferences from actuality to possibility, plus claims licensed by testimony from experts. And this is precisely what modal skepticism says. So, even if there aren’t any problems with Williamson’s account, it’s still a cautious modal epistemology: whatever it says about (G), there are going to be lots of other extraordinary propositions about which it recommends silence. For example, since I’m not up on my metallurgy, I couldn’t use Williamson’s account to say whether there could be transparent iron. So, this account won’t help the Optimist to challenge modal skepticism, and therefore poses no threat to TT.

But there are problems with Williamson’s account. I think that it’s either (a) incomplete or (b) too permissive. Consider this: why can’t we imagine how things would be if gold were the element with atomic number 78? It certainly *seems* as though we can ‘hold fixed’ the object of our imagining – in this case, gold – and this whether or not we also hold fixed its being the element with atomic number 79. Williamson blocks this by saying that we’re supposed to hold *constitutive* facts fixed. But why, exactly? If we know that we should, then *of course* we’ll be able to derive contradictions, since the antecedent will always (implicitly or explicitly) contradict the consequent: e.g., ‘if gold – i.e., the element with atomic number 79 – were the element with atomic number 78, then a contradiction would be true’. But then it’s our knowledge of this obligation that’s of interest to the modal epistemologist, and not the machinations in the

counterfactual imagination. How do we come by this knowledge? Is it *a priori*? Does it fall out of our concept of a thing that has a constitution? Is it gleaned from some abductive argument? In the absence of answers to these questions, Williamson's doesn't actually have a modal epistemology for extraordinary modal claims; he just has a schematic that could be fleshed out in any number of ways. Alternately, suppose that we *aren't* required to hold such constitutive facts fixed. Then whence the contradiction? Now it looks like his account is far too permissive: what *won't* we judge to be possible under this interpretation? (Note that it won't do to appeal here to the necessity of identity, at least if that principle is motivated by the same intuitions that lead Williamson to say that we should hold constitutive facts fixed. Since those intuitions are in question, an argument for their veracity is wanted before we take them seriously.)

By my estimation, Williamson assumes just what needs to be shown – namely, that certain properties are essential (and so should be held fixed) while others can be varied. TT explains why this should be so: the essential properties are the ones that are invariant in the theory of the target system; the contingent ones are the rest (at least by the lights of the theory). Absent an analogous account, what Williamson provides is a schema that could be filled out in a number of ways. Indeed, you might even regard TT as one such way. But whether or not you take this line, it remains that Williamson hasn't said enough to underwrite knowledge of extraordinary modal claims. In particular, he hasn't given us enough to underwrite knowledge of (G), since we have no guidance as to what we should hold fixed when trying to decide whether the following proposition is true:

(G*) If God were to create a world *sans* gratuitous evil, then a contradiction would be true.

Should we hold fixed facts about God's attributes? Some or all? If only some, then which ones? Should the laws of nature be held fixed? Some or all? If only some, then which ones? I see no

way to answer these questions based on what Williamson says, and no way to evaluate the counterfactual without those answers.

6.5 Conclusion

This concludes my argument to the effect that three major modal epistemologies cannot underwrite knowledge of propositions like (G). And TT, of course, says that we aren't justified in believing them. This provides some evidence that propositions like (G) are beyond our ken. But if this is so, then why do so many of us believe otherwise? I think that TT can shed some light on this question. The answer, in short, is that an extraordinary proposition can sometimes look to be backed by a theory when either (a) there is no theory behind it at all or (b) there is a theory in the neighborhood, but it is not one that we're actually justified in believing – just one with which we're familiar that is easy to use.¹²⁶ If TT is true, then the truth of (a) or (b) explains the appearance of justification, despite its absence. I'm not aware of any other epistemology that ties our modal knowledge to our theories in a way that allows for the same explanation.

The Optimist's Objection is this. Modal skepticism says that we don't know the truth value of any extraordinary proposition, but we normal folk know the truth value of extraordinary proposition *p*; hence, modal skepticism is false. Whatever entails a falsehood is false, and TT entails modal skepticism. So, TT is false.

The aim of this chapter has been to turn back the Optimist's Objection. To do this, canvassed three prominent modal epistemologies, arguing that none provides a plausible grounding for knowledge of extraordinary propositions. I also argued that each suffers from

¹²⁶ I think that (b) is very common. Suppose, for example, that Jones proposes a theory of the mind (e.g., a version of functionalism) according to which it's possible to realize a mind in some system other than the one that first realized it. But then Jones is then presented with a battery of devastating arguments against his theory of the mind, and he abandons it. Nevertheless, he hangs on to the modal claims that a mind could have a different realizer, despite the fact that his only basis for this claim has been lost.

some serious problems. The upshot of all this, I submit, is this: it's implausible that we have the knowledge that the Optimist attributes to us; so, it's no shortcoming of TT that it disagrees with the Optimist about the scope of our knowledge. From this encounter, anyway, TT emerges unscathed.

7. CONCLUSION

7.1 The Case for TT

TT, you'll surely recall, goes like this. Where p is any modal claim:

[TT] You justifiably believe that p iff (a) you justifiably believe a theory according to which p is true, (b) you believe that p on the basis of this theory, and (c) you have no defeaters for the belief that p .

If we are realists about modality, should we believe TT? I am very pessimistic about anyone's ability – not least mine – to lay out an argument that can coerce belief in a positive philosophical thesis.¹²⁷ However, let's see what can be said on its behalf.

7.2 Modal Skepticism and the World's Rules

We can extract one argument for TT from the analogy with *Clue* that I developed in Chapter 1. There, I pointed out that the rules of *Clue* fix what's possible in the game. So, knowing the rules is crucial to assessing *Clue*'s modal facts. After all, if you don't know any of the rules, then you have no grip on the factors that determine what's possible; and if you have no grip on the factors that determine what's possible, then surely you are just *guessing* when you make judgments about what's possible in *Clue*. I am not (epistemically) risk averse, but surely *this* level of risk is unacceptable. Likewise, the world has rules: logical, mathematical, natural, and metaphysical (and for all I know, others too). These rules are the world's invariant features, and they determine possibility spaces. One of the spaces – in my view, the metaphysical – is the space of *genuine* possibilities; this space includes all the way that things *really* could be. Our task – as when uninformed about *Clue* – is to determine the rules that govern that space (or at

¹²⁷ Robert Nozick thinks that it is virtually impossible. And short of achieving this goal, he wonders what might suffice in its stead: "Perhaps philosophers need arguments so powerful that they set up reverberations in the brain: if the person refuses to accept the conclusion, he *dies*" (Nozick 1981, 4).

least the rules that govern one of its subregions). Insofar as we are successful, we can have justified beliefs about modal matters. If it's indeed the case that we have no special sort of access to the world's rules – and why suppose otherwise? – we can do no better than to theorize about them, doing the best we can with any information available, reasoning about modal matters accordingly.

I think that this analogy provides good reason to accept TT and its version of modal skepticism. The argument is a dilemma. Either we know the factors that determine the space of possibilities, or we don't. If we do, and if we use that knowledge to reason to a given modal claim, then we are operating as TT says we should. If we don't know the factors that determine the space of possibilities, then we shouldn't venture opinions about what's possible, since those opinions will be nothing more than wild guesses. So, we ought to accept TT.

Someone is bound to ask how I *know* that there are factors that determine the space of possibilities. I don't.¹²⁸ However, if there aren't any such factors, then I think that we should despair of having any modal knowledge whatever. If modal reality is not tolerably well-*organized* – if we can't extrapolate some facts about the space of possibilities using principles that hold in the actual world – then we should simply stop talking about the ways that things could and couldn't be. Again, think about the space outside our lightcone. If weren't willing to say that certain general truths that hold within our lightcone also hold outside it, then on what basis would we say anything at all about the distant parts of the universe? Surely none. And the same point applies, *mutatis mutandis*, to the idea that we might have some basis for venturing an opinion about a chaotic space of worlds.

Here is a different analogy that makes the point more clearly still. Imagine that you are a spy who is captured in the middle of a daring mission. Your captors drug you and take you to an

¹²⁸ Or if I do, it's only because of the work that this assumption does in some grand explanatory framework.

undisclosed location (undisclosed to you, anyway). You wake up slumped in the corner of windowless room; two of your captors are standing in the opposing corner, guns drawn. They say:

You are in a building. This may or may not be the building's only room. Give us your estimate of the contents of this building, being careful to indicate, for any two objects, whether or not you judge them to be in the same room. And by the way: you may not leave this room, we will not answer any questions about the building or its contents, and you should not assume that the other rooms – if others there are – bear any resemblance to this room.

And of course, because this is a spy story, they go on to say:

If your estimation is largely correct, then we will release you immediately; if it is not, then we will torture you until you give up national security secrets.

There should be little doubt that you are about to be tortured. There is virtually *no way* that your estimation will be correct; you have no information from which you can plausibly extrapolate the facts about the building. But now let's suppose that realism about modality is true, and that we are causally isolated from the truthmakers for modal claims, and that the worlds are not organized in any discernable way (i.e., if they have no invariant features that we can discover by theorizing about actuality). If all this is true about modal reality, then what's the difference between your epistemic situation in our spy story and your epistemic situation vis-à-vis modal reality? To my mind, none whatever.

So in my view, we have to choose between (a) gambling on the assumption that there are (detectable) factors that (jointly) determine the space of absolute possibilities and (b) denying that we have any modal knowledge. Whether we're willing to make the assumption in (a) will probably come down to our threshold for epistemic risk. But for those of us who cannot see how to do without modality, it will be obvious what to say.

7.3 Conservatism

Here is a second argument for TT. All other things being equal, we should take on conservative hypotheses; as Quine and Ullian point out, a conservative hypothesis “sacrifices as little as possible of the evidential support, whatever that may have been, that our overall system of beliefs has hitherto been enjoying” (Quine and Ullian 1978, 67). TT is indeed a very conservative modal epistemology; therefore, we should accept TT.

(Of course, TT also is conservative in the sense that it is not the Optimist’s modal epistemology, but that is not the relevant sense of ‘conservative’ here. My present point is that TT fits very well with what are, for many of us, prior commitments.)

If you think that we justifiably believe our best theories, and if you are sympathetic to the claim that we can represent the content of a theory with a set of models, then you have the building blocks for TT. These are not wildly contentious views. Any scientific realist will affirm that we justifiably believe at least *some* of our best theories, and scientific realism is surely the dominant view amongst philosophers and non-philosophers alike. And although the semantic view has its detractors, and while there is disagreement about how its details should be worked out, it still appears to be the new orthodoxy amongst philosophers of science. TT does not demand much more of us than that we conjoin these two views in a particular way. Therefore, it does not require us to multiply sources of justification; it does not even require us to extend or retool epistemological principles to which we are already committed. TT simply points out that a plausible account of theories and their justification *includes* a modal epistemology: if our ordinary and scientific theories are justified, and if their content can be represented by sets of models, then we have essential components for TT. The only additional claim – or, rather, the only essential additional claim – is that this route to modal knowledge is

the only one available.¹²⁹

I think that this additional claim is less contentious than it may seem. First, note that when someone maintains that p is possible, and p is contentious, we naturally wonder about the theses to which that person is thereby committed. If Jones says, for example, that he could be disembodied, we would be right to ask what he believe in virtue of which this seems possible to him. Does he believe in Cartesian souls? Aristotle's substantial forms? Does he think that the mind is to the brain as software is to hardware? (If so, what runs the software when he is in a disembodied state?) Does he think that he is an abstract object? A set of time-slices, perhaps? (If a set, then isn't it true that, strictly speaking, he is *necessarily* disembodied?) Suppose that Jones denies that he believes any of these things, and he is unwilling to offer an alternative hypothesis. What, then, should we make of his claim that he could be disembodied? I suspect that many of us will doubt that he is entitled to it; we will doubt that he has really thought the matter through. To my mind, this is some evidence that we are looking for a theory behind Jones' modal claim – for some hypothesis about the world that lends credibility to his modal claim. When we fail to find one, we stop taking his claim as seriously.

Second, suppose that we propose a modal claim, citing our ability to imagine that p as some evidence for p 's possibility. Jones resists: "Of course I can *imagine* that p ," he says, "but so what?" If we want to persuade Jones, we can do no better than to show him that p 's possibility follows from a theory that he accepts. (If you doubt this, go talk to someone who isn't a philosopher. When I was recently asked by an engineer to give an example of a possibility claim, I pointed to a picture frame and said, 'This frame could hover six inches away from the wall, unsupported by any visible means.' She denied that this was possible, I noted that

¹²⁹ ...with the possible exception of trivial inferences from actuality to possibility and properly basic modal beliefs. I wonder, though, whether the former can be seen as following from a general theory of modality, itself justified on abductive grounds.

she was certainly able to imagine the frame hovering away from the wall, and she then reacted precisely as Jones does. I replied by pointing out that quantum indeterminacy allows for such oddities. She conceded.) If Jones initially rejects p 's possibility, why does he defer to the theory later on? One appealing explanation is that he *always* defers to some theory or other. However, because our folk theories are largely shared, and because we rarely consider weird modal claims, it usually suffices to appeal to the imagination, the deliverances of which are shaped by the theories that we accept. But when confronted with more unusual claims, we aren't sure whether to trust our imaginations; we want some confirmation that they are not leading us astray. Hence, we make explicit recourse to our theories in those contexts. Insofar as this explanation is plausible, it lends some credibility to the claim that our theories are our only route to modal knowledge, and thereby to my suggestion that TT harmonizes with our antecedently-held beliefs about these matters.

Finally, while on the topic of TT's conservatism, let me point out that it would be very strange to be a complete skeptic about our knowledge of this world while being wildly optimistic about our knowledge of other worlds. I think it likely that our views about the potential scope of our modal knowledge will be correlated with our views about the potential scope of our knowledge of actuality.¹³⁰ If TT is true, then this is so. If you think, for example, that we are really only justified in believing what our best science tells us, then TT explains why it's appropriate for you to take a similarly high view of the modal claims that our best science

¹³⁰ I don't know whether Cartesian skepticism is the exception to my claim that there is a natural correlation between (alleged) knowledge of actuality and (alleged) knowledge of possibility and necessity. The skeptic denies that we justifiably believe anything of which we are not certain. Presumably, this rules out ordinary and scientific theories but rules in some logic, and therefore the skeptic is likely to regard consistency as a guide to possibility. At first blush, then, we should say that the skeptic is entitled to the modal premises that she employs in her skeptical arguments, since – for example – it does seem that the evil demon scenario is consistent. If this is right, then skepticism is the exception to my rule. However, I wonder whether the skeptic should be *certain* that the evil demon scenario is consistent. What would it take to show this? I am not sure, and I doubt that she is either. This doesn't create a problem for skepticism – since it may suffice for her purposes if the evil demon scenario is merely epistemically possible – but it is enough to keep my correlation principle intact.

underwrites *and* a correspondingly low view of modal claims that follow from non-scientific theories (i.e., folk and metaphysical theories). But if you think, on the other hand, that some of our metaphysical theories are justified, then TT explains why you should take a more generous stance on the potential scope of our modal knowledge, allowing for belief about the physically impossible but metaphysically possible (on the assumption that your metaphysical theory does not entail that what's physically necessary is necessary *simpliciter*).

So, the case for TT's conservatism is quite good. Most of us think that we are justified in believing that (at least some of) our best theories are (approximately) true. The semantic view of theories enjoys wide support. When some proposes or rejects a contentious modal claim, we often appeal to theories to settle the dispute. And, finally, it's plausible that our views about the potential scope of our modal knowledge should track our views about the potential scope of our knowledge generally. I conclude, therefore, that TT has conservatism on its side.

7.4 Disagreement

There is one last virtue that I want to claim for TT: namely, that it's useful for understanding (and sometimes resolving) modal disagreements.

We know how to adjudicate perceptual disagreements. If I say that my keys are on the table, and my wife says that I left them on the dresser in the bedroom, then it doesn't take much to figure out how to settle our dispute. It would be nice to have a modal epistemology that made it equally easy to resolve disagreement. This is probably a pipe dream. Still, it's no doubt a virtue of modal epistemology if it makes it *easier* to sort out differences of modal opinion – and the more tools it offers, the better.

TT gives us a battery of questions to ask when faced with competing modal claims. The

first thing to ask is whether the two parties are reasoning from the same theory. If so, then is it actually clear that the theory speaks to this issue? Is the theory inconsistent? Are the models in question clearly supposed to represent the target system? Has someone made a computational mistake? Suppose that the two parties hold different theories. In that case, after confirming that both theories make a pronouncement about the topic at hand, we should begin to consider their respective merits. What phenomena do the explanations purport to explain? Are they offered to explain the same data? If not, then is there any basis for preferring one set of data to the other? And if they are offered to explain the same data, then how well does each explanation score on the virtues? Can you show that one is superior to the other, and hence that the modal claims following from it are more worthy of belief? If you can't, then is this a case of reasonable disagreement, or should the dispute undermine the confidence of both parties? No doubt there are still more questions that we could ask, but the ones mentioned are illustrative of how we can use TT to locate the source of a modal disagreement – and the prospects for resolving it.

Plainly, TT's approach to modal disagreement is based on there being a tight relationship between the justification of a modal belief and the justification of the corresponding theory. Most modal epistemologies treat the justification of our modal beliefs as largely independent of the justification of our non-modal beliefs.¹³¹ (You might think that identities are the exception to this generalization, although I would deny that identity claims are non-modal – at least if by 'identity' we mean more than any relation that is actually symmetric, transitive, and reflexive.) And if you think that our modal and non-modal beliefs are justified in (more or less) independent ways, then you'll have no trouble with the notion that there can be arguments in which all the disputants agree about the non-modal facts while disagreeing about the modal facts. If TT is

¹³¹ I should say *allegedly* non-modal beliefs, since it is unclear to me whether there is any interesting sense in which there are claims with no modal content whatever.

true, though, then this sort of case is very unlikely – if it is even possible. And this is as it should be, since it is very plausible that the modal supervenes on the non-modal – i.e., there is no difference in the modal facts without a difference in the non-modal facts.¹³² TT directs us to refocus our attention on the non-modal facts, since there very likely *isn't* genuine agreement about them, and that disagreement is likely to provide some insights as to why the disputants are at loggerheads over this or that modal claim. After all, the corresponding non-modal claims will be about objects and properties that were posited for theoretical purposes, and even slight disagreements about the relevant explananda, or about what counts as the best explanation thereof, may lead to radical differences down the line.

I think that modal disagreement is perhaps the area where TT can be most fruitful, although I grant that much more should be said about it. Still, the above should be enough to indicate how TT might be deployed in cases of modal disagreement, and the tools that it offers certainly seem promising.

7.5 **Taking Stock**

At the outset, I raised these four questions:

1. Under what conditions are we justified in believing that *p* is possible?
2. What constitutes evidence for *p*'s possibility?
3. What are we justified in believing about modal matters?
4. How can we resolve disagreements about modal matters?

Shortly thereafter, I said this about my project:

Realism about modality is non-negotiable, and I am not questioning the view that we are causally isolated from the modal facts. Moreover, I have no interest in views that clear

¹³² I take it that, for every (purportedly) non-modal facts, some modal fact or other supervenes on it. When a supervenience relation is universal in this way, one begins to wonder about the plausibility of the claim that the supervenience base and the supervening facts are truly distinct. And insofar as a claim's being non-modal depends on its subject matter being non-modal, the supervenience of the modal on the non-modal provides reason to deny that there are truly non-modal claims.

these two hurdles by postulating a special faculty or cognitive capacity or divine intervention. And even if we have some *a priori* knowledge, I have a strong preference for a view on which our modal knowledge is *a posteriori*. My aim here is to construct a general, simple, and conservative modal epistemology that respects these considerations.

How did I do?

TT certainly provides answers to the above questions. TT is a direct answer to the first question, stating the conditions under which we are justified in believing a modal claim. It also offers a clear response to the second question: since we have empirical evidence for our theories, TT implies that we have empirical evidence for modal claims. Much of what I say in Chapters 1 and 2 answers the third question: we can be justified in believing that our best theories are true, and given TT, we should look to those theories for the specific modal claims that we are currently justified in believing. Essentially, Chapter 4 fills out this answer, clarifying the implications of the slogan, ‘no theory, no justification’. (Granted, what I really provide is a method for answering the third question, not an actual answer; however, short of offering a ridiculously long list, I don’t see what else could reasonably be expected.) In this Conclusion, I offer the beginnings of my answer to the fourth question: in short, we should bring to bear on modal disagreements those tools that we use in theory selection.

I also respected the constraints that I said I would respect. I never questioned realism about modality, nor the view that we are causally isolated from the modal facts. I did not postulate any special means by which have access to modal reality, and I offered an account on which modal knowledge is *a posteriori*. It seems to me, moreover, that TT is not myopically focused on one part of the epistemology of modality, nor is it particularly complex. And finally, as I’ve suggested in this Conclusion, there are some strong considerations in TT’s favor. First, it’s plausible that our epistemic access to non-actual possibility is based on our knowledge of those factors that determine the space of worlds. Essentially, TT turns this observation into a

modal epistemology. Second, for many of us, TT is a conservative theory, fitting neatly with a number of our prior commitments. Third, TT suggest promising methods for understanding and resolving modal disagreement.

Perhaps the Optimist will still deny that these qualities make TT attractive enough to offset the modal skepticism that TT requires. If so, then the Optimist has his work cut out for him: he owes us an epistemology that delivers what TT does not. If the arguments of the last chapter hold up, then it will be no trifling matter to find such an epistemology. Perhaps the best bet for the Optimist to reject the terms of those arguments, contending that consistency is indeed a good guide to possibility. His task then will be to block the arguments for the necessary *a posteriori*. And if this where optimism leads you, then I'm glad not to go along.

I conclude, therefore, that TT has a strong claim to be the epistemology of modality for realists about modality. If there is a better option out there, I am not aware of it.

REFERENCES

- Alston, William P. 1991. *Perceiving God : the epistemology of religious experience*. Ithaca, N.Y.: Cornell University Press.
- Armstrong, D. M. 1989. *A combinatorial theory of possibility, Cambridge studies in philosophy*. New York: Cambridge University Press.
- Baker, Lynne Rudder. 2007. *The metaphysics of everyday life : an essay in practical realism, Cambridge studies in philosophy*. New York: Cambridge University Press.
- Benacerraf, Paul. 1973. "Mathematical truth." *The Journal of Philosophy* no. 70 (19):661-679.
- Benacerraf, Paul, and Hilary Putnam. 1983. *Philosophy of mathematics : selected readings*. 2nd ed. New York: Cambridge University Press.
- Burgess, John P., and Gideon A. Rosen. 1997. *A subject with no object : strategies for nominalistic interpretation of mathematics*. New York: Oxford University Press.
- Chalmers, David J. 1996. *The conscious mind : in search of a fundamental theory, Philosophy of mind series*. New York: Oxford University Press.
- Chisholm, Roderick. 1973. *The Problem of the Criterion*. Milwaukee: Marquette University Press.
- Churchland, Patricia S. 1986. *Neurophilosophy: Toward A Unified Science of the Mind-Brain*. MIT Press.
- Churchland, Paul M. 1989. *A neurocomputational perspective : the nature of mind and the structure of science*. Cambridge, MA: MIT Press.
- Cohnitz, Daniel. 2002. "Modal skepticism: Philosophical thought experiments and modal epistemology." *Vienna Circle Institute Yearbook* no. 10:281--296.
- Colyvan, Mark. 2001. *The indispensability of mathematics*. New York: Oxford University Press.
- Day, Timothy, and Harold Kincaid. 1994. "Putting inference to the best explanation in its place." *Synthese* no. 98:271-295.
- DePaul, Michael. 2001. "Value Monism in Epistemology." In *Knowledge, Truth, and Duty: Essays on Epistemic Justification, Responsibility, and Virtue*, edited by Matthias Steup, 170-183. New York: Oxford University Press.
- DePaul, Michael R., and William Ramsey. 1998. *Rethinking intuition : the psychology of intuition and its role in philosophical inquiry, Studies in epistemology and cognitive theory*. Lanham, Md.: Rowman & Littlefield.

- Devitt, Michael. 2010. *Putting metaphysics first : essays on metaphysics and epistemology*. Oxford ; New York: Oxford University Press.
- Devitt, Michael, and Kim Sterelny. 1999. *Language and reality : an introduction to the philosophy of language*. 2nd ed. Cambridge, Mass.: MIT Press.
- Divers, John. 2007. "Quinean scepticism about de re modality after David Lewis." *European Journal of Philosophy* no. 15 (1):40–62.
- Einstein, Albert, Max Born, and Hedwig Born. 1971. *The Born-Einstein letters: correspondence between Albert Einstein and Max and Hedwig Born from 1916-1955, with commentaries by Max Born*. London: Macmillan.
- Field, Hartry H. 1980. *Science without numbers : a defence of nominalism*. Princeton, N.J.: Princeton University Press.
- Frigg, Roman. 2002. Models and representation: Why structures are not enough. In *Measurement in Economics and Natural Sciences*. Tinbergen Institute, Amsterdam.
- Fumerton, Richard. 1980. "Induction and reasoning to the best explanation." *Philosophy of Science* no. 47 (4):589-600.
- Gale, Richard. 1996. "Some difficulties in theistic treatments of evil." In *The evidential argument from evil*, edited by Daniel Howard-Snyder, 206-218. Bloomington, Ind.: Indiana University Press.
- Geirsson, Heimir. 2005. "Conceivability and defeasible modal justification." *Philosophical Studies* no. 122 (3):279-304.
- Gentner, Dedre, and Albert L. Stevens. 1983. *Mental models, Cognitive science*. Hillsdale, N.J.: Erlbaum.
- Giaquinto, M. 2008. "The linguistic view of a priori knowledge." *Philosophy* no. 83 (1):89-111.
- Giere, Ronald N. 1979. *Understanding Scientific Theories*. New York: Holt, Rinehart, and Winston.
- Goldberg, Elkhonon. 2005. *The wisdom paradox : how your mind can grow stronger as your brain grows older*. New York: Gotham Books.
- Gopnik, Alison, and Andrew N. Meltzoff. 1997. *Words, thoughts, and theories, Learning, development, and conceptual change*. Cambridge, Mass.: MIT Press.
- Gunner, D. L. 1967. "Professor Smart's 'Sensations and brain processes'." In *The Identity Theory of Mind*. St. Lucia: University of Queensland Press.

- Hanrahan, Rebecca Roman. 2007. "Imagination and possibility." *The Philosophical Forum* no. 38:125-146.
- Hart, W. D. 1988. *The engines of the soul, Cambridge studies in philosophy*. New York: Cambridge University Press.
- . 1989. "The price of possibility." *Pacific Philosophical Quarterly* no. 70:225-239.
- Hawke, Peter. 2011. "Van Inwagen's modal skepticism." *Philosophical Studies* no. 153 (3):351-364.
- Hempel, Carl G., and Paul Oppenheim. 1948. "Studies in the Logic of Explanation." *Philosophy of Science* no. 15 (2):135-175.
- Huemer, Michael. 2001. *Skepticism and the veil of perception, Studies in epistemology and cognitive theory*. Lanham, Md.: Rowman & Littlefield Publishers.
- Huggett, Nick. 2010. *Everywhere and everywhen : adventures in physics and philosophy*. Oxford ; New York: Oxford University Press.
- Jackson, Frank. 1998. *From metaphysics to ethics: a defence of conceptual analysis*. New York: Oxford University Press.
- Jenkins, C. S. 2008. "Modal knowledge, counterfactual knowledge and the role of experience." *Philosophical Quarterly* no. 58 (233):693-701.
- Johnson-Laird, P.N. 2004. "The history of mental models." In *Psychology of reasoning: theoretical and historical perspectives*, edited by K. Manktelow and M. C. Chung, 179-212. New York: Psychology Press.
- Kagan, Shelly. 1988. "The additive fallacy." *Ethics* no. 99 (1):5-31.
- Kim, Jaegwon. 1994. "Explanatory knowledge and metaphysical dependence." *Philosophical Issues* no. 5:51-69.
- Klein, Colin. 2008. "An ideal solution to disputes about multiply realized kinds." *Philosophical Studies* no. 140 (2).
- Koons, Robert C. 2000. *Realism regained : an exact theory of causation, teleology, and the mind*. New York: Oxford University Press.
- Kripke, Saul A. 1980. *Naming and necessity*. Cambridge, Mass.: Harvard University Press.
- LaPorte, Joseph. 2004. *Natural kinds and conceptual change, Cambridge studies in philosophy and biology*. New York: Cambridge University Press.

- Lemos, Noah Marcelino. 2004. *Common sense : a contemporary defense, Cambridge studies in philosophy*. New York: Cambridge University Press.
- Leon, Felipe. 2009. *Armchair Knowledge and Modal Skepticism: A Rapprochement*, Philosophy, University of California, Riverside, Riverside.
- Lewis, C. I. 1970. "The pragmatic element in knowledge." In *Collected Papers of Clarence Irving Lewis*, edited by John D. Goheen and John L. Mothershead, 240-257. Stanford: Stanford University Press.
- Lewis, David. 1986a. "Causal explanation." In *Philosophical Papers Vol. II*, edited by David Lewis, 214-240. New York: Oxford University Press.
- Lewis, David K. 1973. *Counterfactuals*. Oxford: Blackwell.
- . 1986b. *On the plurality of worlds*. New York, NY: Blackwell.
- Lloyd, Elisabeth Anne. 1994. *The structure and confirmation of evolutionary theory, Princeton paperbacks*. Princeton, N.J.: Princeton University Press.
- Maddy, Penelope. 1990. *Realism in mathematics*. New York: Oxford University Press.
- McGinn, Colin. 1999. *Knowledge and reality : selected essays*. New York: Oxford University Press.
- . 2000. *Logical properties : identity, existence, predication, necessity, truth*. New York: Oxford University Press.
- Molnar, George, and Stephen Mumford. 2003. *Powers : a study in metaphysics*. Oxford ; New York: Oxford University Press.
- Moore, G E. 1963. "Four Forms of Scepticism." In *Philosophical Papers*. New York: Collier Books.
- Nolan, Daniel. 1997. "Impossible Worlds: A Modest Approach." *Notre Dame journal of formal logic* no. 38 (4):535-573.
- Nozick, Robert. 1981. *Philosophical explanations*. Cambridge, Mass.: Harvard University Press.
- . 1993. *The nature of rationality*. Princeton, N.J.: Princeton University Press.
- . 2001. *Invariances : the structure of the objective world*. Cambridge, Mass.: Belknap Press of Harvard University Press.
- O'Connor, Timothy. 2008. *Theism and Ultimate Explanation: The Necessary Shape of Contingency*. Blackwell Pub.

- O'Leary-Hawthorne, John. 1996. "The epistemology of possible worlds: A guided tour." *Philosophical Studies* no. 84 (2-3):183-202.
- Oreste Fiocco, M. 2007. "Conceivability, imagination and modal knowledge." *Philosophy and Phenomenological Research* no. 74 (2):364–380.
- Peacocke, Christopher. 1999. *Being known*. New York: Oxford University Press.
- Plantinga, Alvin. 1993. *Warrant and proper function*. New York: Oxford University Press.
- Plantinga, Alvin, and Michael Tooley. 2008. *Knowledge of God, Great debates in philosophy*. Malden, MA: Blackwell.
- Psillos, Stathis. 1999. *Scientific realism : how science tracks truth, Philosophical issues in science*. London ; New York: Routledge.
- Putnam, Hilary. 1979. *Mathematics, matter, and method: philosophical papers, vol. 1*. 2d ed. New York: Cambridge University Press.
- Quine, W. V. 1981a. *Theories and things*. Cambridge, Mass.: Harvard University Press.
- Quine, W. V., and J. S. Ullian. 1978. *The web of belief*. 2d ed. New York: Random House.
- Quine, WV. 1981b. "Five milestones of empiricism." In *Theories and things*, 67–72.
- Rea, Michael C. 2002. *World without design : the ontological consequences of naturalism*. New York: Oxford University Press.
- Riggs, Wayne D. 2008. "Epistemic risk and relativism." *Acta Analytica* no. 23 (1):1-8.
- Rowe, William. 1979. "The problem of evil and some varieties of atheism." *American Philosophical Quarterly* no. 16 (4):335-341.
- Salmon, Nathan U. 2007. *Content, cognition, and communication, Philosophical papers (Oxford University Press)*. New York: Oxford University Press.
- Salmon, Wesley C. 1989. *Four decades of scientific explanation*. Pittsburgh: University of Pittsburgh Press.
- Sidelle, Alan. 1989. *Necessity, essence, and individuation : a defense of conventionalism*. Ithaca: Cornell University Press.
- Sosa, Ernest. 2000. "Modal and Other A Priori Epistemology: How Can We Know What Is Possible and What Impossible?" *The Southern Journal of Philosophy*.

- Stich, Stephen P. 1990. *The fragmentation of reason : preface to a pragmatic theory of cognitive evaluation*. Cambridge, Mass.: MIT Press.
- Strawson, P. F. 1959. *Individuals, an essay in descriptive metaphysics*. London: Methuen.
- Suppe, Frederick. 1977. *The Structure of scientific theories*. 2d ed. Urbana: University of Illinois Press.
- . 1989. *The semantic conception of theories and scientific realism*. Chicago: University of Illinois Press.
- Suppes, Patrick. 1993. *Models and methods in the philosophy of science : selected essays*, *Synthese library* v. 226. Dordrecht ; Boston: Kluwer Academic.
- Swinburne, Richard. 2001. *Epistemic justification*. New York: Oxford University Press.
- . 2004. *The existence of God*. 2nd ed. New York: Oxford University Press.
- Tahko, Tuomas E. 2009. "On the Modal Content of A Posteriori Necessities." *Theoria* no. 75 (4):344-357.
- Thomasson, Amie L. 2007. *Ordinary objects*. New York: Oxford University Press.
- . Forthcoming. "Modal Normativism and the Methods of Metaphysics." *Philosophical Topics*.
- Thompson, Paul. 1989. *The structure of biological theories, SUNY series in philosophy and biology*. New York: State University of New York Press.
- . 2007. "Formalisations of evolutionary biology." In *Philosophy of Biology*, edited by Mohan Matthen and Christopher Stephens, 485-523. Oxford: North-Holland.
- Vaidya, Anand Jayprakash. 2010. "Understanding and essence." *Philosophia* no. 38 (4):-.
- Van Fraassen, Bas C. 1980. *The scientific image*. New York: Oxford University Press.
- . 1989. *Laws and symmetry*. New York: Oxford University Press.
- van Inwagen, Peter. 1977. "Ontological arguments." *Nous* no. 11 (4):375-395.
- . 1979. "Review: The Coherence of Theism, by Richard Swinburne." *The Philosophical Review* no. 88 (4):668-672.
- . 1995. *God, knowledge & mystery : essays in philosophical theology*. Ithaca: Cornell University Press.

- . 1997. "Materialism and the psychological-continuity account of personal identity." *Philosophical Perspectives* no. 11:305-319.
- . 1998. "Modal Epistemology." *Philosophical Studies* no. 92 (1-2):67-84.
- . 2006. *The problem of evil : the Gifford lectures delivered in the University of St. Andrews in 2003*. New York: Oxford University Press.
- . 2009. *Metaphysics*. 3rd ed. Boulder, Colo.: Westview.
- Vogel, Jonathan. 1990. "Cartesian skepticism and inference to the best explanation." *The Journal of Philosophy* no. 87 (11):658-666.
- Williamson, Timothy. 2007. *The philosophy of philosophy, The Blackwell/Brown lectures in philosophy*. Malden, MA: Blackwell Pub.
- Woodward, James. 2003. *Making things happen : a theory of causal explanation, Oxford studies in philosophy of science*. New York: Oxford University Press.
- Yablo, Stephen. 1993. "Is conceivability a guide to possibility?" *Philosophy and Phenomenological Research* no. 53 (1):1-42.
- Yagisawa, Takashi. 1988. "Beyond possible worlds." *Philosophical Studies* no. 53 (2).
- . 2010. *Worlds and Individuals, Possible and Otherwise*: Oxford University Press.
- Yandell, Keith. 1988. "Divine Necessity and Divine Goodness." In *Divine and human action : essays in the metaphysics of theism*, edited by Thomas V. Morris, 313-344. Ithaca, N.Y.: Cornell University Press.

Dr. Robert William Fischer

Texas State University-San Marcos
Department of Philosophy
601 University Drive
Psychology 110
San Marcos, Texas, 78666

585.739.9686
robert.william.fischer@gmail.com

Employment

Senior Lecturer, Texas State University-San Marcos	2011-Present
--	--------------

Education

Ph.D., Philosophy, University of Illinois at Chicago	2011
M.A., Philosophy of Religion, Trinity International University	2006
B.A., English / Philosophy, State University of New York at Geneseo	2004

AOS: M&E / Philosophy of Science

AOC: Philosophy of Religion / Logic / Ethics

Talks

“The Theory Theory: A Neglected Modal Epistemology”
APA Eastern Division Meeting, December 2011

“Modal Knowledge, in Theory”
Southwestern Philosophical Society Meeting, November 2011
To be published in *Southwest Philosophy Review*

“Interpreting the ‘Wall of Separation’”
The Dialogue Series, Texas State University-San Marcos, October 2011

“On Hawthorne on Lewis on the Case for Modal Realism”
Central States Philosophical Association Meeting, September 2011

“Comments on Muntean’s ‘Genetic Numerical Simulation’”
Indiana Philosophical Association Spring Meeting, April 2011

“The Theory Theory: A Neglected Modal Epistemology”
Indiana Philosophical Association Spring Meeting, April 2011

“Comments on Lee’s ‘Are Analyticity and Necessity a Posteriori Mutually Exclusive?’”
Joint Meeting of Illinois and Indiana Philosophical Associations, November 2010

“A Better Basis for Modal Skepticism”
Joint Meeting of Illinois and Indiana Philosophical Associations, November 2010

“A Better Basis for Modal Skepticism”
Alabama Philosophical Society’s 48th Annual Conference, September 2010

Papers Under Review for Publication

“A Theory-based Epistemology of Modality” (*Australasian Journal of Philosophy*)
“Why It Doesn’t Matter Whether Simplicity Is Truth-Conducive” (*Synthese*)

Reviews

“Lance Rips’ *Lines of Thought: Central Concepts in Cognitive Psychology*.”
Philosophical Psychology (Forthcoming).

“Bas van Fraassen’s *Scientific Representation: Paradoxes of Perspective*.” *Heythrop Journal* 52, no. 2 (2011): 301.

“Timothy O’Connor’s *Theism and Ultimate Explanation: The Necessary Shape of Contingency*.” *Faith and Philosophy* 27, no. 4 (2010): XXX-XXX.

“Louise Antony’s *Philosophers Without Gods: Meditations on Atheism and the Secular Life*.” *International Journal for the Philosophy of Religion* 66, no. 2 (2009): 119-123.

“Plantinga and Tooley’s *Knowledge of God* (Blackwell Great Debates in Philosophy).” *Heythrop Journal* 50, no. 3 (2009): 513-515.

“Amy Thomasson’s *Ordinary Objects*.” *Metaphilosophy* 40, no. 2 (2009): 296-302.

“Joseph LaPorte’s *Natural Kinds and Conceptual Change*.” *Erkenntnis* 69 (2008): 415-419.

Courses Taught

Ethics and Society
Introductory Symbolic Logic
The Philosophy of Death & Dying
Philosophy of Science

Introduction to Ethics
Introduction to Philosophy
Philosophy and Revelation
Applied Philosophy (Animal Ethics)

Dissertation: “Modal Knowledge, in Theory”

We have some justified beliefs about modal matters. I articulate an account of that justification. On my view, we are justified in believing a modal claim only if we are justified in believing a theory according to which that claim is true. Defending this requires (1) an account of theories on which they have modal content and (2) an account of how we can be justified in believing that theories are true. I use the semantic view of theories to satisfy the former requirement and inference to the best explanation (IBE) to satisfy the latter. Since IBE has been roundly criticized, I devote considerable attention to its defense. I also explore a number of potential defeaters to our having justified modal beliefs, including the modal analogue of Benacerraf’s Dilemma. Ultimately, though, I show that my leads to the view that the scope of our modal knowledge is modest, and I argue that modal modal epistemologies that are more optimistic are implausible on that basis.

Committee: W. D. Hart (Chair), Colin Klein, Walter Edelberg,
Karen Bennett, Daniel Sutherland

Doctoral Course Work

Theories of Vision (David Hilbert); Hegelian Critiques of Empiricism (Sally Sedgwick); Non-Reductive Physicalism (Colin Klein); Philosophical Foundations of Human Rights (Tony Laden); Personal Identity (Marya Schechtman); Modal Epistemology (W. D. Hart); Function in Biology and Psychology (Colin Klein); Berkeley (David Hilbert); Direct Reference (Walter Edelberg); Truth and Beauty in the Medieval Tradition (Georgette Sinkler); Metalogic I (Daniel Sutherland); The History of Modality (Peter Hylton); Survey of Contemporary Metaphysics (Walter Edelberg)

Other

Compiled index for Samuel Fleischacker’s *Divine Teaching and the Way of the World: A Defense of Revealed Religion* (New York: Oxford University Press, 2011).

Provided Oxford University Press with instructor feedback on Russ Shafer-Landau’s *The Fundamentals of Ethics* (New York: Oxford University Press, 2010).

References

W. D. Hart, Professor Emeritus
Department of Philosophy, M/C 267
1400 University Hall
University of Illinois at Chicago
Chicago, IL 60607-7114
hart@uic.edu
312.413.1278

Colin Klein, Associate Professor
Department of Philosophy, M/C 267
1420 University Hall
University of Illinois at Chicago
Chicago, IL 60607
cvklein@uic.edu
312.413.1801

Walter Edelberg, Associate Professor
Department of Philosophy, M/C 267
1406 University Hall
University of Illinois at Chicago
Chicago, IL 60607
edelberg@uic.edu
312.996.3022

Marya Schechtman, Professor
Department of Philosophy, M/C 267
1404 University Hall
University of Illinois at Chicago
Chicago, IL 60607
marya@uic.edu
312-413-7565