



AUTHOR: J. ADAM CARTER

MATRIC. NO.: 0787306

TITLE: THE VALUE OF KNOWLEDGE AND THE PROBLEM OF EPISTEMIC LUCK

DEPARTMENT: PHILOSOPHY

SCHOOL: PHILOSOPHY, PSYCHOLOGY AND LANGUAGE SCIENCES

SUPERVISOR/S: PROFESSOR DUNCAN PRITCHARD

EXAMINERS: ERNEST SOSA, JESPER KALLESTRUP

ACKNOWLEDGMENTS

I dedicate this work to my advisor, Duncan Pritchard, to whom I am profoundly indebted for his inspiration, unwavering help, and more importantly, for the influence he has had on me as a philosopher and as a person.

TABLE OF CONTENTS

INTRODUCTION

- 0.1 Motivations for the Project
- 0.2 Project Overview
- 0.3 Chapter Descriptions

CHAPTER ONE: “The Value of Knowledge and the Nature of Knowledge”

- 1.1 Introduction: The Place of Normativity in 20th Century Epistemology
- 1.2 A New Way of Thinking: The *Meno* Assumption and the Link Between Knowledge and its Value
- 1.3 Three Master Questions
 - 1.3.1 First Question: What is the *Meno* Assumption?
 - 1.3.2 Second Question: Is the *Meno* Assumption True?
 - 1.3.3 Third Question: How Should a Theory Be Held to Account for MA?
- 1.4 Concluding Remarks

CHAPTER TWO: “Swamping, Additive Value and Epistemic Parts”

- 2.1 The swamping problem
- 2.2 Challenging the Swamping Thesis
 - 2.2.1 Invariabilism and Additivism
 - 2.2.2 The Value of Reliability
- 2.3 Avenues of Defence
 - 2.3.1 The Value of Reliability as Expected Value
 - 2.3.2 Additivism and Epistemic Parts
- 2.4 Diagnosing the Appeal to Swamping

CHAPTER THREE: “Subjectivism, Objectivism and the Epistemic Value of Tokens and Types”

- 3.1 Introduction
- 3.2 Subjectivism and Objectivism
- 3.3 Invariance
- 3.4 Subjectivism, Objectivism, Types and Tokens
- 3.5 Subjective Value and the Token Criterion
- 3.6 The Proportionality Thesis and the Token Criterion
- 3.7 Concluding Remarks

CHAPTER FOUR: “The Epistemic Point of View”

- 4.1 Introduction
- 4.2 The Strong Case for Veritism
- 4.3 The Normative and Teleological Accounts of Truth-Directedness
- 4.4 A Different Case for Veritism
- 4.5 Kvanvig, David and the Knowledge Objection to Veritism
 - 4.5.1 Marian David’s Circularity Response to the Knowledge Objection
 - 4.5.2 Kvanvig’s Response to David
 - 4.5.3 What Does the Knowledge Objection Tell Us About Veritism?
- 4.4 Ecumenical Veritism
- 4.5 The Collapsing Defence
- 4.6 Veritism and Epistemic Responsibility
- 4.7 Knowledge and the Epistemic Point of View
- 4.8 Concluding Remarks

CHAPTER FIVE: “Evidentialism, Reliabilism and the Value of Knowledge”

- 5.1 Introduction
- 5.2 Evidentialism
- 5.3 Mentalist Internalism
- 5.4 Reliabilism
 - 5.4.1 Olsson, Reliability and Stability
 - 5.4.2 Goldman & Olsson on the Value of Reliability: Two Attempts
 - 5.4.2 (a) Reliability, Type- and Token-Instrumental Value
 - 5.4.2 (b) The Conditional Probability Solution and the Value of Future True Beliefs
- 5.5 Concluding Remarks

CHAPTER SIX: “Knowledge, Ability and Final Value”

- 6.1 Introduction

- 6.2 Epistemic Value Pluralism: Belief and Agency
- 6.3 Rabinowicz and Rønnow-Rasmussen on Final Value
- 6.4 Knowledge and Final Value
- 6.5 Addressing the Type Criterion and the Token Criterion

CHAPTER SEVEN: “Knowledge, Credit and Testimony”

- 7.1 Introduction
- 7.2 Lackey’s Argument
- 7.3 First Line of Response
 - 7.3.1 Testimony and the Epistemic Situation
 - 7.3.2 Lackey’s Case Revisited
 - 7.3.4 The Structure of Knowledge Transmission
 - 7.3.5 Testimony and the Structure of Reasons
 - 7.3.6 Morris’s Evidence: A Closer Look
- 7.4 The Second Line of Response
 - 7.4.1 Recasting the Objection
 - 7.4.2 Against the Disjunctive View of Testimonial Credit Attributions
 - 7.4.3 Agency, Credit and Luck
- 7.5 Concluding Remarks

CHAPTER EIGHT: “John Greco on Credit, Salience and Causal Explanation”

- 8.1 Introduction
- 8.2 Greco’s New Proposal
- 8.3 Salience and Causal Explanation
- 8.4 Causal Explanation and Context Dependence
- 8.5 Salience, Contextualism and Interests: Three Lines of Defence
 - 8.5.1 Luck and the Gambler
 - 8.5.2 Default Salience
 - 8.5.3 Interests and Purposes as Indexed to a Practical Reasoning Environment
- 8.6 Knowledge, Interests and Objectivity
- 8.7 Perceived Interests
- 8.8 Concluding Remarks

CHAPTER NINE: “The Problem of Epistemic Luck”

- 9.1 Introduction
- 9.2 Lack of Control Accounts of Luck
- 9.3 What Does LCAL Tell Us About Epistemic Luck?
- 9.4 Motivating an Account of Luck that Doesn’t Appeal to Control
- 9.5 A Conundrum for Pritchard’s Safety: Lottery Propositions and the Garbage Chute Case

- 9.6 Kvanvig's Closure Objection
- 9.7 Becker on Luck and Reliability
- 9.8 Lackey's Necessity and Sufficiency Objections
- 9.9 Concluding Remarks

CHAPTER 10: "Knowledge in the Balance of Luck and Ability: Toward a Robust VE"

- 10.1 Introduction
- 10.2 Choosing the Best Version of ALVE
 - 10.2.1 Against Single-Condition (Anti-Luck) ALVE
 - 10.2.2 Against Dual-Condition ALVE
 - 10.2.3 Single-Condition (Ability) ALVE
- 10.3 Motivating a Robust Virtue Epistemology
- 10.4 A Positive Defence for RVE4
 - 10.4.1 Lottery Propositions
 - 10.4.2 The Garbage Chute Case
- 10.5 Putting it All Together
- 10.6 Concluding Remarks

Introduction

o.1 Motivations for the project

Imagine that you've just spent the last several months reading *Don Quixote*—and that you're all but fifty pages away from finishing. Unfortunately for you, the book was due back before you could finish, and so

begrudgingly, you turn it back in, having not known what happens in the end. Riddled with curiosity, you make your best guess about Quixote's eventual fate and suppose it is the most likely scenario. Entirely unbeknownst to you, it turns out that you were right; Quixote's ultimate destiny was just what you had supposed it would be! What luck!

Quite naturally, we would say that (despite how impressed we are that you rightly anticipated Cervantes), when all is said and done, *knowing* what happens to Don Quixote in the end is surely better than merely *believing truly* what happens in the end—the predicament you find yourself in having not actually finished the book. After all, it was just by dumb luck that you guessed the ending right—a point you could deny only on the pain of some embarrassing claim of clairvoyance.

We might put the idea more generally by saying that from a purely cognitive standpoint, it is better to *know* the truth than to stumble upon it by luck. This general idea betrays two distinct insights about knowledge. The first is the insight that while true belief is valuable, knowledge is distinctively so—*knowing* the truth is valuable in a way that merely *having* a true (but not-known) belief is not. The second insight here is that you lack knowledge if it's just by dumb luck that the belief you have is true. Call these the *value insight* and the *anti-luck insight*:

Value insight: Knowledge is distinctively valuable.

Luck insight: Knowledge excludes luck.

In contemporary epistemology, and especially over the past five years, separate projects have arisen in correspondence with these distinct intuitions: *value-driven epistemology* is concerned with issues surrounding the first insight, and projects under the description of *anti-luck epistemology* have arisen out of the second.¹ Now we might reasonably suppose that whatever it is that makes knowledge relevantly *un-lucky* would be something

¹ Both the value insight and the luck insight are important for epistemologists as each provides a glimpse into what it is a theory of knowledge should be capturing when engaging in the task of saying what knowledge is. Projects in value-driven epistemology (e.g. Wayne Riggs 2005; John Greco (*forthcoming*); Duncan Pritchard (2007, *forthcoming*); Jonathan Kvanvig (2003, 2007, 2008, *forthcoming*), Guy Axtell (2005, 2006, 2008); Stephen Grimm, Ernest Sosa (2007, 2008) have made great strides in helping us to understand the insight that knowledge is distinctively something better to have than true beliefs that fall short. More generally, these projects have helped us to understand the nature and sources of epistemic value and how these come to bear on the normativity of knowledge. Projects under the heading of *anti-luck epistemology* (Pritchard (2005, 2007, 2007, *forthcoming*) as well as recent work on epistemic luck more broadly (e.g. Kelly Becker 2007, Neta & Hiller 2007, John Greco *forthcoming*; Juan Comesaña 2004; Jennifer Lackey 2007; Christoph Kelp 2008; Jon Kvanvig 2008; Wayne Riggs 2007) have been very useful for the purpose of getting clear about what luck is, just *how* it is that knowledge is supposed to be incompatible with luck, and what condition an analysis of knowledge requires in order to *rule out* the kind of luck thought to be incompatible with it.

we could cite in accounting for what makes knowledge distinctively valuable. This natural idea reflects the thought that the insights about value and luck should not be entirely disconnected. There is a problematic tension though between this reasonable expectation and the resources epistemologists have provided for us to accommodate it. What explains the tension is the fact that *value-driven* and *anti-luck* projects in epistemology have by and large been developed apart from each other, each focused on one of our two guiding insights at the expense of the other. Consequently, value-driven epistemology's focus on the normative but not modal properties of knowledge leaves these two aspects of knowledge disconnected much in the way that *anti-luck* epistemology's focus on the modal but not the normative properties of knowledge leaves these same aspects disconnected.²

The lacuna here between what value driven epistemologists tell us about what makes knowledge valuable and what the anti-luck epistemologists tell us about what makes knowledge exclude luck is troubling. We may resist that there should be such a disconnect if we avoid the common flaw shared by each of these projects pursued in isolation from the other. The flaw here is essentially one of naivety: that of supposing that we can give an account of knowledge exclusively in terms of conditions that would accommodate *one* of the two insights, while still managing to account for the other insight—which itself we did not appeal to directly in our theory of what knowledge is—e.g. in the analysis provided of it. This is the flaw behind the *value-driven* approaches that think about knowledge in terms of valuable properties they can't explain to ensure modal robustness and the anti-luck projects that think of knowledge in terms of modal properties into which we can't well smuggle the normativity needed to explain its value. To avoid this flaw, then, we should let *both* of these insights dictate the conditions our analysis offers as essential to knowing.³ The project with which I'll be engaging here develops substantially on this widely overlooked and promising idea.

0.2 Project overview

² Unsurprisingly, then, the *theories* of knowledge motivated by the *value insight* generally lack any convincing story about how these accounts ensure that known true beliefs are modally stable. And likewise, modal theories of knowledge motivated by the *anti-luck* insight develop their accounts of knowledge out of fundamentally non-normative (modal) insights about what knowledge is—and this becomes a problem when we expect modal theories should then build normativity into the picture so as to explain what makes knowledge distinctively valuable.

³ This will be the overarching task of the project pursued here—a task motivated by the thought that (i) *anti-luck* and *value-driven* theories of knowledge are by themselves unsuccessful, (ii) that this un-success is explained by the disconnect that stands between explanations for distinctive value of knowledge and the luck it is thought to exclude; and (iii) that this disconnect arises from the flaw of supposing that we could preserve both the *value* and *luck* insights while analyzing knowledge in terms of conditions that ensure that it will accommodate just *one* of the insights; and (iv) finally, that to avoid this flaw, both the *value* and *anti-luck* insights must dictate the conditions our analysis offers as essential to knowing.

The broad strategy I pursue for offering an analysis of knowledge in terms of properties that would ensure it to be *both* distinctively valuable and exclude luck begins by clarifying the *value insight* and what would be needed in order to accommodate it. Once the value insight is clarified, I argue that a theory could accommodate it only if it satisfies what I call the *Meno Requirement*, and further, that the two criteria a theory must meet to satisfy the Meno Requirement are ones that point strongly to a theory of knowledge on which knowledge consists essentially in some connection between an agent's true belief and her cognitive abilities that gave rise to it. From the idea that a theory of knowledge in terms of a connection between true belief and ability would be necessary for meeting the criteria for satisfying the Meno Requirement, I go on to show that it would also be *sufficient* for doing so. However, a theory of knowledge along the lines of true-belief-through-ability must also be one that we can demonstrate to ensure that beliefs connected to ability will not depend on luck in the way that the luck insight suggests would be incompatible with knowing. Put another way: we need to show that a connection to ability, which we showed could make a true belief distinctively valuable, can be defended as a connection that would *also* make the true belief insulated from knowledge-undermining luck. The next step then will be to provide an account of luck and then to specify the sense in which luck is incompatible with knowledge. Once these tasks are met, it will become clear that the anti-luck insight is essentially a modal insight that can be preserved only by a theory of knowledge that ensures that true beliefs will count as knowledge only if they are appropriately safe.

What this points to is the thought that a theory must preserve that a true belief counts as knowledge only if it satisfies an ability condition (i.e. it must be appropriately connected to ability) as well as a safety condition. And so we see that a theory of knowledge that preserves both the value and luck insights must fall within the description of what Duncan Pritchard (*forthcoming a*) has called *anti-luck virtue epistemology*. I argue, though, that those with a common commitment to placing an ability and anti-luck condition on knowledge can disagree about whether either of these conditions entails the other. This motivates the thought that there will be three distinct kinds of *anti-luck virtue epistemology* (ALVE). I argue that the only one that could plausibly satisfy the Meno Requirement will be a variety of anti-luck virtue epistemology according to which satisfying the ability condition entails that the belief is appropriately *safe*. And so, the only defensible variety of ALVE turns out to be *robust virtue epistemology*—a view that is committed to showing that the ability condition, as it were, does all the work—including insuring that a belief so produced is appropriately safe. Drawing from the insight that a true belief depends on luck and ability gradiently (and not rigidly, as some have thought) I go on to develop and

defend a type of *robust VE*, one motivated by the thought that a true belief is through ability (and is as such finally valuable) just when the success (truth) of your belief depends on your abilities more so than luck.

0.3 Chapter descriptions

0.3.1 Chapter One: The Value of Knowledge and the Nature of Knowledge

Here I show how the insight that knowledge is distinctively valuable is one that a theory of knowledge must be amenable to. To this end, I aim to answer three questions: (i) How should we articulate this insight? (ii) Is the insight actually correct? (iii) If correct, then how is a theory of knowledge supposed to account for it? Taken together, the answers imply that a theory of knowledge must satisfy what I'll call the Meno Requirement.

0.3.2 Chapter Two: Swamping, Additive Value and Epistemic Parts

A recent and popular idea has been to consider whether a theory of knowledge accommodates the value insight by determining whether the theory succumbs to the Swamping Problem. In this chapter, I argue that the 'Swamping Test' is an inadequate test for assessing whether or not the Meno Requirement is met. I reveal what I call the Swamping Thesis to be committed to additivism about value as well as to certain mereological commitments concerning knowledge and its presumed-to-be-valuable parts. I show that these commitments essential to getting the Swamping Thesis, and thus the Swamping Problem, off the ground have unacceptable implications. Consequently, we can't simply appeal to the Swamping Problem as a way to adjudicate whether a theory satisfies the Meno Requirement, and I conclude that we'll have to look elsewhere for such a test.

0.3.3 Chapter Three: Subjectivism, Objectivism and the Epistemic Value of Types and Tokens

Toward the end of finding a criterion other than the Swamping Problem to determine whether a theory of knowledge satisfies the Meno Requirement, I consider that we've not yet addressed the matter of what *constitutes* the value of knowledge. Subjectivists and objectivists about value disagree about the constitutive grounds of value—whether they are located in valuable objects themselves or in subjects attitudes toward them. In investigating the implications of these positions, we shall see that it matters whether the value insight is understood as an insight about the comparative value of knowledge and true belief *tokens* or as an insight about knowledge and true belief as *types* of epistemic standings. This distinction will be shown to motivate two distinct

criteria—a type criteria and a token criteria—that a theory of knowledge must meet in order to satisfy the Meno Requirement.

0.3.4 Chapter Four: The Epistemic Point of View

In order to apply the two criteria for satisfying the Meno Requirement defended in Chapter 4, we'll need to know what makes something more or less valuable from the epistemic point of view. But this is a matter of contention. *Veritists* and *pluralists* have rival accounts on this score, and the aim of this chapter will be to determine which account is correct. I argue that veritism runs in to troubles leaving it indefensible and that the most plausible account of the epistemic point of view will be a variety of *pluralism* on which two distinct sources of epistemic value are recognized: one governing correctness of belief and the other correctness of agency.

0.3.5 Chapter Five: Evidentialism, Reliabilism and the Value of Knowledge

Having contended that the pluralist account of epistemic value defended in Chapter 5 suggests that what makes knowledge distinctively valuable is best explained by supposing knowledge derives value from *both* of the two distinct sources of epistemic value (correct belief and correct agency), it is concluded that the Meno Requirement appears to call for a theory of knowledge on which knowledge is understood in terms of a connection between correct belief and correct agency. It would be too quick though to conclude this before considering the prospects of rival accounts of knowledge. In this chapter, I consider the prospects of internalist evidentialism and externalist reliabilism. Seeing why both fail, it will appear now more clear that an account of knowledge in terms of a connection between true belief and ability would be required in order to account for the distinctive value we take knowledge to have.

0.3.6 Chapter Six: Knowledge, Ability and Final Value

Whereas the previous chapter gave stronger traction to the thought that a theory of knowledge along the lines of true belief through ability would be *necessary* for answering the Meno Requirement, the aim of this chapter will be to show such an account would also be *sufficient*. I defend here the idea that true belief through ability constitutes a cognitive achievement, and that true beliefs that are the product of cognitive achievements can be shown to have *final value* that mere true beliefs lack. Thus, if knowledge is understood along the lines of true belief through ability, then knowledge can be defended as inheriting the final value of cognitive achievements—which I argue at the end would satisfy the criteria for answering the Meno Requirement.

o.3.7 Chapter Seven: Knowledge, Credit and Testimony

If cognitive achievements are indeed finally valuable as we've said, then it should be pointed out that this counts in favor of the thought that a theory of knowledge along the lines of true belief through ability could satisfy the Meno Requirement only if knowledge requires cognitive achievements. Indeed, if there can be knowledge in the absence of cognitive achievement, then the arguments developed in Chapter 7 will fail to explain why knowledge *in the absence* of cognitive achievements is distinctively valuable. And this would take the final value solution off the table. Jennifer Lackey has recently offered a challenge to the necessity of cognitive achievements for knowledge, and the aim of this chapter is to rebut this challenge. I show that the example she gives fails to show that knowledge and cognitive achievements come apart.

o.3.8 Chapter Eight: John Greco on Credit, Salience and Causal Explanation

A theory of knowledge along the lines of true belief through ability is one we can appeal to in answering the Meno Requirement only knowledge can be *defended* as consisting essentially in such a connection. John Greco (2009a) has recently offered a comprehensive attempt to defend such a proposal, and in this chapter, I examine the view he develops in some detail; what I conclude from Greco's own attempt to understand knowledge as true belief through ability is that the connection between true belief and ability cannot plausibly be defended as a causal connection nor as one that relies on the notion of explanatory salience.

o.3.9 Chapter Nine: On the Problem of Epistemic Luck

In order to improve on Greco's attempt to defend knowledge as essentially consisting in a relationship between true belief and ability, we'll need to show how such a connection could exclude luck. I begin here by offering an account of luck and then specifying the sense in which luck should be thought incompatible with knowledge. What is concluded is that luck is essentially a modal notion and that this will motivate a safety requirement on knowledge. I consider several challenges to a safety requirement on knowledge and conclude that the variety of safety best suited to handling these objections is Pritchard's (2007) more recent account.

o.3.10 Chapter Ten: Knowledge in the Balance of Luck and Ability: Toward a Robust VE

What this points to is the thought that a theory must preserve that a true belief counts as knowledge only if it satisfies an ability condition (i.e. it must be appropriately connected to ability) as well as a safety condition. And so we see that a theory of knowledge that preserves both the value and luck insights must fall within the description of what Duncan Pritchard (*forthcoming a*) has called *anti-luck virtue epistemology*. I argue, though, that those with a common commitment to placing an ability and anti-luck condition on knowledge can disagree about whether either of these conditions entails the other. This motivates the thought that there will be three distinct kinds of *anti-luck virtue epistemology* (ALVE). I argue that the only one that could plausibly satisfy the Meno Requirement will be a variety of anti-luck virtue epistemology according to which satisfying the ability condition entails that the belief is appropriately *safe*. And so, the only defensible variety of ALVE turns out to be *robust virtue epistemology*—a view that is committed to showing that the ability condition, as it were, does all the work—including insuring a belief’s safety. Drawing from the insight that a true belief depends on luck and ability gradiently (and not rigidly, as some have thought) I go on to develop and defend a type of *robust VE*, one motivated by the thought that a true belief is through ability (and is as such finally valuable) just when the success (truth) of your belief depends on your abilities moreso than luck.

1

The Value of Knowledge and the Nature of Knowledge

1.1 Introduction: The place of normativity in 20th-century epistemology

I want to take as a starting point Jonathan Kvanvig's observation that 'A correct account of the nature of knowledge must resist counterexamples, but it also ought to be amenable to an account of the value of knowledge.' (Kvanvig 2003: 5) This somewhat innocent if not obvious point stands, within the tradition of analytic epistemology, as something of an anomaly. Up until the last decade or so of the 20th century, projects in epistemology were conducted without any notable concern about values—at least not in any way that would have given any legitimacy to Kvanvig's suggestion, one that would require that we countenance value claims (e.g. epistemic evaluations) as the sort of claims we must be sensitive to when offering an account of knowledge.

The previous century's movements, debates and problems stand as a testament to this point. A useful starting point will be to note Quine's call to 'naturalize' epistemology around mid-century. In his paper *Epistemology Naturalized*, Quine writes:

Epistemology, or something like it, simply falls into place as a chapter of psychology and hence of natural science. It studies a natural phenomenon, viz., a physical human subject. (Quine 1994: 273).

On the picture Quine gives us—one later adopted by Hilary Kornblith (2002) among others—the needed amenability of a theory of knowledge to corresponding claims we make about its value would seem an odd requirement at best. Indeed, on such a picture, Kvanvig's claim would amount to suggesting something as strange as supposing that the projects of empirical science are hostage to claims of axiology. Viewing

epistemology, then, as a ‘chapter...of natural science’ would be enough to have made Kvanvig’s point, at mid-century, absurd.

Though Quine’s conception of epistemology fell within a commitment to a rejection of some of the fundamental assumptions (e.g. the distinction between analytic and synthetic judgments) that pervaded logical empiricist tradition that had risen in the decades following the early work of Wittgenstein, Russell and the Vienna Circle, a naturalized epistemology was nonetheless a continuation of the inherited tradition of the logical positivists insofar as it afforded *no* significance (in the theory of knowledge) to considerations of normativity and value.

An important challenge to positivist and naturalizing approaches to epistemological theory was leveled by Roderick Chisholm, whose influential works *Perceiving: A Philosophical Study* (1957) and *The Theory of Knowledge* (1966) were instrumental in having distinguished the project of analyzing knowledge as self-standing and interesting in its own right—one whose questions shouldn’t be thought of as exhausted by those that epistemology happened to inherit as a bridesmaid to the philosophy of science. Chisholm, instead, thought it was important that we can give a correct analysis of knowledge because he was motivated by the view that such an analysis would be necessary for the purpose of *vindicating* external world knowledge against the sceptic in the spirit of G.E. Moore (1939). With this aim in mind, he maintained (contra his predecessors) that epistemology, and knowledge in particular, is best studied *via* investigating epistemic principles.

With the thought, then, that epistemology’s aims be aligned moreso with Moore’s than with the Vienna Circle’s, Chisholm’s principle-based approach to analyzing knowledge constituted two important breakthroughs: first, his preferred principles rationalized the thought that centrally important to knowing is the extent to which our beliefs are justified in virtue of evidence that is reflectively accessible to us; and secondly, this sort of *internalism*, coupled with his related endorsement of *foundationalism* as an account of the ‘structure’ in which beliefs are justified, encouraged that the theory of knowledge be conceived of as one whose central questions are, as Richard Feldman (2007) says, ‘purely epistemological’ in nature, as opposed to (as it was previously thought) fundamentally psychological or empirical. Feldman, whose present evidentialist theory is itself a development upon Chisholm’s internalism, notes even more that:

Chisholm held that epistemologists did not need the assistance of the empirical sciences in answering their purely epistemological questions. (Feldman 2007: 1)⁴⁷

⁴ Although present-day discourse permits the thought that *internalist* theories of knowledge rely on an internalistically construed notion of “justification” that itself captures some kind of normativity (although the sense in which justification here is normative is widely disputed),

Had Chisholm's internalism had the effect of provoking his contemporaries to focus on issues about the relationship between knowledge and the ethics of belief (issues that occupy present-day internalists) then considerations of epistemic normativity might well have seemed unavoidable for framing these worries properly. However, this was not to be. Analytic epistemology in the 1960s was shaped (perhaps) equally by Chisholm's revolutionary approach to the theory of knowledge as it was by Gettier's (1963) watershed paper 'Is Justified True Belief Knowledge?' which was—among other things—a bold challenge to the thought that knowledge could be correctly analyzed as Chisholm would have it: as a true belief supported by internalistically-construed justification. Gettier's paper gave way to several decades—these led by Alvin Goldman (1967; 1976; 1979; 1992)—in which the nature of knowledge was predominantly thought of in terms of causal connections and the reliability of belief-forming processes. The 'post-Gettier' focus upon process reliability, defeasibility, truth-tracking and causal connections, though a move forward insofar as it (among other things) carved out and sharpened our intuitions about why justified but luckily-true beliefs fall short of knowledge, was at the same time a move back to the pre-Chisholmian view that would have us investigate the nature of knowledge by asking questions to which claims about epistemic values would fail to even resemble answers.

We can see then that from the early 20th century and right on through the decades that followed, epistemology's leading figures and movements did not take the theory of knowledge to be something we investigate through any sort of axiological lens.⁵ Indeed, Hans Reichenbach (1951), in *The Rise of Scientific Philosophy*, sums up curtly (at about mid-century) what our brief tour of recent epistemology suggests to have been an accepted sentiment: Says Reichenbach:

...The Modern analysis of knowledge makes a cognitive ethics impossible: knowledge does not include any normative parts and therefore does not lend itself to an interpretation of ethics. (Reichenbach 1951: 273).

Against this background, we have, I think, the right perspective for appreciating the significance of Kvanvig's observation. If an account of the nature of knowledge really must be amenable to an account of the

Chisholm never developed this line of thinking either in his ensuing work on epistemology or value theory. Thus, while his departure from previous conceptions of epistemology made "purely epistemological" notions such as "justification" and "reasonableness" relevant to the project of analyzing knowledge (and thus, made epistemology at least a discipline *open* as it wasn't before to normative assessments and corresponding discussions of value), actual *claims* of epistemic value were themselves not given any development to speak of. Consequently, Chisholm's break from the empirical purview of epistemology toward one whose questions became purely epistemological was not one on which claims about the *value* of knowledge would have any role to play.

⁵ Likewise, specific debates themselves mirrored on this score the wider movements within which they cropped up. For example, debates between foundationalists and coherentists, like the Chisholmian project that gave birth to them, did not stand or fall on the basis of anything to do with value. Neither did the influential debates Sellars inspired having attacked the myth of the given. Other important epistemological problems over the past century—(e.g. Chisholm's problem of the criterion, Moore's proof of the external world, problems arising out of inference and induction, etc.) are in this respect no different from the debates that generated them, or for that matter, the wider movements that generated those debates.

value of knowledge, then value-based claims that a long tradition of epistemology would have taken to be irrelevant or at best peripheral fall squarely within the range of considerations that should determine what we say knowledge *is*. And for the purpose of analyzing knowledge, these considerations are quite important.

1.2 A New Way of Thinking: The Meno Assumption and the Link Between Knowledge and its Value

As Wayne Riggs (2008) has observed, epistemology has undergone a shift in thinking over the past two decades—a shift he’s labeled the *value turn* in epistemology. (Riggs 2008: 3) What’s notable about the value turn is that, after a century of insignificance in the debates and problems of epistemology, the *normative* aspect of knowledge and related epistemic notions is now thought important, and increasingly so.⁶ Importantly responsible for having inspired the recent surge of value-driven epistemology—especially during the past five years that have seen its popularity surge—is the peculiar implications shown to arise from ideas associated with simple idea with which the present chapter began: Kvanvig’s (2003) observation that ‘A correct account of the nature of knowledge must resist counterexamples, but it *also ought to be amenable to an account of the value of knowledge*.’ (2003: 5) Indeed, the influx of literature we can locate within debates about the implications of this observation is remarkable. Especially remarkable, though, is the amount of attention that has been given to a *particular* claim about the value of knowledge, one that epistemologists have increasingly been lead to believe that a theory of knowledge must take special care to accommodate. The claim I have in mind here is the widely held assumption that (to put the idea roughly) knowledge is more valuable than ‘mere’ true belief. What epistemologists are increasingly worried about then is whether their theory of knowledge is ‘amenable’ to the thought that *knowledge*—as their own theory defines it—is something demonstrably more valuable than what their theory counts as a ‘mere’ true belief.

This issue—the issue of whether knowledge should be preferred to true opinion—was first investigated first by Plato himself in the *Meno*, as part of an exchange that is now essential reading for those working on the problem⁷. It is Kvanvig (2003; 2009*a*; 2009*b*), though, who is primarily responsible for reintroducing attention to that particular exchange in the *Meno* and for highlighting the force of a set of puzzles he pulls out of ideas developed in the *Meno* exchange concerning the values of knowledge and true belief. The puzzles actually

⁶ Axtell & Carter (2008) have identified the rise of virtue epistemology in the early 1990s as important in bringing issues of epistemic normativity and value into the mainstream literature.

⁷ It is interesting to note that, prior to Kvanvig’s (2003) work on the value of knowledge, the *Meno* was best known in epistemology for giving rise to the paradox of inquiry, whereas now the dialogue (I suspect) is equally if not more frequently identified as providing the inaugural discussion about the value of knowledge and true belief. But I say this with reservation as a point that, if true, would be so only very recently.

‘become’ puzzles for anyone who insists, as Socrates did, that knowledge should have a value true opinion lacks, and in virtue of which it is something better, comparatively, to have.

Before the puzzles associated with accepting this assumption had become obvious, only Swinburne (1999) and Jones (1997) had, since Plato, seemed to ‘care’ about the implications of thinking knowledge better than true opinion, though their points about this were for the most part overlooked. Kvanvig, however, raised the stakes so that the matter of whether the assumption should be thought true was no longer of academic interest only (as we might assume it was for Plato).

The way that Kvanvig succeeded in challenging epistemologists to care about the this assumption and the implications its endorsement would have for a theory of knowledge owes partially to the fact that (as we might suspect) few epistemologists are prepared to deny it; (and at the least, most all admit that it gets at something quite intuitive.) But this in itself would not carry any call to action were it not for the fact that Kvanvig, in his seminal 2003 *The Value of Knowledge and the Pursuit of Understanding*, argued quite cleverly (and certainly originally) that any theory that tries to preserve this assumption will ultimately fall prey to what he dubbed the *Swamping Problem*. The Swamping Problem (in its capacity as a *bona fide* problem, not specific to reliabilism) was designed to show that none of the prominent theories of knowledge of the day could account for how the condition(s) knowledge (on any popular account) requires in addition to true belief were ones that added any *value* to the true belief itself. And problematically, if a theory of knowledge is one according to which knowledge turns out to be *no more valuable* than mere true belief, then a consistency problem emerges: such an account would be incompatible (i.e. logically inconsistent) with the thought that knowledge has a value greater than true belief.⁸

Following a somewhat popular precedent, call this assumption we’ve been discussing the *Meno Assumption* (hereafter, MA). It’s an assumption whose denial is not without its own problems. To start, were we to try to dodge the puzzles Kvanvig raises by simply rejecting MA, we would be doing so at the cost of denying what is widely taken to be an intuitive if not obvious observation about knowledge. More worryingly though, if MA is not *positively* endorsed, then the reasons we would be inclined to cite when defending the importance we explicitly (or implicitly, by way of our projects) give to knowledge in epistemology would lose a lot of their purchase power. After all, if knowledge is no better to have than mere true belief, then as Duncan Pritchard

⁸ Most epistemologists, with some exceptions (e.g. Jason Baehr 2000), seem to agree with the reasoning behind this observation—at least insofar as they appear worried by the thought that their theory of knowledge couldn’t be exonerated from the charge that it is incompatible with an endorsement of knowledge as more valuable than true belief.

says, the historical focus of epistemologists on knowledge is ‘somewhat mysterious.’ (Pritchard 2008*b*: 1) At the very least, it would fail to be obvious why knowledge is something we should distinctly care about, or even more, why we should as epistemologists be all that concerned to say what it is.⁹

Now Kvanvig himself, after arguing at length that MA is not something a theory can accommodate, simply bites the bullet here and announces his discovery that knowledge was never so valuable as we’ve always thought it was. He suggests that we do the same—that we recognize that the inability of any plausible theory of knowledge on the market to account for why a true belief would be made more valuable by the satisfaction of the theory’s knowledge conditions gives us reason enough to deny MA¹⁰ and perhaps (as Kvanvig has) turn our attention to something that *really* has distinctive epistemic value. In Kvanvig’s book, that would be *understanding*.¹¹

But Kvanvig’s challenge has had the entirely *opposite* effect on mainstream projects¹². Rather than to have engendered a movement away from thinking of knowledge as distinctively valuable, his challenge has motivated epistemologists to endorse this claim explicitly. Nowadays, the *Meno* Assumption has become something that theories of knowledge—particularly, reliabilist and virtue-theoretic theories—have engrafted on the list of other desiderata that a theory of knowledge should be able to accommodate. John Greco (2009*a*) goes so far as to place the problem of accommodating the *Meno* Assumption among what he calls the ‘perennial

⁹ Indeed, if we demur from a positive endorsement of MA, then we give up at the same time any good defense for why we should ever be concerned, from a perspective that’s epistemic, pragmatic or otherwise, whether a state counts as knowledge rather than mere true belief. Without such a defense, we might well face a further problem: we’d have no obvious resources for explaining why the concept of knowledge plays any useful role in language—that is, we would be in a puzzling position when asked to explain the linguistic convention that permits us to use “knowledge” it as a means of picking out something we aim to be distinguishing within our cognitive economy. After all, to borrow from Edward Craig, the concept of a knower is useful for picking out an informant in a way that captures something that the concept of ‘true believer’ seems inept to. I raise these points as a way to dispel the tempting thought that denying MA might be a quick and painless way to sidestep the puzzles associated with accepting it.

¹⁰ Now as it’s turned out, few have accepted Kvanvig’s open invitation that we should, at the end of the day, abjure the *Meno* Assumption and leave behind the spoils of its endorsement—the spoils being entitlements to continue caring about knowledge as others traditionally have and to justify the epistemic worth of the projects we’ve undertaken in analyzing it.

¹¹ Notice that, interestingly, the argument I’ve attributed to Kvanvig here is not one that concludes or (strangely) even suggests that prominent theories of the day are actually false. Nowhere does he try to give a counterexample to the theories he runs through the “swamp”, and nowhere do his arguments depend on the falsity of any particular analysis of knowledge. Nonetheless, leading epistemologists such as Ernest Sosa, John Greco and Alvin Goldman have all found Kvanvig’s argument threatening enough to have ‘defended’ their proposals in some detail against his unique challenge, and I think this is telling. It reflects the point that I’ve claimed represents a monumental break from decades of previous tradition—the very point that Sosa, Greco and Goldman (among many others) have, by defending their view against Kvanvig implicitly assented to: that a theory of knowledge must—in *addition* to being free from counterexamples—also be ‘amenable’ to the claims we make about the value of knowledge.

¹² I qualify this; Kvanvig’s work on understanding has itself been influential and in this respect he’s not had the opposite effect as that which he intended. However, his attempt to repudiate the value of knowledge platitude has probably done more to lead people to defend it than anything else has.

problems' for a theory of knowledge, alongside the Gettier Problem, the Lottery Problem and the problem of radical scepticism.

My intention in having sought to locate historically the problem that the *Meno* Assumption poses for a theory of knowledge was motivated in part by the thought that doing so might indicate the special sense in which it is a high-stakes problem with unique implications. Now, I've at this point admittedly only engaged with the details of the puzzles surrounding the connection between MA and a corresponding theory of knowledge superficially, as was my intention. The next point of order will be to develop this connection and its associated puzzles with more precision. In particular, it will be important to get clear about (among other things) how we should articulate the precise claim we should think of MA as making when we take it to be capturing something intuitive. Further, it's needed that we understand just how it is that philosophers like Greco (2009*a*; 2009*b*)¹³, Goldman (2009)¹⁴ and Sosa (2007)¹⁵ take their theories to be held accountable to this assumption, a concession they make in having defended their theories against the Swamping Problem.

A further point about MA: Kvanvig claims it to be such that our theories of knowledge should be 'amenable' to the endorsement of it, but this does little to help us take the idea seriously. If 'amenable to' is supposed to just pick out a relation of logical consistency, then it should seem that a theory of knowledge must be 'amenable' to not only claims about the value of knowledge, but to any claims that are *true* (full stop). Amenability in this weak sense, then, invites the question of whether there is any special connection between a theory of knowledge and claims about its value as the puzzles inspired by MA imply.

From here, then, we have a perspective from which we should be able to see that the problem arising from the MA occupies a special place in the history of epistemology, and that this is not just because of the attention it has received in the contemporary debate. Beyond that—for reasons considered in 1.1—it was shown why the various paradigms in epistemology over the course of the past century prevented this now-important problem from ever seeing the light of day. In this respect, the restriction MA appears to pose on the theory of knowledge is uniquely a matter that carries with it high stakes and, unlike other problems with a lot riding on them, no long body of literature to accompany these high stakes: barring two articles, everything published has been from 2003 onward. I'll be drawing from most all of these resources in this and the chapters

¹³ Greco makes this claim in a chapter draft on the value of knowledge in his forthcoming *Achieving Knowledge* (2009), a version of which is also forthcoming in Pritchard, Haddock and Millar (2009) *The Value of Knowledge*.

¹⁴ With Erik Olsson, Goldman defends reliabilism against the challenge taken to be posed by the Swamping Problem.

¹⁵ Sosa (2007) takes this problem to be binding in Ch. 4 on "Epistemic Normativity" where he goes on to offer a way out of the problem.

that follow, though as we shall see, there are quite a few gaps in the explanations corresponding to some of the questions we've raised. I'll turn my attention now to these.

1.3 The Three Master Questions

What I've been calling the 'problem' arising out of the endorsement of MA is, in reality, three separate problems that are closely interconnected. I say *three* problems because we are not in a position to understand with any clarity how the MA is one that a theory of knowledge must be 'amenable' to without positive answers to all three questions. While each of the three central questions has been addressed (in varying degrees of sophistication) in the literature on epistemic value, there has yet to be any attempt to deal with them all in any sort of uniformity. Consequently, then, there is to date no complete account of how it is that a theory of knowledge should be held to account for the MA, and this is so despite the fact that (as I've indicated) a substantial number of writers have gone to lengths to show that their theories would not be threatened by an endorsement of it. Unsurprisingly, the remainder of this chapter will explore these three master questions.¹⁶

1.3.1 First Question

Question (i): What is the right way to articulate the motivating assumption behind the MA (the thought that knowing is better than (merely) truly believing)

Pritchard (2008*b*) has hit upon something that should be made clear upfront regarding what it is we ought to be after in our attempt to clarify MA. In reflecting on the problem of explaining why knowledge has a distinct value that true belief lacks, he makes the following insight about the intuition itself:

Notice that the value problem, as it stands, is ambiguous in an important respect. Are we simply seeking an explanation of why, *whether rightly or wrongly*, we regard knowledge as distinctively valuable? Or are we seeking an explanation of why knowledge *is* distinctively valuable which can in turn explain why we regard knowledge as distinctively valuable? (Pritchard 2008*b*: 1)

According to Pritchard, it is the *latter* we're after. This is surely right. After all, insights about knowledge

¹⁶ In order to best engage with them, I offer that we bring on board a device for clarity—one that will help us to organize what we're investigating. Think of what it is we're interested in as a three-place relation; there is a (thus-far underdeveloped) 'is amenable to' relation, and the idea is that a theory of knowledge and a particular assumption about its value stand to each other in that relation. To clearly understand this relation, we should have a grasp of each of its three components. It should now be a point of obviousness that of the three master questions, each will stand connected to one of the three components. Taking the theory component first: it should be pointed out that none of those aspects of the relation suggested in the previous section to be 'mysterious' owed to any confusion about what it is that must be 'amenable' to the MA. So 'theory' is one part of the three-place relation that, by clarifying it, we'll have failed to gain any traction. It's the other two components that *really* need unpacking. Questions (i) and (ii) pertain strictly to the MA while Question (i) concerns the nature of the accountability in question. I'll address each of separately though with a more extended focus on the first two as I think they invite the most confusion.

and its value are ones that our theories should explain and preserve only to the extent that they are *correct* insights. Indeed, one reason in favor of thinking MA captures a correct insight is that it is widely endorsed. But this is a motivation for figuring out what's correct about the insight—and upon doing so, we may then articulate the claim that reflects *that* accordingly. And so it's important that we take Pritchard's warning and distinguish this task from the task of simply giving an accurate representation of what people seem to think and nothing more.

Variations of Different Strengths

Answering what I'm calling Question (i) is something Kvanvig (2009*a*; 2009*b*) has given a lot of recent thought to, and this is for good reason. Greco (2009*a*), among others, has pointed out that Kvanvig was not consistent in his 2003 in the way he presented the assumption throughout the book. A particular problem was that Kvanvig wavered between articulations of the assumption that were themselves of varying degrees of ambition. For example, he started things off by considering whether theories could be demonstrably amenable to the claim that knowledge is more valuable than true belief but later shifted to pose the question as though grounded in the assumption that knowledge is more valuable than any proper subset of its parts.

The two problems are different. Duncan Pritchard (2007*a*) calls the problem he takes to be associated with the first assumption (comparing knowledge and true belief) the Primary Value Problem and the latter problem, which compares the value of knowledge and any proper subset of its parts the Secondary Value. At points, Kvanvig seems like he has in mind what is perhaps an even stronger version of the assumption, which Pritchard calls the Tertiary Value Problem—the problem Pritchard takes to correspond with the assumption that knowledge has a *distinctive kind* of value over and above any proper subset of its parts—a value that is not merely one captured in terms of degree. Greco (2009*a*) reads Kvanvig as trading in his book on all three of the assumptions Pritchard specifies in distinguishing his three distinct value problems, though it's not clear to me that Kvanvig commits himself anywhere to the third reading (at least, on any charitable reading).

Regardless, it's important that there is no ambiguity about the ambition of the *Meno* Assumption insofar we are taking seriously the idea that a theory should have to account for it. I'll set aside presently the two stronger versions of the assumption Pritchard has developed and focus on the first. It is, after all, the weakest, least controversial and most widely endorsed of the three assumptions and so we should think that it, to a greater degree than the others, hits at something right.

*Pragmatic, All-Things Considered and Epistemic Value**(A) Practical value and the Larissa exchange*

The MA admits of another ambiguity, though, which trades on a distinction between *practical value*, *all-things-considered value* and *epistemic value*. It's quite important that these values not be conflated. But even if we grasp the important difference, there remains the issue of determining which of the three distinct values frequently invoked when discussing the Meno Assumption is apposite to the idea the MA should be thought of as capturing.

First, then, let's rid ourselves from whatever inclinations might have us *conflate* the relevant kinds of value at issue. We can begin by considering that distinctions in value are familiar to us (social, pragmatic, aesthetic, religious, epistemic, etc.). Equally familiar is the thought that a given thing might generate positive value that is of one sort and negative value of another. For example, imagine a painting by a skilled artist depicting some grotesque act of gratuitous blasphemy. This painting might generate negative religious or political value while at the same time maintaining some positive, purely aesthetic value. The all-thing-considered value of the painting would just be (on the most basic accounts) a function of the total positive and negative values generated. In the case of knowledge, then, the all-things-considered value will be a function of both pragmatic value, epistemic value and other values.

An equally simple way to assess whether something has value of a particular sort will be to employ a (modified here) test Kvanvig (2008*b*) endorses: if we want to assess the purely epistemic value of something, we simply control for all other values.¹⁷ It should be obvious that Kvanvig's thought experiment is an imperfect one, however, it at least provides an organized and I think for the most part correct way of drawing distinctions between pragmatic and epistemic value and as such is useful as a model for judging what sort of value is at issue in the Meno Assumption.

Now that we have a working idea how to go about distinguishing different kinds of values, let's move on to consider which kind of value best captures the preference we hold for knowledge over true belief. Some writers, including Goldman (2009), Olsson (2007) and Ahlstrom (2006), have written in a way that betrays that

¹⁷ For example, the purely epistemic value something has would be determined by considering the value it would have in a world in which the possibility of non-epistemic (i.e. practical, etc.) value is controlled for—for example, a world in which (i.e. affective) states that bring about the possibility of practical values are controlled for. In short, something would be purely epistemically valuable only if it would continue to be valuable even if you didn't have any desires or cares at all. Similarly, purely pragmatic value, on this view, will be value that something has in a world where everything is controlled for except our interests

they suppose the assumption that their theories of knowledge must account for is one that requires only that they show how, on their theories, knowledge is more practically useful than mere true belief. These authors then are thinking that the MA is an assumption about practical value. Rene Van Woudenberg (2006) and Jason Baehr (2009) also view the assumption this way.¹⁸

This raises then an important question: does the discussion of practical value present in the Meno dialogue motivate a pragmatic reading of the value assumption we've set out to articulate? Let's skip to where the conversation gets good:

Socrates: ...If a man knew the way to Larissa, or anywhere else, and went to the place and led others thither, would he not be a right and good guide?

Meno: Certainly.

Socrates: And a person who had a right opinion about the way, but had never been and did not know, might he be a good guide also, might he not?

Meno: Certainly

Socrates: And while he has true opinion about that which the other knows, he will be just as good a guide if he thinks the truth, as he knows the truth?

Meno: Exactly.

Socrates: Then true opinion is as good a guide to correct action as knowledge...¹⁹

Meno: I admit the cogency of your argument, and therefore, Socrates, I wonder that knowledge should be preferred to right opinion—or why they should ever differ. (Jowett: 26-27)

Notice here Socrates' final point and how it causes Meno's mind to wander dangerously²⁰: the point is that knowledge is an equally good practical guide as true belief. Both get you to Larissa just the same. From here

¹⁸ Both Van Woudenberg and Baehr appeal to Plato's *Meno* itself, pointing out that the most famous and oft-cited issue within that particular exchange¹⁸ turned on whether knowledge or true belief was a better practical guide. The Larissa example from the Meno was indeed the starting point Kvanvig used for his 2003 book, although Kvanvig himself has always insisted that the sort of value at issue in the assumption that theories aim to account for is purely epistemic, a view Pritchard has also suggested in a recent paper on the Swamping Problem.

¹⁹ The fact that Meno has here made an inference about the nature of knowledge from a claim he believed about that value of knowledge is especially interesting, as Kvanvig has noted. However, the sophistication of this point brings up issues outside the present focus.

²⁰ Of course, Meno's subsequent musing betrays an illicit move in his reasoning: he reasons that if *A* is no more valuable than *B* insofar as *A* and *B* are practically useful then (perhaps) *A* is no more valuable than *B* (full stop). This would be a valid inference only if the value of something all-things-considered were reducible to the value the thing has as a practical guide. But our earlier discussion should be enough to spot this as a mistake. Meno's misstep here is frequently (and unconsequentially) overlooked though because of the fact that, through his illicit reasoning, he has done us all the service of stumbling his way to raising the *very interesting* question of whether knowledge should be preferred

Meno speculates that knowledge might not be better to have than true belief after all, or even different from it for that matter.²¹

Socrates at this point is a champion of common sense. He claims to know that Meno is wrong *both* to suppose that knowledge is no different from true belief and to suppose that knowledge is not rightly preferable to true belief; knowledge is, Socrates insists, 'more honourable and excellent' by comparison. (Jowett: 27) He then boldly proclaims that he knows these things if he knows anything at all (how unlike him!)

So what can be drawn from this glimpse back to the Greeks? Pritchard's (2007*d*) suggestion is one that draws from both the question Meno invited and Socrates' instinctive response. He says:

Given that we clearly do value knowledge more than mere true belief [Socrates' point], the fact that there is no obvious explanation of why this should be so [The point Meno stumbles to] creates a problem. (Pritchard 2007*d*: 1)

Pritchard (2007*d*) calls this the *Meno* Problem. At this stage though we are concerned specifically about the nature of the assumption and not the problem (which he takes to be first and foremost a problem of explanation). So where do we find an instance of the assumption that we think a theory of knowledge should be held accountable to? On a closer inspection it's not clear how anything within the discussion about getting to Larissa constitutes the sort of platitude that fits the profile of the one we're concerned with: one that is both very compelling and such that Kvanvig would say you can't preserve it. Look as we might, it isn't there.

The assumption we looked to the dialogue to clarify and sharpen—one that is highly intuitive and one that Kvanvig denies our theory can preserve—is actually located in Socrates' instinctive, mysterious and somewhat dogmatic response rather than in any of their respective observations about getting to Larissa. The assumption was also importantly one that we should note Socrates held *in the face of* Meno's observations about there being no difference in practical value (or value full stop) between knowledge and true belief.

My reasons for pulling out the dialogue should now be made a bit clearer. First, I want to stress that MA—in the sense that we've been concerned with it from the outset—isn't necessarily supposed to pick out anything Meno or Socrates said or assumed, and so we must resist whatever temptation there might be to think that the 'real' Meno assumption lies somewhere in the spirit of the text. Quite recently, *Meno* has become

to true opinion *at all*. Though an interesting question in its own right, it is especially so in this context given that Meno has just laid down a restriction for anyone who would answer the question in the affirmative. The restriction is that, if knowledge should be preferred to true opinion, this wouldn't be (supposing that we grant Socrates' point) because it's a better practical guide.

²¹ While Alan Millar and Timothy Williamson (2000), Erik Olsson (2007) and Goldman & Olsson (*forthcoming*) demur here, I see no reason to and have argued this point elsewhere (see Chapters 2 and 6).

synonymous with the value of knowledge debate²² and is used to describe positions in that spirit. We could get on just as well referring to it as the Value Assumption.

What's (at this point) *essential* to the claim is only that it represents the intuitive idea that the value of knowing exceeds that of merely truly believing. The purpose of exploring the *Meno* text was thus with the intention of diffusing what we saw was a very common mistake of thinking that the claim that a theory of knowledge should be held to account for is a claim about comparative practical value simply because there are considerations involving practical value present in the topic's inaugural debate.

At present, it's useful to see that the value problem, when posed in the practical context of the Larissa exchange, would mislead one to think just this—and in doing so, overlook what the challenge has really become. Meylan (2007) sums up this point nicely here in her observation that:

Plato was concerned by the different *practical* value of knowledge in comparison to true belief. The contemporary debate focuses on the distinctive *epistemic* values of knowledge and true belief. This means that a contemporary solution will try to appeal exclusively to epistemic values...and not to practical values...to explain why it is better to possess knowledge than mere true belief (Meylan 2007: 86)

From this Meylan observes, therefore, that:

...the contemporary debate is not perfectly pointed by our initial question but more accurately by the following: (Q) Why is it epistemically better to know that p than to possess a mere true belief that p? (Meylan 2007: 87).

If this is right, then have we learned nothing about the *Meno* Assumption from the *Meno* text other than that Socrates was prepared to assert with conviction the very assumption we're out to articulate more clearly?

Perhaps this is too quick. The dialogue does at least invite us in the right direction. Kvanvig (2009a) writes:

When it is pointed out that true opinion gets us all the practical benefits that knowledge does, a quite natural response is to insist that practical value isn't the kind of value one has in mind when one claims that knowledge is more valuable than mere true belief. (Kvanvig 2009a: 2)

If the *practical* value reading is indeed undermotivated, then does this count in favor of Meylan's suggestion that the MA be thought of as a claim about *epistemic value* (which she is taking to be a mark of the contemporary debate), or is an *all-things-considered* interpretation equally as viable? Prima facie, neither stands out simply in virtue of the fact that the pragmatic account failed. And so an independent argument will be required; I'll now offer such an argument, one that will show that an epistemic-value reading of the assumption is the only

²² See Pritchard (2008b, 2009a), Kvanvig (2003, 2009a; 2009b), Riggs (2005), Carter (2008), Meylan (2007).

plausible candidate. This will become evident though only after we first consider carefully the reasons why an *all-things-considered* reading fails.

(B) *Should MA be understood as a claim about all-things-considered value?*

Bearing in mind that we're out to clarify MA insofar as it should be picking out a *correct* insight about knowledge and truth, a point of note will be that whether MA will appear to capture something correct in a given case can depend on how we're qualifying the value in question. And so we want to know which interpretation picks out a *correct* insight, and in this section, we'll examine whether an all-things-considered interpretation of the value at issue in MA does just this.

I'll begin with an analogy that takes as basic the notion of preference. Just as, in the cognitive arena, knowledge is something we prefer to mere true belief, we are similarly inclined to prefer, physically speaking, strength to weakness. The task of articulating the insight about knowledge and true belief that would explain this preference would be akin to the task of articulating the insight about strength and weakness that would explain *that* preference. That insight would be of the same structure as MA and thus would be:

(S) The value of strength is greater than the value of weakness

Now, let's assess the plausibility of the *all-things-considered* reading of (S). As the *all-things-considered* reading would be a way to read the evaluative claim made by the *comparison* itself, the application will be wide-scoped rather than narrow-scoped.

(S*): *All things considered*, the value of strength is greater than the value of weakness.

If S* captures what we should think *correct* about strength and weakness that rationalizes our preference of the former to the latter, then plausibly, this would count in favour of an *all-things-considered* reading of MA as well.

That said, S* does seem to capture something true in the case of Achilles, for example. Achilles was strong, and we're inclined to think that, given his noble successes in war, this strength was something valuable—something such that, once everything else is taken into account as concerned his strength and what he accomplished with it, weakness would have been a (far) less valuable trait. S* clearly accommodates what we take to hold true about Achilles. But our preference of strength to weakness is grounded in the idea that strength wouldn't just be better than weakness for Achilles, but that it should also be preferred to weakness by

those would rarely if ever need to exercise it. Plausibly, then, the thought that strength should be preferred to weakness is a thought for which our endorsement is *unreserved*: we think that strength is something that *anyone* should prefer to weakness, and it is in this spirit that we uphold the value of strength over weakness.

That said, imagine that we are in a world that is exceedingly corrupt—a world where everyone seeks to use her advantages to malicious ends. In such a world, S^* is false; strength in such a world, *all-things-considered*, generates significantly more negative value than weakness. This we may suppose to be a contingent fact that holds true for all individuals who possess strength. Now, there is a notable tension then between (i) the fact that our endorsement of the value of strength over weakness is one we hold freely without reservation and (ii) the thought that the way to articulate this endorsement would be to qualify it as a claim about *all-things-considered* value. If the *all-things-considered* reading were the correct one, then it's unclear why the thought that strength is more valuable than weakness is one we endorse in the same spirit that we think that, *for anyone*, strength would be better than weakness.

Different things might be concluded from this tension. First, let's consider what we might conclude if we continue to maintain that S^* is the right articulation of our original endorsement of the value of strength over weakness. Holding that fixed, we might draw from the recognition that, all things considered, strength is not more valuable than weakness in worlds such as our malicious world, that therefore, S^* must be false, and if S^* articulates our original intuition, then consequently, our original endorsement of the value of strength over weakness must have been mistaken. A consequence of drawing this conclusion would be that we should suppose that strength is less valuable than we'd originally suspected. But this is not an inference we are led to make at all.²³

Putting this all in perspective: the sense in which we value strength over weakness is one expressed by an endorsement that is unreserved, and this explains why the case considered does not compel us to revise our original endorsement or think that strength is any less valuable than we had supposed. This means that the insight that the value of strength is greater than that of weakness is *not correctly captured by an all-things-considered reading* of the claim, a reading that the example *does* require that we revise (and indeed reject as false).

²³ The mere recognition that strength could be put to bad ends should not require that we revise what we originally thought about its value over weakness. After all, the possibility of a world in which all advantages are used for ill purposes would also be one in which we could maintain that knowledge provides some cognitive advantage over true belief only if granting that, all things considered, knowledge will (in virtue of being an advantage in this world) generate comparatively greater negative value than would mere true belief. But reflection on the possibility of such a world does little by way of changing our original intuition (MA) about the value of knowledge.

By analogy, we can draw the same conclusion for MA. The intuition that knowledge is more valuable than mere true belief is expressed by an endorsement that is similarly unreserved, and this explains why the recognition that in some worlds knowledge (compared to true belief) generates considerable negative value doesn't lead us to think of knowledge as any less valuable as compared to true belief than we did before, *even though* it requires us to reject an *all-things-considered* reading of MA. Correspondingly, then, this means that the insight that the value of knowledge is greater than that of truth is not correctly captured by an *all-things-considered* reading.

Leaving behind then the thought that the *all-things-considered* interpretation does justice to MA, we face a question: why, after all, did a case that depicted a circumstance under which knowledge was not more valuable than mere true belief appear to be *incompatible* with the all-things-considered interpretation of MA while, at the same time, it *did not* appear to threaten MA in whatever sense it is that we seem to be endorsing it? After all, nothing said to this point licenses us to assume that the scope of MA is supposed to be restricted, and indeed, the fact our endorsement would be offered unreservedly (i.e we endorse it without first worrying about who it applies to) speaks to the idea that we intend it in an unrestricted sense. So *why* did an apparently unrestricted claim not appear threatened by an example that would appear to have posed a direct challenge to such a scope. That is, the malicious world example would seem to require that we think of MA as a claim restricted to worlds where advantages are not used for malicious ends. But it *does not* have this effect. And further, the spirit in which we endorse MA without reservation in the actual world is one in which we'd be happy to allow that at least some people would be inclined to use knowledge more so than they would mere true beliefs to malicious ends.

An explanation for this would have to be one that allows for the value we have in mind to be unrestricted in scope, defeasible, and at the same time not undermined by these instances of defeat. These considerations point directly to what Kvanvig defends as a *prima facie* reading of MA, wherein the value at issue is *prima facie* value as opposed to *all-things-considered* value. Because there is common confusion between the concepts of *prima facie* and *pro tanto*²⁴ (especially as these descriptions may apply to value, as opposed to reasons), I want to be precise about what *prima facie* value is supposed to be (in the context that Kvanvig has intended it, and I'll be intending it) so that there will not be confusion. To this end, consider what Kvanvig says about pain:

²⁴ *Prima facie*, as opposed to *pro tanto*, value better characterizes value that is (though unrestricted) defeasible for the reason that value can be defeated by being overridden *or* undercut. The notion of *pro tanto* value, though of the sort that could be defeated by being overridden, could not be undercut. On this point, see Shelly Kagan (1989): 17, also for discussion on undercutting defeaters, see Kvanvig (2008b).

Pain is a bad-making feature of a case, always and everywhere, but sometimes it occurs in circumstances where its intrinsic badness is overridden or defeated by other factors. We often use the term ‘prima facie’ to capture this defeasible character of the badness in question. (Kvanvig 2009a: 2)

Pain is prima facie, rather than all-things-considered, bad. And like the badness of pain, the value of knowledge over truth (as captured by MA) is not restricted, even though, like in the case of pain, the prima facie value at issue can be overridden or undercut; this is just to say that the value at issue is *defeasible* as it applies unrestrictedly. Accordingly, there will be cases in which, all-things-considered, the value of knowledge does not exceed the value of true belief. In these cases—such as our case of the world of malicious advantage—Kvanvig points out that:

The prima facie value (of knowledge over true belief) *does not disappear* in the presence of untoward consequences, but rather, is defeated by these other factors. When the story is told about the all-things-considered disvalue of knowing that specific claim, the story includes the prima facie value of knowledge and the fact that it is overridden by the negative practical consequences engendered by such knowledge. (Kvanvig 2009a: 3, *my italics*)

I should also dispel what might be a point of confusion by considering how a prima facie reading of MA stands in a different relation to defeaters as do claims that hold about most, but not all, things of a kind. Take, for instance, the claim that *All swans are white*. Intended to range over all swans, this claim would be undermined by the presence of a black swan. But unlike a black swan, which was always black and so did not have whiteness to be “defeated” (i.e. it wasn’t white and then painted black) in the first place, knowledge does possess a positive value over mere true belief in the world of malicious advantage, but because the knowledge was but to ill ends, that positive value was defeated by the overriding negative value generated once used for such ends. This point captures a familiar idea: those features that could make an otherwise good man better (i.e. wit, charisma, intelligence, power) can make the bad man all the worse, and yet this fact would not lead us to value wit any less over stupidity, or charisma any less over repulsiveness.

(C) *MA as a prima facie claim about epistemic value*

An argument for reading MA as a claim about *epistemic value* has not yet been independently motivated, although we shall see now why it is that our understanding of MA as a *prima facie* claim will make obvious that the type of value MA is concerned with is epistemic. First, consider that the fact that, all things considered, the value of knowledge can be defeated in cases does not lead us to think that knowledge is any less valuable in any way that would encourage us to revise our intuition about the value it has in comparison to true belief. This tells us that a reading of MA will correctly capture our insight only if it is a reading that specifies the kind of value at issue in

MA as being whatever value knowledge *always* has over mere true belief, including those cases in which this value is ultimately defeated by other negative values.

Here an analogy to beauty is useful. The claim that the value of beauty exceeds that of ugliness, like MA, should be read as *prima facie*. Worlds in which the beautiful are tortured would be ones in which all-things-considered it would be better to be ugly, but in these worlds, the thought will not be that beauty lacks any value over ugliness, but rather, that whatever value beauty has over ugliness is a value that is overridden by the negative practical value associated with being beautiful. Now, let's investigate what kind of value beauty has over ugliness in such worlds that stands to be overridden, all things considered. Given that we distinguish between types of value—for example, moral value, practical value, religious value, etc.—the notion of aesthetic value should be a familiar one. When we go about assessing the aesthetic value of something, we take ourselves to be evaluating it from the point of view from which it is only *aesthetic considerations* that matter, and nothing else. We are, as it were, holding other considerations fixed. From a *purely aesthetic point of view*, the value of beauty exceeds that of ugliness, even in worlds in which the beautiful are tortured. With this observation, we are in a position to give a correct articulation of the intuition that the value of beauty exceeds ugliness. We intend this claim to be *prima facie* in the sense that we accept that in some cases the all-things-considered value of beauty can be overridden while, at the same time, the intuition we want to capture is nonetheless an unrestricted one—one that reflects that we think there is *some* sense in which beauty always has some value over ugliness. The correct way to articulate the insight we have about the comparative value of beauty and ugliness then would be by way of a *prima facie* claim according to which the relevant kind of value at issue is aesthetic value.

It is a natural move then to see that the widely-held insight that the value of knowledge exceeds that of mere true belief is one we can capture only by (i) distinguishing between all-things-considered and *prima facie* value; and then (ii) distinguishing purely epistemic value from other values that we hold fixed. In doing this, it becomes clear that the ubiquitous preference for knowledge to true belief, and our thought that this preference would be appropriate for anyone, rationalizes an articulation of MA that is *prima facie*, defeasible, and that the kind of value we have in mind in taking the claim to be unrestricted is purely *epistemic value*.

1.3.2. The Second Master Question: Is the claim made by MA true?

Our second master question asks that we determine whether what I've defended in the previous section as the correct reading of MA is itself *true*. If it's not true, then this will be a game-changer. Indeed, if the assumption legitimizing our preference of knowledge to true belief turns out to be false, then it will not be included within

the set of claims a theory of knowledge must account for. And then it simply wouldn't matter whether we can explain why, on our theory of knowledge, known true beliefs are more valuable than unknown true beliefs—and even further, much of the work aimed at showing this loses its relevance. So, if we think of the previous section as clarifying the right way to capture the intuition so widely held about the value of knowledge, we could think of this section as the task of simply figuring out whether this widely-held insight is true.

A useful way to evaluate whether it's true will be to consider arguments epistemologists have leveled against it. The arguments I'll consider take aim at cruder expressions of MA than the more specific expression I argued in the last section to be correct. Naturally, then, rejecting these arguments will sometimes require that we appeal to considerations about *prima facie* and *epistemic* value to dissolve the worries. But this won't always be the case. As we'll see, some of the reasons offered for rejecting MA are quite straightforward and don't stand or fall on the matter of whether the value at issue is epistemic or the claim a *prima facie* one. We shall turn our attention now to these.

John Greco and the over-intellectualization argument

In his forthcoming monograph *Achieving Knowledge*, John Greco (2009a) casts doubt on the thought that MA—even in the broadest sense that we might interpret it—captures anything our ordinary thinking speaks to. Here's Greco:

Put more generally, the issue is whether ordinary thinkers assume that knowledge is more valuable than any subset of its constituents, whether this is understood in terms of the JTB+ account or in terms of some other account. And this is what I am saying is dubious. Ordinary thinking, I am suggesting, contains no convictions about that issue. (Greco 2009a: Ch. 9)

First, we should note that Greco's worry is directed only to the assumption motivating what Pritchard called the 'secondary' value problem and not to the assumption we've been concerned with—one which compares the value of knowledge and mere true belief only. It would be a serious mistake, though, to think that this point would be enough to warrant us casting Greco's worry to the side. Pritchard gives a compelling argument for why this is. His reasoning takes as a premise a thought that I shall not contend—which is that ordinary thinking licenses us to suppose that knowledge is of *distinctive* value to us. On Pritchard's view, the assumption that knowledge is distinctively valuable motivates all three versions of the value problem (Primary, Secondary and Tertiary) that were brought up in the previous section. Now, Pritchard will agree that the *weakest* articulation of the assumption that knowledge is of distinctive value is captured by the insight that it is preferable to true belief.

I argued in Sec. 1.3.1 that the best way to understand our central issue—of clarifying how a theory of knowledge might be held to account to our claims about its value—will be to focus on the weak formulation on

which knowledge is taken to be preferable to mere true belief. After all, it's the formulation most likely to express a claim a theory of knowledge *really must* be amenable to. Now that said, it nonetheless stands that our weaker version of MA (as it concerns specifically knowledge and true belief) is something we assent to because we *do* think of knowledge as something distinctively valuable to us. And as Pritchard points out, an explanation for why knowledge stands to be more valuable than mere true belief will be one that explains why knowledge is distinctively valuable to us only if it could also explain why knowledge would be preferable to any proper subset of its parts which themselves fall short of knowledge. This seems right when we reflect on the fact that 'mere' true beliefs consist simply in beliefs that are merely true but not known. Among the set of merely true but not known beliefs will be true beliefs that are justified and unjustified to varying degrees. Pritchard (2008*b*) infers from this that:

The primary value problem thus naturally leads to the secondary value problem, and it seems that both will need to be answered if we are to account for the distinctive value of knowledge. (Pritchard 2008*b*: 3)

Because we should expect that a theory of knowledge that is 'amenable' to claims about its value should be able to provide some explanation for *how* it is (this will be a matter discussed more in the next section), we should appreciate Pritchard's point that the project of explaining our weak assumption collapses into the project of explaining a stronger formulation of it (the one motivating the Secondary Value Problem). It follows then that *even though* MA as we're understanding it is expressed in a way that specifies only true beliefs (and does not specify justified, true but not known beliefs—which are themselves squarely within the range of proper *subsets* of the constituent parts of knowledge) as what it is the value of knowledge is taken to exceed, an argument such as Greco's that suggests the stronger assumption (*pave* Pritchard call this the Secondary Assumption) to be illegitimate also threatens the primary assumption we're interested in—indeed, both are products of the guiding thought that knowledge is of distinctive value to us.

With this much said, let's revisit the point Greco makes with a perspective from which we can rightly appreciate the threat it poses to us.

Greco's argument

- (1) Ordinary thinking does not carry with it any conviction that the value of knowledge exceeds any subset of its constituents (Greco 2009*a*: Ch. 9).
- (2) A theory of knowledge is not required to accommodate assumptions about knowledge that are not countenanced by ordinary thinking (The idea from which Greco reasons that a theory of knowledge need not

accommodate the claim in (1).²⁵

(3) A theory of knowledge is not required to accommodate the assumption that knowledge exceeds any subset of its constituents (from 1 and 2).

And given Pritchard's point that the guiding thought that knowledge is of distinctive value to us will leave an explanation for the weaker assumption (MA) satisfactory only if it also serves to explain the assumption Greco specifies in (1), it follows that:

(4) The assumption that the value of knowledge exceeds that of mere true belief is one that a theory would accommodate satisfactorily only if it also accommodates the assumption in (1).

And from (4) and (2) we get:

(5) Therefore, ordinary thinking does not contain the conviction that the value of knowledge exceeds that of mere true belief.

To see how we might resist this conclusion, let's consider carefully (1) in Greco's argument, as it stands to be the premise that carries with it the most threatening implications:

(1) Ordinary thinking does not carry with it any conviction that the value of knowledge exceeds any subset of its constituents. (Greco 2009*a*: Ch. 9)

Greco's defence for (1) is quite simple. The argument goes: (i) If (1) is true, then ordinary thinking has convictions about whether the value of knowledge exceeds the value of Gettiered true belief; (ii) Ordinary thinking does not carry with it such convictions; (iii) Therefore, ordinary thinking does not carry with it convictions about whether the value of knowledge exceeds the value of its parts.

There is a sense in which this sub-argument is compelling. However, it will be compelling to us only if we fail to recognize the important sense in which we are *not* always able to provide a descriptive account of the

²⁵ A subtle point about (2): As we said in considering Pritchard's point in Sec. 1.2.1, the interesting sense in which MA is important is not one that owes simply to the fact that some majority of folk talk as though it's true. The majority might be wrong. We must proceed carefully then and reason this way: if knowledge is widely preferred to mere true belief, then this is certainly an indicator of its truth, and *if it is true* then it is argued that a theory must be amenable to it. Thus, (2) of Greco's reasoning would have to be changed to reflect that a theory should be held to account for only *correct* convictions about knowledge—to which Greco could add that, if ordinary thinking doesn't hold convictions on the point in question, then this would appear to undercut a central reason we suspect it gets at something *correct*, and *ipso facto*, would undercut a motivating reason for thinking that a theory of knowledge should be held to account for it. With that clarified, the idea becomes as follows: if ordinary thinking has no convictions on the matter of whether the value of knowledge exceeds true belief [the claim in (5)], then this will undercut significantly the motivating reasons we have for thinking that a theory of knowledge must be held to account for it (i.e. be at least amenable to it).

objects of our convictions, and when we can't, that fact itself wouldn't threaten *what it is* that our convictions are about. Ordinary folk, for example, might whilst at the casino have convictions about the gambler's fallacy or Dutch book arguments even if they lack the ability (or conceptual competence) to articulate these convictions in the way that, for example, a seasoned Bayesian would. Similarly, it wouldn't follow from the fact that ordinary thinkers lack convictions about Gettier cases under any sophisticated description that ordinary thinkers lack convictions about Gettier cases, and these are the convictions we're ultimately interested in when purporting to gauge folk intuitions. It is, after all, in part because the folk are thought to have philosophically innocent intuitions (unclouded by preconceived philosophical opinion) that we look to them.²⁶

Jason Baehr and the Argument from Scope

While Greco's intellectualization argument took aim at MA in a way that didn't require us to distinguish between *prima facie* and *all-things-considered* value or, for that matter, to appeal to the point also made in 1.2.1 that MA is fundamentally a claim about *epistemic* value, Baehr (2009) has offered an argument against MA that will require that we make use of these distinctions.

What Baehr has claimed is that we're flat wrong to suppose that the insight driving MA is anywhere near as strong as most writers in the contemporary debate have taken it to be (Baehr 2009: 1-3). His argument begins with an interesting inference: Baehr infers from the fact that epistemologists frequently try to avoid Kvanvig's Swamping Problem that they must be holding an antecedent commitment to thinking that an endorsement of the value of knowledge over true belief is itself an endorsement *strong enough* (whatever that may

²⁶ Despite the fact that Greco's reasoning fails to support his conclusion, we would be remiss not to take his conclusion at face value and try to offer some reasons for thinking it's false. The simple way to do that is to show that ordinary thinking carries with it a preference for knowing to having what *epistemologists* are inclined to refer to as Gettiered true beliefs. I think the point can be made in a rather straightforward way, and by appealing to the sort of cases in which the ordinary thinkers familiarly find themselves. Most any person (in this case, call our hero 'Joe') you ask will tell you that whether his mother loves him is something he cares about, and that it is something he *knows* to be true. He'll further deny that this is a matter on which he might easily have been mistaken. Suppose we point out to Joe that, over the last two years, his mother was unbeknownst to him taking a medication that was found to have quite dangerous side effects—80% of those who took that medicine suffered a severe lack of empathy coupled with devastating memory loss. Were his mother to have manifested these side effects, then plausibly, she would have lacked (both emotionally and recognitionally) the capacity to love Joe. Fortunately, though, we tell him that his mother did not experience these side effects. We point out further to Joe that, consequently, his claim to know his mother loved him was false; he had what epistemologists call a Gettiered true belief and what we explain to him as a true belief that he was justified in holding but which could have easily been false. Setting aside (to avoid conflating what's at issue) whatever convictions Joe holds about the importance that his mother is well, a natural conviction of Joe's will be that it would have been better to have *known* whether his mother loves him than to have been in the situation he was—one in which his claim to know was not safe. Because ordinary thinkers presumably hold a great deal of true beliefs about the world, and on the matter of *whether* these true beliefs *are* true is a matter that ordinary thinkers will say they care, the point made in the case of Joe generalizes to the convictions ordinary thinking has about the value of knowing—indeed, ordinary thinking endorses a preference for the epistemic benefits of knowing to the recognizably lesser benefits of justified correctness that itself narrowly avoids mistake.

be) to be such that, by endorsing it, we'll have placed a formal constraint on whatever analysis of knowledge we give. I think this observaton is correct.²⁷

From here, Baehr argues that the popular position is—as the assumptions made by the reliabilist convey—to suppose that the guiding insight about the value of knowledge meets this needed strength. He expresses his central worry as follows:

Again, if the guiding intuition were just an intuition to the effect that knowledge is *sometimes* more valuable than true belief, or that certain *kinds* or *instances* of knowledge are more valuable than true belief, or that knowledge is more valuable than true belief on account of certain features that some kinds or items of knowledge might *lack*, then it would fail to indicate that knowledge is always or categorically more valuable than true belief, and thus would fail to motivate a constraint on an *analysis* of knowledge, which again is an attempt to specify the *necessary* (and sufficient) features of knowledge. (Baehr 2009: 4)

Though the precise *nature* of the constraint at issue is not one that we've considered beyond the point that—by Kvanvig's notion of 'amenability'—we mean at least that a theory must be logically consistent with the corresponding claims we make about its value, we can set aside any sophisticated worries about this here, as I shall be raising them in the next section as they stand at the focus of our third master question. A lack of any sophisticated account here doesn't matter for our present point, which can be addressed just as well by understanding Baehr's notion of formal constraint as *at least* a constraint of logical consistency. Baehr's point can then be recast as follows: a theory of knowledge should be required to be logically consistent with the claim that the value of knowledge is greater than that of mere true belief only if knowledge is 'always or categorically' more valuable than mere true belief.

If we let K =knowledge and TB =mere true belief, the point of Baehr's reasoning is that the guiding intuition $V(K) > V(TB)$ could be read in the strong sense in which a theory of knowledge would be constrained by it—a reading according to which:

A. (CONSTRAINING): $V(K) > V(TB)$ functions as a universally quantified claim which could be expressed: $\forall p [(V(Kp) > V(TBp))]$

²⁷ After all, when the reliabilist, for example, takes labour to show (as have Goldman & Olson (2006), Olson (2008), etc.) that reliability adds value to true belief, their stated intention for doing so premised upon the thought that this is something their theory *needs* to show (a matter on which reliabilism has been challenged on the grounds that it *can't* (e.g. Zagzebski (2002), Kvanvig (2003), Brogaard (2006), Meylan (2007)). When reliabilists take themselves as required to *vindicate* the value of reliabilist knowledge over that of true belief, they are in the same breath conceding that it would be a problem for their view if they couldn't—that is, that whether a reliabilist theory of knowledge is true depends on whether the insight is one they can preserve. And by defending reliabilist against the Swamping Problem, reliabilist betrays an antecedent admission to the thought that the guiding insight about the value of knowledge over true belief is of *whatever* strength would be needed to make its endorsement count as a threat to whether their theory is correct.

Or, alternatively:

B. (NON-CONSTRAINING): $V(K) > V(TB)$ *really* (according to our guiding intuition) is a weaker claim—the claim that for any p , knowledge of p is most of the time more valuable than a corresponding true belief that p . This would be expressed then as an existential generalization: $\exists p [(V(Kp) > V(TBp))]$

The idea Baehr has in mind with this distinction is not hard to grasp: the thought is that an analysis of knowledge has as its aim the specification of conditions the satisfaction of which will be necessary and sufficient for knowing, and a claim about some feature knowledge has most of the time, but which is itself an *inessential* feature of knowledge, will not be one that will constrain our task of specifying just the *essential* (e.g. both necessary and sufficient) conditions of knowledge. And with this in mind, he says that the *guiding intuition* that the value of knowledge exceeds that of mere true belief only licenses us to endorse (B), the existentially quantified expression that he says would be *non-constricting* on a theory.

In the previous section, my argument for the *prima facie* reading of our guiding intuition (what I'm calling MA) revealed my willingness to suppose that our specification of the claim be unrestricted. Whether Baehr's rejection of (A) threatens this unrestricted reading depends crucially on whether it is *all-things-considered* value that stands within the scope of the universal quantifier. If the value that our guiding insight is concerned with is *all-things-considered* value, then our guiding intuition counts in favour of (B) rather than (A). But I've shown in 1.3.1 that the insight we hold clearly isn't one that an all-things-considered reading would accommodate.²⁸

It is problematic for Baehr then that he supports his thought that (A) is too strong by appealing to cases in which practical considerations make the all-things-considered value of knowing no greater than that of truly believing. He says that

...as a matter of fact knowledge is not always more valuable than true belief. This is evident, first, in certain isolated cases in which a demand for knowledge rather than mere true belief would require the forfeiture of certain other important goods. (Baehr 2009: 5)

But this fact, we've shown, is unproblematic on the *prima facie* reading of the claim. Kvanvig (2009*b*), in his more recent work on the topic expresses forcefully this point made in the previous section:

²⁸ Recall here that the guiding intuition is both one that (i) we assent to without reservation—i.e. we endorse it in the spirit of supposing that knowledge is something that *anyone* should prefer it to true belief; and (ii) it is also one that we are not inclined to revise when it is pointed out that, in some cases, the all-things-considered value of knowledge does not exceed that of true belief.

Suppose my entire family would be killed immediately if it is discovered that I know, rather than merely truly believe, that they are in Dallas today. Then, all things considered, it is not better to know. But such a conclusion shouldn't be thought to show anything about the universality or necessity of the value of knowledge. What it shows is that we shouldn't think of the value of knowledge in terms of all-things-considered value. Instead, the obvious reply is to distinguish between *prima facie* and *ultima facie* value. (Kvanvig 2009b: 14)

Because the distinction between *prima facie* and *all-things-considered* (*ultima facie*) value allows us to admit that Baehr is right to reject (A) on an *all-things-considered* reading without moving straight to (B), we can now see more clearly the more general point that citing instances where the all-things-considered value of knowledge is zero or negative isn't a cook-book recipe for generating counterexamples to the thought that the scope of the guiding intuition is unrestricted nor that the insight is supposed to be about only some instances of knowledge (and thus about some aspect of knowledge inessential to it).²⁹

1.3.3. The Third Master Question

The first two questions called for an (i) articulation of the specific claim MA should be thought of as capturing; and (ii) some reasons for thinking that the very *truth* of that claim could be upheld against various points of criticism. The next logical step will have us shift our focus from the nature of the insight itself to the nature of the relation it bears to a theory of knowledge.

Third Master Question: How might a theory of knowledge be restricted by endorsing the insight (MA) that, in the previous two sections, has been clarified and then defended. Put another way: what is it to hold a theory of knowledge *accountable* to this insight?

Accountability

What's at issue then is the curious idea that the MA is one that we should hold a theory of knowledge *accountable* for. What we need to know before worrying about anything specific to MA is how theoretical accountability

²⁹ Now it is worth noting here that Duncan Pritchard has posed a concern about the strength of the guiding insight which is itself closely related to Baehr's and which shares a common starting point. Pritchard's target is what he calls the *austere reading* of MA, which he takes to be the claim that "knowledge that *p* is *always* more valuable than mere true belief that *p*."²⁹ Taking the *primary value problem* to be the problem of *explaining* this, Pritchard reasons that "It is far from clear why we should endorse the austere reading of the primary value problem. For one thing, it seems that the primary value problem, construed austere, is unanswerable, since there are surely bound to be cases in which it would be better to merely truly believe a proposition rather than know it (e.g., where one's life depends on it, for example)." The primary value problem would indeed be unanswerable *only if* that which it must answer to is a claim that carries with it the implication that knowledge is more valuable than true belief even when knowing (and presumably not merely truly believing) will cost you your life. Leaving behind the *all-things-considered* reading for the *prima facie* reading of our intuition, the worry that the austere reading of our intuition is itself unexplainable dissolves.

might be extended to claims or insights more generally. That is, we need to know what the conditions are under which, for *any* claim, we should hold a theory accountable for it.³⁰

Now by invoking the thought that recent philosophers writing on the value of knowledge and truth are interested in whether, how to what extent the MA insight is something that we should be holding a theory to *account* for (the language used in the third master question), I'm intending a sense of accountability that is a familiar one in philosophy. I'll aim to elucidate this sense of accountability by cashing it out in terms of some equally common yet comparatively more basic notions we employ when assessing theories.

Let's take then as a starting point the thought that, as philosophers, we frequently aim to adjudicate between theories, and we do this in part by evaluating competing theories as they stand up to various conditions of adequacy, one of which is the preservation of certain platitudes.³¹ Now we might be tempted to think that a theory of Φ should only be held accountable for preserving platitudes that are, in some sense we might think is obvious, *about* Φ . But that's not quite right. After all, if a moral philosopher offered a theory of justice that carried with it the strange (and unwelcome) implication that we must be skeptics about external world knowledge, then all the worse for the theory of justice. This point about platitudes—e.g. widely-held convictions, such as that we have at least some knowledge of the external world—can be put as follows: platitudes constrain all theories, even though for any given theory, only some platitudes will be *about* that which that theory is out to explain. This observation about platitudes is important for our purposes in that it invites us to notice two things: first, a key subtlety about the way that theoretical *constraints* are connected to what they must *preserve*; and secondly, just how this subtle point about the connection between *constraints* and *preservations* encourages that we sharpen the notion of theoretical accountability that will frame the ensuing discussion we'll be undertaking about the value of knowledge and true belief.

Drawing then from the point that platitudes do not constrain *only* those theories of things the platitudes are strictly *about*—as was shown in theory of justice case—we might nonetheless resist the strong suggestion

³⁰ It should be obvious that pronouncing on *whether* or *why* the accountability relation holds between a theory of knowledge and any particular assumption about knowledge (be it MA or otherwise) amounts to a spinning of wheels in the absence of a view about what that relation is. We'll see that it would be a mistake to think as other writers have (with the exception of Kvanvig) that the sense in which we are interested in holding a theory of knowledge to account for the value assumption is one that permits us to leave by and large the accountability relation on pre-theoretic grounds.

³¹ Perhaps the most basic condition of adequacy a theory will be measured for concerns logical consistency. Taking logical consistency as a paradigmatic condition of theoretical adequacy, we can see how accountability, in the sense it shall be used here, is tied to the equally familiar and related notions of *preservation* and *constraint*. Logical consistency is, among other things, something that a theory must *preserve*. Correspondingly, *that it preserve logical consistency* is something we will naturally refer to as a *constraint* on a theory—something that limits what an adequate theory of Φ can say about Φ . The intended notion of *accountability* at issue will be a function of the comparatively more basic concepts of *preservation* and *constraint*. Suppose that P is a platitude about Φ and it is Φ that we want to give a theory of. We hold a theory of Φ *accountable* for P by placing on it the *constraint* that P be preserved.

that platitudes constrain *all* theories by drawing attention to a point that can be extracted from the very same case. The point might be: *even if* in moral philosophy, a theory of justice must preserve the platitude that we have at least some knowledge of the external world, the ensuing constraint that arises out of that platitude constrains a theory of knowledge differently.³² As a rule of thumb, it should seem that the strength of the constraint that a platitude places on a theory—and thus the sense in which we should say that that platitude is one the theory should be held accountable for—will be a function of two things: (i) the extent to which the platitude is *about* the thing we’re giving a theory of; and (ii) the strength and scope of platitude.

As for (i): consider that we don’t say that the anti-skeptical platitude (one clearly not about justice in any relevant sense) is one that a theory of justice should be held to account for, *even though* a theory of justice is constrained by the platitude in the weak sense that it must be logically consistent with it. But we do hold a theory of justice accountable to the insights we have about fairness—as these insights are ones that are about something we think is at the core of justice.

Regarding (ii): Even if some claim is relevantly *about* what it is we’re giving a theory of, we don’t hold a theory accountable for it unless it is, additionally, a claim that is not restricted in a way that would have it apply only to some instances of that which the theory is about. For example, we wouldn’t hold a theory of justice accountable to the claim that justice is a virtue of effective rulers—even though this claim is probably true of most effective rulers and clearly *about* that which a theory of justice is about. This was a point alluded to by Baehr in his thought that claim (A) but not (B) formally constrains a theory of knowledge.

Drawing from these two aspects of accountability, we get the following:

(ACCOUNTABILITY-1): We can appropriately hold a theory of Φ accountable to some true claim (or platitude) P if and only if: (i) P is about Φ ; and (ii) The scope of P is not restricted to only some instances of Φ ; it must be unrestricted in a way that holds for all Φ s.

³² Whereas a theory of justice is constrained by the external-world knowledge platitude only in the trivial sense that it must not imply the falsity of it (i.e. be logically consistent with it), a theory of justice is at the same time *not* constrained by the external-world knowledge platitude in any way that would lead us to think that a theory of justice preserves this platitude only if the theory *positively supports* it. After all, the mere fact that it’s a useless weapon against the external world skeptic is no flaw of a theory of justice. But the same is not true for theories of knowledge. The platitude that external-world knowledge is possible does not *only* constrain a theory of knowledge by requiring that a theory of knowledge be logically consistent with it (i.e. not imply the skeptical conclusion). Additionally, the anti-skeptical platitude constrains a theory in a way that requires some positive support (over and above a mere *lack* of incompatibility). A qualification: the constraint consisting in the requirement that it positively support the anti-skeptical platitude does not require that the theory *prove* the platitude to be true. Instead, we think of the anti-skeptical constraint on a theory of knowledge as occupying some middle ground: stronger than *lack of incompatibility* but weaker than *proof*. With this said, we should find it plausible that the most attractive notion of theoretical *accountability* will be appropriately sensitive to the strength of the constraint imposed by the platitude at hand.

Aside from the conditions under which we could rightly hold a theory accountable to a platitude, we also can conclude from the previous discussion what this will entail:

(ACCOUNTABILITY-2) We hold a theory of Φ accountable to some true claim (or platitude) P only if we require that the theory be both (i) logically consistent with P , and (ii) can be shown to give some positive support for P .

From these two equations, we can outline the notion of accountability at issue between MA and a theory of knowledge that wishes to preserve it. Borrowing from the first equation, we can say that MA is a platitude for which a theory of knowledge should be rightly *accountable* (and not merely constrained in the sense of being required not to contradict) because the MA is straightforwardly *about* knowledge, and further, the claim MA makes about knowledge is one whose scope is unrestricted. And, borrowing from the second equation, we can say what a theory of knowledge must do in order to account for MA. Let's call this the *Meno Requirement*:

MENO REQUIREMENT: A theory of knowledge positively *accounts* for MA only if it is both (i) logically consistent with MA and (ii) is one we could use to provide *positive support* MA.

A theory of knowledge satisfies the Meno Requirement then only if it meets *both* of these conditions. Condition (i) will be satisfied by a theory only if it is logically consistent with MA, and this means that a theory of knowledge must be one on which, from a *purely epistemic point* of view, knowledge (on the theory) has a value that exceeds that of mere true belief. An endorsement of MA alongside an inability to show how (i) of the Meno Requirement is satisfied is the mark of those theories Kvanvig argues succumb to the Swamping Problem—a problem we'll investigate in more detail in Chapter 3. Condition (ii) of the Meno Requirement is the requirement Pritchard calls the *Meno Problem* (and also the Primary Value Problem). It is fundamentally an explanatory problem, one that requires us to not simply 'have' a theory of knowledge consistent with MA, but that we also be able to explain *what makes* knowledge have always and everywhere epistemic value that exceeds that of true belief. If we can't provide a theory of knowledge that we can appeal to in explaining this, then we'll have failed to address the second part of the Meno Requirement—the part which holds that MA isn't *just* a claim (like, for example, some platitude about justice) with which a theory of knowledge should be logically consistent. The fact that MA is both about knowledge and is unrestricted in scope means that our theory must be one we can appeal to in giving some *positive* explanation for why it is so.

1.4 Concluding Remarks

I've to this point argued for a particular characterization of the *Meno Assumption* that best captures the widely held intuition that the value of knowledge exceeds that of mere true belief; I then argued that the claim, articulated as a prima facie claim about epistemic value, is strongly motivated and defensible against several recent challenges. Finally, I showed just how we should think of a commitment to the Meno Assumption as formally constraining a corresponding account of the nature of knowledge.

2

Swamping, Additive Value and Epistemic Parts

2.1 The Swamping Problem

Let us return now to Kvanvig's (2003) claim that a theory of knowledge must not only be counterexample-free, but also amenable to what we say about the value of knowledge. (Kvanvig 2003: 6) The notion of 'amenable' here is imprecise, and so he rightly qualifies it:

What do I mean by 'amenable to'? At the very least, the theory must be logically consistent with an account of the value of knowledge, but perhaps something stronger is required. (Kvanvig 2003: 6)

If what we're saying about the (epistemic) value of knowledge is that it exceeds that of mere true belief, then accordingly, a theory of knowledge will have to be amenable to that claim; at the very least, then, a theory's conditions for knowledge must be able to, when satisfied, make a true belief more valuable epistemically than it would be otherwise (e.g. if the true belief were not known).

Recent work on the value of knowledge has led epistemologists to consider whether their theories of knowledge meet this minimal requirement. When theories are said to fail it, it is usually by reference to the Swamping Problem, though this charge is (though popular) not without some ambiguity.

The Swamping Problem, defended most prominently by Kvanvig (2003) but also discussed in work by Swinburne (1999), Jones (1997), Riggs (2005), Pritchard (*forthcoming a*) and others, is sometimes discussed as though the relevant 'problem' it refers to is a philosophical problem or challenge: the 'problem' of figuring out a way to show how, for example, justification (or other epistemically valuable properties) could add value to a true belief.

Another sense in which the Swamping Problem is discussed, though, understands the problem pejoratively: a theory of knowledge falls to the Swamping Problem by failing to define knowledge in a way such that true beliefs could gain epistemic value from whatever other conditions the theory requires for knowing. In this sense, a theory that succumbs to the Swamping Problem will be one on which knowledge turns out to be

(on that theory) no more valuable than mere true belief. Such a theory then is not even logically consistent with an account of the value of knowledge insisting that knowledge is better than mere true belief. And so, as it were, you aren't entitled to uphold the value of knowledge to true belief unless your theory of knowledge avoids the Swamping Problem. In this respect, the Swamping Problem can be seen as a 'test'—and the prize for passing this test is free reign to endorse the value of knowledge over that of mere true belief, an endorsement that would be otherwise off-limits—and that's indeed a problem. Following Pritchard (*forthcoming a*) here, I think it can be useful to break the Swamping Problem into three theses, at least one of which must be given up. This characterization of the problem offers a 'big-picture' approach to seeing just what generates the Swamping Problem.

The first thesis here would be some broad endorsement of the assumption that in Chapter 1 I discussed and defended in detail: put simply, that knowledge is more valuable (epistemically) than true belief.

First Thesis: Knowledge is more valuable than mere true belief.

The second thesis generating the Swamping Problem relates to the familiar idea that truth is the fundamental epistemic aim, or put differently, the primary goal of inquiry. An important implication of this thesis is that properties of beliefs such as justification (for example) would be epistemically valuable for a belief to have insofar as justified beliefs are likely to be true. This idea falls in line with Sosa's (2007) conception of epistemic normativity, which he describes as 'a kind of normative status that a belief attains independently of pragmatic concerns.' (Sosa 2007: 88) In particular, what the second thesis here claims is:

Second Thesis: Epistemic value is instrumental value that something has as a means to the end of truth.

The third thesis is what I'll call the *Swamping Thesis*: Pritchard articulates this thesis as follows:

Swamping Thesis: If the value of a property possessed by an item is only instrumental value relative to a further good and that good is already present in that item, then it can confer no additional value. (Pritchard *forthcoming a*: 5)

It might not be at all obvious how the Swamping Thesis is supposed to connect to the other two, or for that matter, how it would come to bear on an account of the value of knowledge. Admittedly, the idea it aims to capture isn't immediately clear, nor is it easy to tell, in the context of the other two theses, whether it's even plausible. Because the Swamping Thesis is meant to be a general thesis about value—we can examine its

plausibility by looking at a concrete case where the value at issue is more familiar. Kvanvig (2003) provides such a case:

Suppose I want chocolate. I google to find places close to me. I get two webpages: one titled places that sell chocolate in Waco; the other “places likely to sell chocolate in Waco...” We then note four important facts about these lists. The first fact is that the second list is better than nothing. If I want chocolate, rather than choosing randomly from all the businesses in Waco, it would be better to select randomly from the list of places likely to sell chocolate in Waco. The second fact is that, for the purposes in question, the first list is more valuable than the second. I do not wish to say that it would be irrational to rely on the second list, given only the purpose of getting chocolate, but it would certainly be better to rely on the first list. The third and fourth facts arise from introducing a third list, which is the intersection of the first two lists: places that both sell chocolate and are likely to. The third fact is that the intersection list is also preferable to the second list: it would be better to rely on a list that included places that both sell chocolate and are likely to than to rely on a list that included only places likely to sell chocolate. The final fact, however, is the swamping fact: the intersection list is no more valuable than the first list, given the interest in question. If all I care about is chocolate, it would be no better to use the list of places that both sell chocolate and are likely to than to use the list of places that sell chocolate. (Kvanvig 2003: 47-48).

The Swamping Thesis is supposed to be validated here by the fact that the list of places likely to get you chocolate, though valuable on its own as a means to getting you chocolate, confers no additional value to a list of places *guaranteed* to get you chocolate. Toward the aim of getting chocolate, then, the value of a list likely to get you chocolate is ‘swamped’ as it were by the value of the list that guarantees you’ll get chocolate.

The Swamping Thesis, Kvanvig thinks, can be applied to the reliabilist theory of knowledge in a way that is analogous to the case just considered, where it’s chocolate that’s sought. Replace the goal of getting chocolate with the goal of getting truth. If you wanted chocolate, the list of places likely to get you chocolate was valuable insofar as it makes getting chocolate likely. But as Sosa (2007) puts the problem:

“How can the truth-reliability of an epistemic source give to the beliefs that it yields any additional epistemic worth, over and above any that they already have in virtue of being true? (Sosa 2007: 72)

If it’s truth you’re after, Kvanvig says, the property of ‘being reliably produced’ is a valuable property for a belief to have insofar as it makes the belief likely to be true. But, as the argument goes, a belief that’s true is just as valuable as a belief that’s both true and likely to be true. So the value of reliability is ‘swamped’ by the value of truth.³³ Reliabilism, thus, is said to succumb to the Swamping Problem: The reliabilist, at least insofar as she

³³ Consider now a more explicit parity of reasoning: If you want chocolate and have a list of places guaranteed to sell it, you can simply throw away the list you have of places likely but not guaranteed to sell it. After all, a list that guarantees you chocolate wouldn't gain any value by being in addition to guarantor of chocolate sellers, a high probabilizer of chocolate sellers. Similarly, if you want the truth and have a belief that's true, your belief can't gain epistemic value by being in addition to true, also 'likely to be true.'

accepts the second and third theses (e.g. veritism and the Swamping Thesis) is forced to concede that reliably produced true belief (knowledge) is no more valuable epistemically than mere true belief. But, as we saw in Chapter 1, that would be a really bad thing to have to give up.

Pritchard's observation that at least one of the three theses constituting the Swamping Problem must be rejected is interesting and I think apt as a way to set straight the rules of the game. The critique of reliabilism just considered showed the reliabilist to be unable to maintain the first claim—the claim that knowledge is more valuable than true belief—because this claim, in conjunction with the other two theses, seemed to imply that knowledge is *no more valuable* than true belief.

It's worth noting here that, according to Kvanvig, the Swamping Problem shows that not only does the reliabilist fail to maintain the first thesis, but further, that no theory of knowledge accepting the second two theses can preserve it; consequently, the Swamping Problem is one that he thinks shows that knowledge, no matter *what* your theory, must be admitted to lack the distinctive value it has traditionally been taken to have.

Now the common way to resist Kvanvig's conclusion has been to try to show that, on one's preferred theory, true beliefs actually *would* become more epistemically valuable than otherwise in virtue of having the properties that, on that theory, a belief must have to count as knowledge. Goldman & Olsson (2007) and Greco (2009) provide good examples of this sort of attempt. But such a strategy, which is one that purports to avoid the Swamping Problem without giving up anything, is at best meretricious; as Pritchard points out, a theory of knowledge cannot engage with the Swamping Problem without giving up at least one of the three theses. While Riggs has tried rejecting the second thesis in the context of discussing the role of truth in the Swamping Problem, no one has, to my knowledge, had a proper go at the third thesis. And it's precisely this thesis, the Swamping Thesis, that I think we should resist.

2.2 Challenging the Swamping Thesis

2.2.1 Invariabilism and Additivism

Recall again the Swamping Thesis:

Swamping Thesis: If the value of a property possessed by an item is only instrumental value relative to a further good and that good is already present in that item, then it can confer no additional value. (Pritchard *forthcoming* a: 5)

The Swamping Thesis seems to get the right result in a variety of cases where we think of the value something as *instrumental* value as a means to some other valuable end. For example, if your goal is to get just a brief idea what Tchaikovsky’s music sounds like, then any Tchaikovsky record would be instrumentally valuable as a means to this end—for any Tchaikovsky record would be good enough to provide you with a brief idea what he sounds like. But suppose you buy one from a store and go home only to find your mother blaring a song that she announces to be Tchaikovsky. Here again, it looks like the Swamping Thesis prevails: given that you just wanted a brief idea of the musician’s sound, the music already playing had given you this, and though the CD you had was valuable instrumentally toward this end, it no longer is.³⁴

As a point of clarification, it should be noted that even though swamped value is indeed ‘lost’, swamped value isn’t *simply* value that something loses, as is the case when a thing becomes damaged or goes out of fashion. The Swamping Thesis is instead one that, we see, applies to things that are valuable in their own right, but which then lose this value in the presence of other valuable things. The Swamping Thesis therefore presupposes some account of value on which, for two things Φ and Ψ , where Φ is instrumentally valuable as a means to Ψ , Φ can possess value apart from Ψ but not otherwise.

Not all metaethical accounts of value are friendly to this essential presupposition of the Swamping Thesis. Consider here *invariabilism*. Let V represent the value function. Let V_{xy} be the value of x as a part of y , and V_x the value of x as a part of x (i.e. x as a part of itself, or as a whole). Invariabilism says:

INVARIABILISM: If x is a part of y , then $V_y(x) = V(x)$

Invariabilism has been championed most notably by Moore (1903), who endorsed this idea alongside his Principle of Organic Unities. The *invariabilist* claims that the value of x as a part of x is equivalent to the value of x as a part of y . Now this is clearly incompatible with what is implied by the Swamping Thesis; the Swamping Thesis says, for example, that when together with the list of places guaranteed to sell chocolate, the list of places likely to sell chocolate contributes no valuable, but when evaluated not as a part of a whole that includes the ‘guaranteeing’ list—but *apart* from such guarantees—the list of places likely to sell chocolate would be a valuable means toward your goal of getting chocolate. So invariabilism has to go; the metaethical account of value the Swamping Thesis must take on board instead of *invariabilism* is *additivism*, which Brown (2006: 456-

³⁴ Of course, the CD would still have some value if you sold it back, but that is beside the point; it does not change the fact that the value it had as a means to introducing you to Tchaikovsky is lost in the context of your actually being introduced to Tchaikovsky.

463) articulates as:

$$\text{ADDITIVISM: } V(x+y) = V_{x+y}(x) + V_{x+y}(y)$$

Accordingly, *additivism* holds that the value of some whole consisting in parts x and y will be a value we get *not* by adding together the value of x as a part of x and the value of y as a part of y , but by adding the value of x *as a part of* $(x+y)$ and y *as a part of* $(x+y)$. Brown (2006: 456-7) offers a nice example here: if we suppose that chips are better when gravy is added, then it would seem that gravy adds value to chips. Ice cream, however, does not become more valuable when gravy is added. The additivist explains this by pointing out that the value of gravy as a part of $\text{gravy} + \text{chips}$ can be one thing, while the value of gravy as a part of $\text{gravy} + \text{ice cream}$ another. The value of gravy thus varies as a function of the whole we're evaluating, and of which gravy is a part.

So *additivism* says that the value of Kvanvig's 'probabilizing' list, when *apart* from the 'guaranteeing' list, can bear some value V as a means to getting chocolate (which is just what Kvanvig claims), even though this same list, as a part of a whole made up of itself and the guaranteeing list, need not bear value V .

Let's revisit now the epistemic case. To make things simple, let 1 represent a positive value and 0 represent no value. Additivism allows that, if your goal is truth, then 'being reliably produced' can have value 1 so long as it is not a part of a whole of which truth itself is also a part. On closer inspection, this idea will prove to be quite confusing. This is not necessarily because additivism is mistaken; rather, additivism simply reveals a curious point about the Swamping Thesis—a point that shows the epistemic case to come apart from the related examples through which the Swamping Thesis was illustrated.

To see where these cases come apart from the epistemic case, consider first that, in the chocolate case, it is the probabilizing list whose value gets swamped; in the music case, it was the CD whose value gets swamped, and in the epistemic case, it is supposed to be the property of being reliably produced whose value gets swamped.

That said, consider a bit further: if you want chocolate, it's easy to imagine one list being valuable on its own, apart from the other. And if you want to be introduced to Tchaikovsky, it's also no trouble at all to imagine the CD being valuable on its own, for example, if you'd taken it to your friend's house to play it rather than to bring it into a context where the music on that CD was already playing.

But if you want the truth, then by parity of reasoning, it should be no trouble to imagine reliability as a valuable property for a belief to have even when reliability is *not* a part of a whole of which truth itself is also a

part.

It would indeed be a serious problem if the Swamping Thesis implied that reliability lacks epistemic value entirely—that there is no context in which reliability is an epistemically valuable property of a belief. Alvin Goldman would be one of more than just a few to object to that. And although proponents of the Swamping Thesis certainly do not intend or admit such a consequence, that’s precisely what the Swamping Thesis implies.

2.2.2 The Value of Reliability

With this said, let’s again imagine reliability as a valuable property for a belief to have when reliability is *not* part of a whole of which truth itself is also a part. Here’s a first pass: reliable beliefs can be either true or false. Reliability is not a part of a whole of which truth is a part in instances where the reliably produced belief at issue is false. Does reliability bear epistemic value here, as part of a reliably produced false belief? Consider the following case:

DRAWING FOR A CAR: Suppose you have two tickets in a fair lottery for which a thousand tickets are sold. The tickets are numbered 1 to 1000; one number will be drawn blindly, and so one ticket will win a prize, in this case, a new car. You have tickets 1 and 2, and you hope that one of them will win you that shiny new Chevrolet. Before the numbers are drawn, tickets 1 and 2 are each instrumentally valuable insofar as you want to win the car, and each carries with it some likelihood that you will. Suppose now that after just the first of the two numbers is drawn, your ticket 1 wins—lucky you! While ticket 2 was valuable before ticket 1 won, it loses its value once it is determined that you’ve already won the car. After all, once ticket 1 wins, your other ticket no longer carries with it a likelihood of getting you a car, and it was in virtue of this likelihood that we said the ticket was instrumentally valuable. Now, in this case then, we should say that ticket 2 was indeed instrumentally valuable before the outcome of the lottery was determined; after the outcome was determined, and ticket 1 won, ticket 2 lost its instrumental value; after ticket 1 wins, you can throw ticket 2 away (whereas you’d be foolish to do so before the outcome was determined.) Now, consider a further point: suppose that the outcome was different, ticket 526 wins. You can now throw *both* tickets 1 and 2 away; 2 loses the instrumental value it had in this scenario not because it is already guaranteed that you win, but instead because you’re guaranteed *not* to win.

Now, because reliability is an epistemically valuable property for a belief to have because it makes truth probable, then by reference to the Swamping Thesis, this value is lost when reliability is evaluated as a property of a *true* belief. Here, the end toward which it is valuable as a means is already *guaranteed*—the belief lottery, as it were, is already determined in virtue of the fact that the reliably formed belief you have is true. But as the DRAWING FOR A CAR case shows, the instrumental value of Φ as a means to Ψ will *also be lost* if $\sim\Psi$ is guaranteed. And so the instrumental value reliability has as a means to truth would disappear in the presence of

falsity just as well as truth. After all, when your belief is false, reliability fails to make that belief ‘likely to be true’ just as, when the outcome of the lottery is already determined, and you’re guaranteed not to win the car, ticket 2 no longer makes it likely at all that you would.

This point generalizes: the instrumental value something has as a means to a further good fails to add value to something that either already (i) *has* or (ii) *can’t have* that good. The Swamping thesis is therefore a thesis that we can plausibly accept only if we are also willing to grant what I’ll call Lost-Goal Swamping:

Lost-Goal Swamping: If the value of a property Φ possessed by an object O is only instrumental value relative to a further good Ψ and O has the property $\sim\Psi$, then Φ can confer no additional value to O .

Now, if reliability has instrumental epistemic value as a means to truth, then it must have instrumental value (like our lottery ticket 2) in a context in which the outcome of whether the belief is true is not decided. After all, once that matter is decided, then if the belief is true, then via the Swamping Thesis, reliability loses its value, and if the belief is false, then via Lost-Goal Swamping, reliability loses its value. Problematically, though, it is here where beliefs and lottery tickets, which can be instrumentally valuable, come apart: it is only once a ticket is drawn that the outcome of a lottery is determined, and before that, the tickets were instrumentally valuable insofar as they made the favorable outcome probable. Reliably formed beliefs, on the other hand, are always going to be either true or false (as opposed to undecided). In either case, reliability fails to be instrumentally valuable as a means to truth. More explicitly, the argument here is as follows: (let R represent the property of being reliably produced).

A Reductio for the Swamping Thesis

1. Epistemic value is instrumental value that something has as a means to the end of truth. (Second Thesis)
2. As a means to the end of truth, R is an epistemically valuable property of some belief p insofar as R increases the likelihood that p is true.
3. R beliefs are either true or false.
4. R does not increase the likelihood that true beliefs are true.
5. R does not increase the likelihood that false beliefs are true.
6. Therefore, R is not an epistemically valuable property of true beliefs. (From the Swamping Thesis)
7. R is not an epistemically valuable property of false beliefs. (From Lost-Value Swamping)

8. Therefore, *R* is not epistemically valuable. (From 3, 5 and

A defender of the Swamping Thesis will not try to get out of this by rejecting (1) because (1) expresses the second of our three theses that were needed to get the Swamping Problem up and running. Rejecting (2) is also off limits, given (1): if *R* is an epistemically valuable property of a belief as a means to the end of truth, then given that *R* doesn't *make* a belief true (after all, some reliably formed beliefs are false), we have to understand *R*'s connection to truth more weakly: as an indicator of truth; and *R* counts as an indicator of truth only insofar as the probability of a belief's being true given *R* is greater than it would be otherwise. (Otherwise, *R* would fail to even indicate truth). And so for one who endorses the Swamping Problem, (1) is off limits, and therefore, (2) will be off limits as well. Rather than to try to take issue with (3), (4) or (5), I suspect that one might try to save the Swamping Thesis against the unacceptable conclusion I've offered by challenging implicit assumptions essential to the argument.

Because I think the Swamping Thesis is a doomed one, I want to take seriously these possible outs and show why they don't work.

2.3 Avenues of Defense

2.3.1 The Value of Reliability as Expected Value

The first escape option might be to distinguish between expected value and actual value. A reliably held belief, it might be argued, could have some expected epistemic value in the space between the agent's forming the belief and finding out whether it is true or false. This line fails for two reasons: firstly, it is consistent with the claim that reliability lacks any epistemic value, which is a claim about instrumental value something actually has as a means to truth. And so considerations about whether reliability has expected value are orthogonal to the argument's conclusion. Moreover, though, it's hard to understand how to even think about *R* in terms of having expected epistemic value; the conditions required for *R* to hold are not ones that believers must be *aware* of, and expected value is value we expect something to have given that we are aware of certain properties it has that are likely to make it valuable. Expected value, then, is not something to which a defender of the Swamping Thesis should be appealing as a way to dodge our conclusion.

2.3.2 Additivism and Epistemic Parts

I think the more promising avenue would be to try to draw out some of the possible implications of the additivist account of value that is being presupposed by the Swamping Thesis. In particular, the additivist would allow that R as a part of T, B (or F, B) could have no value whilst nonetheless bearing some positive value as a part of itself (that is, as a part of R).

This line gets interesting and also somewhat complicated. For how could the property of ‘being reliably produced’ be something on its own? Are there conditions under which the reliability of some belief isn’t a part of the reliably produced belief itself? This might seem like a strange question to have to ask here, but we must ask it in order to make sense of how a defender of the Swamping Thesis could uphold that reliability could *ever* be an epistemically valuable property of a belief. And so we are led from a specific question about the value of reliability to a more general question of ontology, and in particular, mereology. Peter van Inwagen calls this the Special Composition Question (Van Inwagen 1990: 21-32): What are the conditions under which two objects form a third object?³⁵ Some answer to this question must be presupposed by any answer to the more specific matter of whether a belief and its property of being reliably produced must be understood together as a whole.

In classical mereology, there is a standard answer to the Special Composition Question—one that is endorsed by (among others) David Lewis (1986) and Ted Sider (1997, 2001); this answer is called *unrestricted composition*, an axiom that states:

UNRESTRICTED COMPOSITION: For any x and any y , there is a mereological sum $(x+y)$ ³⁶

Unrestricted composition is simple, elegant and widely-held reply to Van Inwagen’s Special Composition Question. The idea is that, for any two things, there is a mereological sum, and that sum is nothing over and above those two things. The sum just is those two things. If unrestricted composition is correct, then for some reliably produced true belief, there is a sum that consists in R , B and T ; for some reliably produced false belief, there is a sum that consists in R , B and F .

Moreover, if unrestricted composition is correct, then even if R could be ‘something on its own’, apart from the belief that was itself reliably produced, there is nonetheless a sum R, T, B or R, F, B (depending on

³⁵ This is the characterization of Ingawen’s question given by Daniel Korman, “Unrestricted Composition and the Argument from Vagueness.”

³⁶ Among those who defend the position are Lesniewski, “On the Foundations of Mathematics;” Leonard and Goodman, “The Calculus of Individuals and Its Uses;” Goodman, *The Structure of Appearance*; Lewis, *On the Plurality of Worlds*; Van Cleve, “Mereological Essentialism, Mereological Conjunctivism, and Identity Through Time;” Hudson, *A Materialist Metaphysics of the Human Person*; and Sider, *Four-Dimensionalism*.

whether the reliably produced belief was true or false), and the instrumental epistemic value of R is lost as a part of such a sum. What the defender of the Swamping Thesis would need to explain, more specifically then, is how R can have epistemic value apart from T,B or F,B *even though* R will always be a part of a sum that includes either T,B or F,B , and R lacks epistemic value as part of such a sum.

We have now a more refined mereological question: Is R something that can be located both within a sum that includes T,B or F,B and also *apart from* such a sum, where we may assess its epistemic value (via additivism) as a part only of itself? Here the defender of the Swamping Thesis might wish to just assert ‘yes’ and invite that the matter be reduced to brow beating. But there are further reasons to think that this rather abstract point of contention isn’t simply hostage to a shouting match. Consider here a standard axiom of classical mereology known as *inheritance of location*:

INHERITANCE OF LOCATION: For any x , x is located where its parts are.

Even on the most liberal, atemporal, aspatial view of locations, it's hard to see how we might 'locate' R apart from the reliably formed belief of which it is a property, given what inheritance of location tells us: that beliefs are located where their parts are. For any reliably formed belief, part of what makes it a reliably formed belief is that it has the property of being reliably formed. Apart from that property it isn't, after all, a reliably formed belief. And so, by *inheritance of location*, wherever R is located, so is the belief of which R is a part. And given that beliefs that are reliably formed beliefs will be either true or false, then by reference to the Swamping Thesis and Lost-Goal Swamping Thesis, it follows that R is not evaluable apart from properties the presence of which make the instrumental value of R otiose.

2.4 Diagnosing the Appeal To Swamping

We have then the following argument:

1. If the Swamping Thesis is correct, reliability is not (ever) an epistemically valuable property for a belief to have.
2. Reliability *is* (at least sometimes) an epistemically valuable property for a belief to have.
3. Therefore, the Swamping Thesis is not correct.

The aim of this chapter has not been to defend (2)—a premise that a defender of the Swamping Thesis also wants to preserve—but to defend (1). What I've shown is that the Swamping Thesis implies that reliability cannot be an epistemically valuable property for beliefs to have at all; I've shown that the commitments of an

endorser of the Swamping Thesis make it impossible, under closer inspection, to salvage the epistemic value of reliability even if desperate attempts are made to do so—attempts that rely on the thought that reliability might have value in its own right, apart from true or false reliably produced beliefs.

If I'm right, then an endorser of the Swamping Thesis must either bite the bullet and reject (2) here, or alternatively, abjure the Swamping Thesis. I should say a bit about why denying (2) would be too great a cost for the sake of preserving the Swamping Thesis. First, consider that, as Kvanvig supposes, the Swamping Problem is supposed to be one that prevents any theory of knowledge—not just reliabilism, which is frequently used as the target—from preserving the distinctive value of knowledge over true belief.

For one who endorses the Swamping Thesis, it should be clear that not just reliability, but *any* justification condition, will fail to be one that could be explained as an epistemically valuable property of a belief. And so once the Swamping Thesis is embraced, along with the second thesis that generates the Swamping Problem (i.e. the thesis about epistemic value as instrumental value as a means to truth), the argument is up and running which would conclude that, for any justification condition, it will lack any epistemic value at all! And so by denying (2) in the above argument, the proponent of the Swamping Thesis will also have to be prepared to deny that justification (full stop) is ever an epistemically valuable property for a belief to have. This is, of course, an absurd position to hold. After all, an account of epistemic value on which lucky guesses are just as good, epistemically, as beliefs that are formed responsibly is conceiving of epistemic value in a way that is quite far removed from ordinary thinking about the values that govern epistemic practice, and how inquiry should be conducted in light of these values. We shouldn't simply *not care*, from an epistemic point of view, whether our beliefs are reliably formed, or for that matter, justified. But indeed, an overlooked though inevitable implication of the Swamping Thesis is just this.

And so my conclusion here is that the Swamping Thesis must be rejected. Before considering the implications of rejecting it, I should aim to give at least some diagnosis for why it is that the thesis has been so seductive. After all, the Swamping Thesis really does seem to get the right result in Kvanvig's chocolate case and also in cases such as the Tchaikovsky case. So the Swamping Thesis must have something going for it, right?

Indeed, I think it does, and the the diagnosis I want to give here is one that will reveal just why the cases where the Swamping Thesis seems to get the right result are different from the epistemic case. To see this difference, consider that in the chocolate and Tchaikovsky cases, the relevant item of instrumental value is thought of as valuable as a means to a certain end which is *itself* not an end that is either already guaranteed or

determined to be unattainable. For this reason, we can make sense of the idea that, for example, the Tchaikovsky CD is instrumentally valuable as a means to your goal of being introduced to Tchaikovsky's sound. A precondition for this CD being instrumentally valuable as a means to this goal is that the goal isn't already satisfied or guaranteed not to be. So part of what is required for the CD to be instrumentally valuable as a means to the goal of hearing Tchaikovsky is that you've not already heard Tchaikovsky, and yet it's *possible that you could*. If, indeed, a demon had ensured it that Tchaikovsky music is 'unhearable', then your CD will have lacked any instrumental value toward the end of hearing what Tchaikovsky sounds like. And so it is in part *because* the relevant end is not yet attained and could *possibly* be, that we can account for why the CD has instrumental epistemic value that could be lost in the presence of hearing your mother play Tchaikovsky. Similarly, the list of places likely to get you chocolate can be explained to have instrumental value as a means to the goal of getting chocolate only because the list is one that we could have in the absence of having already met the goal of getting chocolate or it being such that it's impossible for us ever to attain it.

The Swamping Thesis thus looks plausible when considering cases of instrumental value in which the end (toward which something is instrumentally valuable as a means) is not yet met nor guaranteed *not* to be met. In these cases, an implication is *not* going to be that (for example) the list of likely chocolate sellers or the Tchaikovsky CD could never have any instrumental value as means to the relevant respective ends. And so we can see how the Swamping Thesis can seem quite alluring when applied to such cases, and also why it has naturally been extended to the epistemic case—given that the epistemic case also appears to be one where some property can be seen as bearing instrumental value as a means to some goal in whose presence the apparent instrumental value appears to be lost. For these reasons, we can see why the Swamping Thesis would seem both plausible and one that could be extended to discussions of epistemic value, which are ones for which—I've shown—an application of the Swamping Thesis does not have the same implications with respect to the possibility of instrumental value.

It is appropriate now to ask what wider implications rejecting the Swamping Thesis will have, particularly in the context of determining whether a given theory is compatible with an endorsement of the claim that the epistemic value of knowledge exceeds that of mere true belief. A direct implication will of course be that we must refrain from simply *testing* a theory of knowledge on this score by asking whether or not it succumbs to the Swamping Problem. And so a consequence of rejecting the Swamping Thesis will be that we must find a different framework within which a given theory of knowledge is said to succeed or fail on this score.

In the next chapter, I'll try to do just this—offer a way of determining whether a theory meets the demands of the Meno Requirement, but in a way that doesn't appeal to the Swamping Thesis en route to a verdict.

3

Subjectivism, Objectivism and the Epistemic Value of Types and Tokens

3.1 Introduction

The aim here will be to find a way to test a theory of knowledge on the basis of whether it satisfies the Meno Requirement. In the previous chapter, we saw that whether or not a theory succumbs to the Swamping Problem is not itself a legitimate test on this score. Leaving swamping behind, it will be required that we identify and defend a different way of testing whether a theory provides what the Meno Requirement calls for.

Let's consider again now the Meno Requirement, which I defended in Chapter 1.

MENO REQUIREMENT: A theory of knowledge positively accounts for the Meno Assumption (MA) only if it is both (i) logically consistent with MA and (ii) is one we could use to provide positive support for MA.

Let's set aside condition (ii) for the moment. We said that condition (i) is going to be satisfied by a theory only if that theory is logically consistent with MA, and this means that a theory of knowledge must be one on which it is preserved that, from a *purely epistemic point* of view, knowledge (on the theory) has a value that exceeds that of mere true belief—(which was the claim expressed by MA).

Suppose that on theory *T*, knowledge is defined as true belief plus conditions C1 and C2. Our question becomes: What must theory *T* do in order to preserve that, from a purely epistemic point of view, knowledge (i.e. true belief plus C1 and C2) has a value that exceeds that of mere true belief?

Suppose a defender of *T* offers us a convincing explanation for why more epistemic value would be generated by Steve's truly believing that the game is at 5:00 when Steve's belief satisfies conditions C1 and C2 than otherwise. Will this be enough to vindicate *T* as a theory that preserves that, from a purely epistemic point of view, knowledge (i.e. true belief plus C1 and C2) has a value that exceeds that of mere true belief?

You'd think this might just be enough; after all, this is a common strategy³⁷. Generalizing from the Steve example, this strategy takes on the following form: theories of knowledge are defended as preserving the assumption that the value of knowledge exceeds that of mere true belief *via* arguments for how particular *tokens* of knowledge are, on a preferred theory, more valuable than corresponding true belief *tokens*—and *this* is itself a claim defended by appeals to particular examples of corresponding knowledge and true belief tokens, the former which are argued to generate more epistemic value than the latter.

Now, I am going to argue that this strategy is not at all a good one.³⁸

Consider here a brief thought experiment:

RACISM: Leroy Bigot was raised in Alabama in the 1940s and has always believed blacks to be less intelligent than whites. When asked to prove this, Leroy provides statistics that show that most blacks are less formally educated than whites and also less proportionately represented in academics. He then points to examples of 15 black people he knows and argues that none is as intelligent as most white people he knows.

Now, suppose that Leroy tries to use the reasons given here to defend the claim that people of a black type are less intelligent than people of a white type.³⁹ Such an explanation would require that Leroy identify

³⁷ This is a general strategy embraced by van Woudenberg (2006), Riggs (2004), Goldman (2009), Olsson (2007) and Williamson (2000), among others.

³⁸ I want to make explicit some of the rather quick moves that are being made. The strategy I've described relies on the thought that if a theory can explain why one token of knowledge (on the theory) generates more epistemic value than would be generated by a corresponding true belief token, then (presumably) this counts in favor of thinking that the theory will also be able to explain why most (if not all) tokens of knowledge (on the theory) generate more value than corresponding true belief tokens. If we were to press the point, a defender of (for example) *T* might simply give us another example: perhaps an equally compelling explanation for why (according to theory *T*) David's true belief about where he left his keys will generate more epistemic value if C1 and C2 are satisfied than his true belief would have otherwise. Presumably, then, it would only be fair for us to concede after enough examples are given that theory *T* could preserve that most *all* tokens of the type "knowledge" are more epistemically valuable than *all* tokens of the type "true belief." And from there, we'd then be pressed to concede that theory *T* has preserved that knowledge, as a *type* of epistemic standing is epistemically more valuable than mere true belief, as a *type* of epistemic standing—which is precisely what we should suspect the Meno Requirement requires insofar as it requires that a theory of knowledge preserve that, from a purely epistemic point of view, *knowledge* (full stop) is more valuable than *mere true belief* (full stop). As I've argued in Chapter 1, our unreserved endorsement of a preference of knowledge to true belief supports just this point—we take it that knowledge is something *anyone* should prefer to true belief; without qualifications that make reference to particular knowers and particular propositions, the claim here is clearly a claim that compares the value of knowledge and true belief qua *types* of epistemic standings.

³⁹ About the best he could hope to do would be to defend that most black people are less intelligent than most white people (suppose, contrary to evidence pointing otherwise, that this were true). Even if this much were true, though, Leroy would lack any explanation for why blacks are less intelligent than whites that could explain why blacks, as a type, are less intelligent than whites, as a type.

some *essential* characteristic of blacks, as a type, and defend that this essential characteristic plays a role in explaining why blacks are less intelligent. Because of course there is no connection between the only essential characteristic of blacks (blackness) and intelligence, Leroy's attempt to defend this claim by appealing to contingent properties of black people fails.

An attempt to show how one's theory of knowledge preserves that knowledge as a type of epistemic standing is more epistemically valuable than mere true belief as a type of epistemic standing had better not make the mistake Leroy makes and assume that this is something we could account for by appealing to what's not essential to knowledge as a type.⁴⁰ Because knowledge, as a *type* of epistemic standing is abstracted from any particular propositional content, (just as blackness, as a mark of a *type* of people, is abstracted from contingent facts about particular black people and their intelligence), we'll avoid Leroy's mistake only if we show that knowledge *essentially* has some valuable property that, abstracted from any particular propositional content, would make knowing more epistemically valuable than mere true belief as a *type* of epistemic standing. What this means then is that, in order to satisfy the Meno Requirement, a theory of knowledge must be one on which it is preserved that, from a purely epistemic point of view, knowledge *qua type* of epistemic standing (on the theory) has a value that exceeds that of mere true belief *qua type* of epistemic standing. Call this the Type Criterion:

Type Criterion (Meno Requirement): The Meno Requirement holds a theory of knowledge to preserve that, from a *purely epistemic point* of view, knowledge *qua type* of epistemic standing (on the theory) has a purely epistemic value that exceeds that of mere true belief *qua type* of epistemic standing.

The Type Criterion is one that a theory of knowledge should aim to show it can meet by pointing to purely formal characteristics of the analysis of knowledge that the theory can explain will make *any belief* with those characteristics more epistemically valuable than a true belief lacking them. Simply defending knowledge tokens as generating more epistemic value than corresponding true belief tokens would not get us anywhere toward showing our theory meets the Type Criterion.

That said, it might be pointed out that *surely* a theory fails to satisfy the Meno Requirement if, on the theory, *S*'s true belief that *p* would generate just as much or more epistemic value than would *S*'s knowing that *p*. That much sounds right. And so part of the Meno Requirement must concern what a theory should preserve

⁴⁰ Now, what this means is that we couldn't defend the claim that knowledge, as a type of epistemic standing, is more valuable than mere true belief as a type of epistemic standing by showing that particular tokens of knowledge are more epistemically valuable (according to our theory of knowledge) than corresponding true belief tokens. After all, the epistemic value generated by, for example, Steve's knowing *p* rather than merely truly believing *p* is value that we can only show knowledge to have by appealing to the value generated by Steve believing that proposition as he did.

about *tokens*, even though the guiding assumption itself was shown to be about knowledge and true belief as *types*. So this means there must be a Token Criterion along with the Type Criterion. In order to understand the Token Criterion, however, and its relationship to the Type Criterion, it's needed that we first understand the special and almost categorically overlooked way that *constitutive grounds* of the value of knowledge and mere true belief qua types is different from the *constitutive grounds* of the value of knowledge and mere true belief qua tokens. But to get clear about *this*, it is especially important that we consider different ways that value might be constituted. Broadly speaking, there are two central metaethical views here: *subjectivists* are guided by the general idea that the value something has is constituted by the fact that a subject *values* it. *Objectivists*, on the other hand, suppose that the source of the value valuable objects have lies entirely within these valuable objects themselves.⁴¹

I contend here that the key difference between knowledge as it bears value qua type as opposed to qua token will be understandable only if we first take great care to make precise what it is that subjectivists and objectivists are actually disagreeing about. The terrain here is at times thorny, but a proper understanding of type and token value isn't possible unless we engage more than just superficially with this axiological dispute.

3.2 Subjectivism and Objectivism

A way that G.E. Moore (1922) used to show how *subjectivism* and *objectivism* come apart was to invoke intrinsic value, and argue that its recognition is possible within an objectivist account but not a subjectivist account, even though the latter speaks 'as though' some things are intrinsically valuable. Moore says:

It is this conviction -- the conviction that goodness and beauty are *intrinsic* kinds of value, which is, I think, the strongest ground of objection to any subjective view. And indeed, when they speak of the "objectivity" of these conceptions, what they have in mind is, I believe, always a conception which has no proper right to be called "objectivity"... The truth is, I believe that... from the proposition that a particular kind of value is "intrinsic" it does follow that it must be "objective" ... (Moore 1922: 2)

The *intrinsic-extrinsic* distinction is not only thought to be connected as Moore suggests to the *subjectivism-objectivism* divide, but also to the debate about *instrumental* and *final* value.

⁴¹ This is the basic characterization of the distinction given by Toni Roni-Rasmussen and Wlodek Rabinowicz, although refinements will be in order (and made clear in the present chapter). These are quite rough definitions of subjectivism and objectivism and are somewhat contentious for reasons we'll see shortly. We'll see also that each position carries with it important implications regarding what *else* can be said about value, and these implications bear directly upon how it is that the value of knowledge qua type and token should be understood and accounted for. The subjectivist/objectivist distinction is rarely appealed to in literature on the value of knowledge (indeed, it is not discussed at all that I am aware of); amongst epistemologists, the *intrinsic/extrinsic* and *instrumental/final* value dichotomies are much more common.

Now it must be clarified at the outset that subjectivism and objectivism are not substantive views, and so do not carry with them any commitments to particular evaluative stances. It could therefore be possible, in principle, for a subjectivist and an objectivist to endorse exactly the same evaluative judgments. (Ronnow-Rasmussen 2007: 248). What they would disagree about would be how to analyze these judgments.

That said, Toni Ronnow-Rasmussen (2007) points out that it is a frequently-made mistake to think that the divide between subjectivism and objectivism is something like this:

- (1) [FALSE] Subjectivists claim that value is located in subjective states, whereas objectivists claim that value is located in valuable objects.

The disagreement in (1) would be a disagreement about what it is that value supervenes on. For some valuable object *O*, a disagreement about what *O*'s value supervenes on would consist in competing positions with respect to what properties of *O* serve as the *supervenience base* for *O*'s value—that is, a disagreement about what properties of *O* are those *in virtue of which* *O* is valuable. As Jonas Olson (*forthcoming*) says, *hedonism* would constitute such a position—a position that states what it is *in virtue of which* something is valuable—by (hedonism's) claiming that:

...the property of containing *n* amounts of pleasure makes *w* valuable to degree *n*. (Olson 2006: 132).

But a claim like this would not be evaluatively neutral, but straightforwardly evaluative. Olson writes:

If asked how to argue for this claim [that claims about what properties of an object value supervenes on are evaluative], I would simply urge that to claim that a property is value-making is to make an evaluative claim (indeed, to specify value-making properties can be seen as the very heart of the matter in substantive value theory—a most certainly evaluative business)... (Olson: 2006: 132).

Ronnow-Rasmussen concurs here (2007: 248).⁴² Accordingly, we see that (1) does not capture what it is about which subjectivists and objectivists disagree—a point frequently confused.⁴³ In order to take 'the metaethical character of objectivism and subjectivism seriously...⁴⁴ Roni-Rasmussen cautions that we must understand subjectivism and objectivism as 'neutral as to what properties of a valuable object are value-making.'

⁴² Ronnow-Rasmussen writes that Judgments such as 'Final value accrues to pleasure', 'Only preference satisfaction is intrinsically valuable', and 'La Gianconda is valuable in itself' are among the stuff about which objectivists and subjectivists disagree as to how it should be *analyzed* (my italics). These claims express judgments about the source or locality of value, i.e., what value supervenes on. Such judgments about value sources are *bona fide* examples of *evaluative* claims. To endorse any of the above judgments is to take a clear evaluative stand (p. 249).

⁴³ Ronnow-Rasmussen mentions that the constitutive grounds of value about which subjectivists and objectivists disagree is spoken of mistakenly as a debate about the supervenience base of value.

⁴⁴ *Ibid.* p. 249.

(2007: 249). To distinguish the positions as disagreeing about what properties of an object are value-making (i.e. what is their *supervenience base*) would be to thus ‘infringe on the neutral character of the metaethical approach’⁴⁵.
(2007: 249)

If not about value-making properties, then *what is it* about which subjectivists and objectivists disagree? Roni-Rasmussen argues that it is precisely the *constitutive grounds* of value with respect to which subjectivism and objectivism are properly demarcated.

Constitutive grounds of value: what it is that *constitutes* values (i.e. the *source* of the value that a valuable object has).⁴⁶

That said, the subjectivist is claiming that:

Subjectivism: Whenever some object *O* is valuable, the value it has is there in virtue of the fact that a subject has (had or would have) a certain attitude toward *O*. (2006: 252)⁴⁷

Objectivists reject this claim; they deny that subjects—and in particular, subjects’ attitudes toward objects—are what it is that constitute the value that some valuable object has. The objectivist, unlike the subjectivist, can allow that something has value even if no one values it; the objectivist may just claim that the constitutive grounds of the object’s value are properties of the object itself. The subjectivist, however, claims that value requires valuing acts, and thus value wouldn’t be constituted merely by some object itself *independent* from whether there exists a certain attitude toward it.⁴⁸

3.3 Invariance

⁴⁵ *Ibid.* p. 249.

⁴⁶ One way to unpack this idea is to show how a disagreement about what *constitutes* values implies nothing about what it is in virtue of which some valuable object is valuable (i.e. what serves as the supervenience base). Consider that two positions might both agree that, for some valuable object *O*, the properties of *O* on which its value supervenes are $P_1, P_2 \dots P_n$, but disagree about what *constitutes* *O*’s value. Let us bring in here subjectivism and objectivism; and, given that these positions are *just defined* with respect to their rival pronouncements about the sources of constitutive grounds, articulating their respective positions on constitutive grounds is just to articulate the positions (full stop).

⁴⁷ This definition is taken from Roni-Rasmussen (2008) p. 252 with only slight (unsubstantive) rephrasing on my part

⁴⁸ Subjectivists disagree among themselves with regards the constitutive act of valuing, i.e., about whether an object is valuable *only while* it is being desired, or just so long as the agent would be disposed to desire it, were she to develop some attitude toward it, etc., however, they are united in maintaining that value is constituted *by* such attitudes. Subjectivists and objectivists could thus agree that the property of a piece of paper signed by Winston Churchill on which value supervenes is the part with the signature (and not the property of being a crisp piece of paper, etc.), but disagree about whether the value of the signed paper is constituted by a subject’s having a favorable attitude toward it. Now that it’s clearer about *what* the disagreement is, let’s consider now what some of the implications would be for supposing that the value of knowledge is constituted one way or the other.

As was mentioned earlier, Moore thought that from ‘the proposition that a particular kind of value is ‘intrinsic’ it does follow that it must be ‘objective’ (Moore 1922: 2) ... but this suggestion turns out to be premature. Consider, for example, Franz Brentano’s (1902) account of intrinsic value, which claims:

Intrinsic Value 1 [IV1]: A state S is intrinsically valuable if and only if mere contemplation of S requires that one approves of it. (Brentano, 1902)

IV1 is an account of intrinsic value compatible with subjectivism; the subjectivist could agree with IV1 and simply add that *when* the mere contemplation of S requires that one approve of it, the contemplation and subsequent approval of S are the *grounds* of S’s being valuable.

Consider also Roderick Chisholm’s (1986) contrastive account of intrinsic value.

Intrinsic Value 2 [IV2] P is intrinsically better than *q* only if it is appropriate for one to prefer *p* to *q* and the contemplation of just *p* and *q* requires one to prefer *p* to *q*. (Chisholm, 1986)

The subjectivist accommodates [IV2] much in the same way as it does [IV1]. The subjectivist can say that *when it is* that it is appropriate for one to prefer *p* to *q*, and this preference would be required just by the contemplation of *p* and *q*, the value *p* has is value constituted by the subject’s preferring *p* in this way—a way such that the preference is required by the contemplation.

Surprisingly, the subjectivist also has no problem, strictly speaking, with Moore’s (1903) own view.

Intrinsic Value 3 [V3] Some value bearer *V* has intrinsic value if and only if its value depends on (or supervenes on) the internal properties of *V*.⁴⁹

When ‘depends on’ in IVE is understood as about the *supervenience base* of an object’s value, IV3 is compatible with both objectivism and subjectivism. If, on the other hand, IV3 is understood as a claim about both the *supervenience base* and the *constitutive grounds* of value (a matter on which Moore isn’t entirely clear), then objectivism would be compatible with IV3, but subjectivism would (obviously) not be.

Suppose, though, we take IV3 to be a claim restricted to the *supervenience base* of value. Though subjectivism would be compatible with this reading of IV3 (a reading silent about *constitutive grounds* of value)

⁴⁹ This is as Moore’s (1903) account in his *Ethics*; my description of it is based on Rabinowicz and Ronnow-Rasmussen’s discussion of Moore in Rabinowicz & Ronnow-Rasmussen (1999)

there is nonetheless one special variety of this interpretation that would be off limits to the subjectivist. Such a variety would hold, in addition, an *invariance* requirement. *Invariance* claims that, for some valuable object V :

Invariance thesis: The value of V is invariant across possible worlds.

The amalgamated version of Moore's account of intrinsic value, along with the *invariance thesis* (to which some suspect he endorses)⁵⁰ would thus be:

Intrinsic Value 4: (i) Some value bearer V has intrinsic value if and only if its value depends on (or supervenes on) the internal properties of V ; and (ii) the value of V is invariant over possible worlds.⁵¹

Specifically, it is the invariance thesis in IV4—that is, (ii) and not (i)—that is incompatible with subjectivism; for the subjectivist could not explain why V maintains its value in worlds in which the constitutive grounds for the value of V are absent—worlds in which no subjects form (or would form) favorable attitudes toward V . Put simply, if intrinsic value implies *invariance*, then the subjectivist cannot account for intrinsic value.⁵²

⁵⁰ It has been suspected that Moore's 'isolation' test for intrinsic value—which is arguably substantively different from his 'analysis' of intrinsic value—commits him to the invariance theses. The isolation test is one that asks us to suppose, for some object O , that O is the only object in the universe. If O is still valuable in such a circumstance, then it is intrinsically valuable. Quite interestingly, Kent Hurtig has suggested that Moore's test here is incoherent: for example, how can 'Bob's knowledge that p ' (as an item of evaluation) be subjected to Moore's isolation test? Can we coherently suppose Bob's knowledge that p exists in some world in which Bob doesn't exist? Hurtig suspects that what Moore should have used for a test (for intrinsic value) is simply a test in which we ask, of a given object, whether it would retain its value in all possible worlds. Confessedly, Hurtig's suggestion might also falter as a test. After all, if we take "Bob's knowledge that p " and subject it to Hurtig's test, then the fact that Bob does not exist in all possible worlds might raise a similar incoherence problem as the one Hurtig levels toward Moore's isolation test. However, Hurtig could point out that we might be able to ask whether Bob's knowledge that p is valuable in all possible worlds without being committed to suppose that Bob is present in all possible worlds, in the sense that Moore's isolation test seems to require that Bob must be present in world in which his knowledge that p is all that exists. I'm not entirely convinced what 'test' for intrinsic value is appropriate, and so for the present purposes, I'm referencing only Moore's analysis of intrinsic value and will remain neutral about what test is best.

⁵¹ IV4 supplements the supervenience version of IV3 with what Ronnow-Rasmussen calls the *invariance thesis* (p. 257).

⁵² Our finding that *invariance* is incompatible with subjectivism might lead us to think that at least one way of demarcating subjectivism and objectivism *really does* require an evaluative commitment about *supervenience bases*—that is, that subjectivism, in virtue of being incompatible with *invariance*, is also incompatible with the claim that value supervenes on *intrinsic properties*, given that theses (i) and (ii) in IV4—*intrinsic value* and *invariance*—appear to mutually imply each other. *Invariance*, however, does not imply the supervenience claim associated with intrinsic value, *even if* this supervenience claim implied invariance. I shall explain this point by first by first introducing two related points:

(Invariance-1): Subjectivism is incompatible with possible world *invariance*.

(Invariance-2): *Invariance* does not imply the claim that value supervenes on an object's intrinsic properties.

Invariance is a thesis about the conditions under which an object's value would endure. Subjectivism is incompatible with *invariance* only because subjects' attitudes could not be said to serve as the constitutive grounds of value in all possible worlds. *Invariance*, as an

Now, this much tells us that—regardless of whether knowledge is intrinsically valuable—the following will be true:

- (1) If the value of knowledge is invariant across possible worlds, then the value of knowledge is not constituted by subjects. Correspondingly:
 - (2) If the value of knowledge is constituted by subjects, then it is not invariant across possible worlds.
- This is a significant result.

3.4 Subjectivism and Objectivism, Types and Tokens

In this section, I'll argue for the claim that knowledge as a *type* of epistemic standing has the value it does regardless of whether anyone values knowledge (objectively constituted value), though the same I shall claim is not true for knowledge *tokens* (subjectively constituted value). The first of these points is much easier to grasp than the latter, and I'll begin with it. Recall from Chapter 1 Kvanvig's (2008*b*) point about controlling for other non-cognitive values in order to assess *purely* epistemic value. His thought was that we might distinguish between the value something has in virtue of its relationship to cognitive states as opposed to affective states (desires, passions, cares, etc.). In order to assess whether something has purely cognitive (epistemic) value, you simply control for the values it would have in virtue of its relationship to desires, passions, cares, etc. What this meant was that, for example, some truth would have purely epistemic value only if it would be valuable for you to believe in a world where you cared about, desired, etc. nothing at all. We may suppose then that the value knowledge has as a *type* of epistemic standing is value that it would have even in worlds in which no one *cared* about knowledge. This is because the value of knowledge as a *type* of epistemic standing simply isn't constituted by subject's attitudes; accordingly, its value is not *subjective*. Unsurprisingly, this would explain why, from a *purely epistemic* point of view, knowledge should always be preferred to mere true belief—unqualifiedly, and even by people who desire not to know (indeed, or even despise knowledge). This is precisely what our Type Criteria for the Meno Requirement demands of a theory: that it can say what knowledge is in such a way that the satisfaction of those conditions it stipulates for knowing will ensure that knowledge is *essentially* more epistemically valuable than true belief.

endurance thesis about value, would not be incompatible with some account just because that account claims the supervenience base of an object's value is to be specified some particular way. *Invariance* is incompatible only with theses that imply that value does not *endure* as *invariance* says it does.

Now, we are in a position to begin resolving the matter of clarifying the relationship between the Type Criterion and a Token Criterion (and what a Token Criterion should look like). First off, an obvious point is that if the Type Criterion is met (by some theory of knowledge), then all tokens of that type (knowledge) will be ones the theory can uniformly explain to be more epistemically valuable than corresponding true belief tokens. After all, the satisfaction of conditions that the theory shows will make knowledge as a type more epistemically valuable than true belief as a type have to be satisfied in all instances where the type is instantiated. Thus, we can see why we should appeal to properties of knowledge as a type of standing to account for why knowledge tokens are epistemically valuable; this *reverses* what is normally done: trying to appeal to the value knowledge *tokens* generate over corresponding true belief tokens to account for why knowledge as a *type* is epistemically more valuable than mere true belief as a type.

3.5 Subjective Value and The Token Criterion

It is perhaps tempting at this point to suppose, in light of this last point, that (whatever the) Token Criterion is, it's something that would be straightforwardly implied by the Type Criterion, and that there's nothing more to be said. That is—if our theory of knowledge is one on which knowledge as a *type* of epistemic standing (in virtue of its formal properties, abstracted from propositional content) is more epistemically valuable than mere true belief as a type of epistemic standing, then we'll have secured the result that, on our theory, all tokens of the knowledge type will be more epistemically valuable than tokens of the true belief type. This much is true, although the Token Criterion I'll argue actually demands more than this, which is why (as we'll see) it stands apart from the Type Criterion as a distinct criterion that must be satisfied by a theory that preserves the Meno Requirement.

The problem of pointless truths presents us with a good place to start for grasping this point. Consider that some truths are so trivial or irrelevant that it seems as though there would be no positive value whatsoever generated by believing them. For example, the truth consisting in how many more grains of sand there are in Africa than there are hairs on all heads in Asia. Plausibly, if it is pointless to believe such a truth, then it will also be pointless to know such a truth—and by pointless I mean that there would be no positive value generated by believing such a truth. Unlike “bad” truths—which were discussed in some detail in Chapter 1—pointless truths aren't such that believing them generates negative value, all things considered. The thought is just that believing them brings about, all things considered, no value at all.

Now pointless truths pose an apparent threat to the claim that the epistemic value of knowledge is unrestricted, and it is a threat Kvanvig (2008*b*) addresses by appeal to an undercutter model of defeaters. According to Kvanvig⁵³, the undercutter model has just the right features here to reconcile *prima facie* positive epistemic value with apparent pointlessness (zero, rather than negative, all-things-considered value). He writes that:

[Undercutting defeaters] eliminate the *prima facie* positive aspect in question, without themselves revealing or supporting a negative aspect. (2008*b*: 14)⁵⁴

The idea is that the positive epistemic value of pointless truths can be undercut by an *absence* of any other positive values associated with believing the truth. Undercutting defeaters, Kvanvig claims, do not jeopardize the unrestricted purely epistemic value of such truths, given that the unrestricted value claim is *prima facie* and defeasible (and not an all-things-considered claim).⁵⁵

I think Kvanvig is right about this, although it's outside the present scope for me to give any detailed defence here. I only bring up the problem of pointless truths because it clarifies how the undercutter model is supposed to work—and the undercutter model shows quite clearly how it is that the epistemic value of knowledge and mere true belief, *qua* tokens, are in a certain way hostage to subjects' attitudes which themselves constitute this value.

This point perhaps requires some more clarification. The thought is that some truths (e.g. long arbitrary conjunctions) are so pointless that zero positive value is generated by knowing them. This does *not* mean that such truths lack any *prima facie* positive epistemic value. The *prima facie* value is simply defeated—it is undercut, rather than overridden (as in cases of bad truths, as was discussed in Chapter 1)—by an absence of any practical value that would be generated by believing them. Now, when truths aren't pointless, there is at least some practical value generated by believing them, usually positive, but sometimes negative, as in the case of bad truths. Believing some truth will generate practical value for me *only if* there are some things that I value (e.g. achieving certain ends that believing this truth promotes). Consequently, the epistemic value generated by

⁵³ I owe this explication to a January 2008 correspondence with Kvanvig.

⁵⁴ Generalizing from the area of epistemic value, we might best convey the undercutter model by drawing from an example from David Lewis. Suppose there is a deadly void—deadly not because of anything it 'does' to you, but deadly because it doesn't provide you with oxygen, and you need oxygen to live. Now let's say that, *prima facie*, it's valuable, from a purely health-based perspective, to have a fully-functional human body. In the void, obviously, the value having such a body is defeated. The conditions necessary for a fully-functioning human body to be valuable are absent, and so the value your body has insofar as it's fully-functioning is undercut. Nonetheless, though, it's *prima facie* valuable (from a health standpoint) to have a fully-functioning human body.

⁵⁵ This is so just as (pace *fn.* 86) the possibility of deadly voids doesn't threaten the *prima facie* health-based value of having a fully functioning body.

my believing a given truth is (unlike the value of knowledge or true belief qua *type* of epistemic standing) *constituted* by my subjective attitudes, as these attitudes determine (among other things) what truths would be pointless for me to believe and which truths will instead have positive epistemic value (e.g. value that is not undercut by a lack of any other positive values generated by believing them).

3.6 Proportionality Thesis and the Token Criterion

In the previous section, I tried to make clearer how we should think of subject's attitudes as the constitutive grounds of the value of *token* beliefs. Having hopefully now motivated this idea sufficiently (and given a idea how this is supposed to work), we are finally in a position to consider the Token Criterion I want to defend alongside the Type Criterion. As the undercutter model suggested, whether or not believing some truth would generate any positive value insofar as my interests are concerned is a matter that itself will itself determine whether the prima facie positive epistemic value all truths have will be undercut or not. Now, recall that if my interests render it that the truth is pointless for me to believe, the prima facie epistemic value is undercut; plausibly, something like the converse it also true. That is—the more non-epistemic positive value that would be generated by my believing some truth, then correspondingly, the more epistemic value will be generated by my believing that truth. This could explain why we're particularly curious about the truth of matters we care about—and these are matters of which the truth is something we take to be *important*. This idea is akin to Sosa's (2007) suggestion that:

Something might be known far better, with greater certainty, and better justification, than something else, while yet the latter knowledge is intellectually finer by far. Compare knowledge that one's back hurts now with some deeply illuminating knowledge about a friend, or a historical period, or a novel, or a scientific theory. (Sosa 2007: 88-89)

Now, I want to take this idea a step further because I think a very natural idea about how to frame the Token Criterion for the Meno Requirement is right in the neighborhood. The idea is that is the more non-cognitive value (i.e. value in virtue of some relation to my affective states, i.e. desires, etc.) that will be generated by believing some truth, the more cognitive value will be generated by believing it—and further, the more epistemic value that would be generated by *knowing* it. This much is rationalized by our diagnosis of pointless truths.

If that's right, then the next idea should follow quite naturally: the idea is that the extent to which knowing some truth is more epistemically valuable than merely truly believing it will be proportionate to how

important (as opposed to pointless) it is that the truth is known. Call this the *proportionality thesis*. What I want to defend as the Token Criterion is explicable directly in terms of the proportionality thesis:

Token Criterion: (Meno Requirement): The Meno Requirement holds a theory of knowledge to preserve that, from a *purely epistemic point* of view, the extent to which a known belief token will be more (epistemically) valuable than a corresponding true belief token will be directly proportionate to how important (as opposed to pointless) it is that the truth is known.

The *proportionality thesis* captures what should be a rather familiar idea, which I'll illustrate as follows. Take a paradigmatically pointless truth: that there are exactly n motes of dust on Mars. There is very little practical (non-epistemic) value that would be generated by knowing this truth. According to the *proportionality thesis*, the extent to which knowing that truth would be, from a purely epistemic point of view, better than merely truly believing it will be marginal. Now suppose we up the stakes: the truth at issue is what time the bank opens (to borrow from Cohen's (1998) classic example), and further, you realize the consequences will be devastating if you are wrong about this (e.g. your brother who you intended to wire money will be killed). Here, the non-epistemic value that would be generated by your believing this truth will be significantly higher than it will be in the mote-of-dust case. According to the *proportionality thesis*, it follows that the extent to which knowing *that* truth is more epistemically valuable than merely truly believing it is greater than the extent to which knowing the mote-of-dust truth will be more epistemically valuable than truly believing that. Thus, a theory of knowledge must preserve this: that a given knowledge *token* is more epistemically valuable than a corresponding mere true belief *token* to an extent proportionate to the non-cognitive value that would be generated by believing the target proposition. After all, a theory of knowledge shouldn't lack an explanation for this and simply imply that it's equally better epistemically to know the bank proposition rather than to truly believe it as it is to know the dust truth rather than to merely believe it.

Taken together now with the Type Criterion, we have two distinct criteria that form a "test" for whether a theory of knowledge will satisfy the demands of the Meno Requirement.

A "Test" for Determining Whether a Theory Satisfies the Meno Requirement

Type Criterion (Meno Requirement): The Meno Requirement holds a theory of knowledge to preserve that, from a *purely epistemic point* of view, knowledge *qua type* of epistemic standing (on the theory) has a purely epistemic value that exceeds that of mere true belief *qua type* of epistemic standing.

Token Criterion: (Meno Requirement): The Meno Requirement holds a theory of knowledge to preserve that, from a *purely epistemic point* of view, the extent to which a known belief token will be more (epistemically) valuable than a corresponding true belief token will be directly proportionate to how important (as opposed to pointless) it is that the truth is known.

3.7 Concluding Remarks

A theory satisfies the Meno Requirement, then, only if knowledge is (on the theory) essentially more epistemically valuable than mere true belief, as the two are compared as *types* of epistemic standings. Further, a theory of knowledge satisfies the Meno Requirement only if a given knowledge token is (on the theory) going to be more epistemically valuable than a corresponding true belief token to an extent that is proportional to the non-cognitive value that would be generated by believing that particular truth.

Now, embedded within these criteria is the notion of *epistemic value*. We've described epistemic value generally as value that something has from the epistemic (rather than practical, aesthetic, etc.) point of view. However, there are competing views about what would make something valuable from the epistemic point of view; veritists think that something is epistemically valuable to the extent that it promotes truth; pluralists deny this claim and argue that there are multiple sources—not limited to truth—in virtue of which something could be epistemically valuable. We'll need to adjudicate between these two competing positions in order to actually defend some theory as meeting the criteria required to satisfy the Meno Requirement. Simply, we need to know what it is in virtue of which something is more epistemically valuable than something else. This will be the topic of the next chapter.

4

The Epistemic Point of View

4.1 Introduction

The topic that will be pursued here is that of epistemic value generally, and more specifically, epistemic evaluations. The aim will be to investigate what makes something more or less good from an *epistemic* point of view—that is, from the point of view from which it is our epistemic purposes that matter.

Epistemologists tend to align themselves with one of two particular ways of thinking about the epistemic point of view. I'll refer to the first way as *veritism* (truth monism)⁵⁶, and the second as *pluralism*. What divides veritists and pluralists in epistemology is the role truth should be thought to play in epistemic evaluations. According to the veritist, truth is the fundamental epistemic aim. This is not the claim that nothing but truth is epistemically valuable; rather, the idea is that something is more or less good epistemically in virtue of its being the sort of thing that promotes truth (or avoids error). Accordingly, then, the veritist says that:

The epistemic point of view (veritism): The point of view from which what matters is maximizing truth and avoiding error.

Veritism has a venerable tradition in contemporary epistemology; as a thesis, it is both simple and forceful. The pluralist view, on the other hand, is comparatively less elegant. According to the pluralist, truth is but one of several, perhaps many, epistemic aims. And accordingly, the pluralist will tell us that the epistemic point of view

⁵⁶ By veritists, I mean epistemic value monists who take truth to be the fundamental epistemic aim or value. Although veritists stand opposed to any variety of epistemic value monists (i.e. epistemic value monists who hold some other end than truth to be the fundamental epistemic aim), the pluralist position is generally thought of as a rival to the variety of epistemic value monism that extols truth specifically. I take the fundamental divide on this score then to be between pluralists and veritists rather than pluralists and any type of epistemic value monist.

is one from which truth is not the only thing that matters, and that truth does not exhaust the possible sources of epistemic value.

Given veritism's comparative explanatory power along with its simplicity, it should be no surprise that this view has been a favourite axiology in mainstream epistemology. Veritism has the advantage of parsimony for one thing, and this parsimony makes possible a straightforward way of explaining why, for example, one belief but not another counts as justified or why one trait is an intellectual virtue and another a vice.

But is veritism true? On the basis of *what* might we think that veritism is true and pluralism is false? This will be the central issue I'll be exploring here. In what follows, I will first make precise what it is we should take the veritist to be arguing for and then consider how the most viable veritist line holds up against a variety of challenges in recent literature. Although no strategy for defending veritism will ultimately turn out promising, I argue that a core idea of veritism can nonetheless be preserved within a particular kind of pluralism—one that upholds the truth goal alongside a distinct goal that governs agency.

4.2 The strong case for veritism

It is common to say that beliefs aim at truth. Perhaps equally common, though, is the thought that the beliefs we hold should be rational or justified. Do we want our beliefs to be rational and justified simply because we want to have rational and justified beliefs? More likely, it seems, we want rational and justified beliefs because they are markers of truth, which is what we really want. If that is right, then a neat and simple picture seems to emerge: we think our beliefs are better or worse to hold on the basis of their propensity to guide us to the truth, and we also think of beliefs themselves that they *aim* at truth. It is quite natural then to connect these observations and suppose that the latter fact must play a role in explaining the former: that it's *because beliefs aim at truth* that we inquire, as believers, with truth as our aim.

It's easy to see how the connection between that at which beliefs aim and that toward which inquirers aspire—truth—could be used not just as a simple picture that explains how we evaluate beliefs, but also as a picture that by extension would apply to evaluations within the wider discipline that evaluates beliefs, epistemology. And indeed it has. Goldman (1999), for example, conceives of epistemology as straightforwardly 'a discipline that evaluates...along truth-linked (veritistic) dimensions' (1999: 41-46) And what's wrong with making this move? After all, not only beliefs, but also cognitive faculties and methods of inquiry seem answerable to the same standard: it is against the background goal of getting to the truth that we're inclined to

judge some cognitive faculties, belief forming processes, methods of inquiry, etc. better than others. It's easy, then, to see how the thought that truth is our fundamental epistemic aim is compelling: beliefs aim at truth, and epistemic agents, qua believers, inherit this aim, which in part explains why we hold as epistemic values such things as rationality, justification, memory, and those traits we call intellectual virtues; they are epistemically valuable because truth is our goal and they help get us to it.

4.3 The normative and teleological accounts of truth-directedness

If the fact that belief aims at truth is supposed to play a role in explaining why truth is the fundamental epistemic value, as the strong case for veritism I just outlined suggests, then we should insist on two arguments: one which clarifies and defends the claim that beliefs aim at truth, and another that establishes how it is that the truth-directedness of belief would stand to support the veritist's claim that truth is the goal of inquiry.

Let's look first at the claim that beliefs aim at truth. As Engel (2004) points out, this claim is at best metaphorical: 'Obviously beliefs do not 'aim' at anything by themselves, they do not contain little archers trying to hit the target of truth with their arrows.' (Engel 2004: 77). So what does it mean then to say that beliefs aim at truth?

According to Engel, the claim *beliefs aim at truth* is best understood as one expressing the idea that truth is the norm of belief. Beliefs, he says, have *normative* properties, which means that 'if certain properties are not instantiated, they are incorrect or wrong.' (Engel 2004: 81). An important normative property of beliefs, Engel claims, is that they are correct only if the proposition believed is true. He calls this the *truth norm*:

Truth Norm (Engel): For all P, one ought to believe that P only if P. (2004: 82)

Along with Engel, Lynch (2008) also endorses the idea that truth is the norm of belief. According to Lynch:

Truth Norm (Lynch): It is prima facie correct to believe <p> if and only if <p> is true. (Lynch 2008: 4)

It's important here to note that both Engel and Lynch take the truth norm to be more than just an observation about beliefs, but moreover a constitutive fact about the very concept of belief. Says Engel:

...a subject cannot ascribe to herself a belief unless she recognizes that belief is constrained by the norm of truth...in this respect the connection between the very concept of belief and the epistemic norm of truth is conceptual or constitutive."

(Engel 2004: 83)

This idea is shared by Velleman (2000) who proposes that ‘the concept of a belief is just the concept of an attitude for which there is such a thing as correctness or incorrectness, consisting in truth or falsity.’ (Velleman 2000: 16) Lynch, drawing from Velleman, follows suit and says that the truth norm ‘states a constitutive fact about belief’ (Lynch 2008: 7).

If Engel, Velleman and Lynch are right that part of what it is for something to even *be* a belief is that it counts as correct (or, *pace* Engel, one you ought to hold) only if true, then the claim that belief aims at truth is a *conceptual* claim; that beliefs are truth-directed is part of what makes them count as beliefs. Now we may ask: does this imply anything about whether truth is the fundamental epistemic aim? It certainly seems plausible that it would; after all, the truth norm is taken to be a normative constraint on belief.

Lynch argues for just such an implication; he thinks that by engaging in the practice of inquiry you thereby assent to recognizing the value that governs the practice, and for the practice of inquiry—which consists in forming beliefs—that value is truth. (Lynch 2008: 14) And so for Lynch, then, you are committed to recognizing truth as the value that governs inquiry just by engaging in inquiry at all.

Lynch’s move here commits what Engel takes to be the mistake of supposing that:

in order to be responsive to the norm of truth, believers must recognize this norm and reflect upon it... (Engel 2004: 88)

Engel adds:

...believing that p entails believing that p is true, but it does not involve any attitude towards the proposition ‘P is true.’ It only involves an attitude towards P.” (Engel 2004: 88).

Engel here distinguishes between the *normative* account of truth-directedness and the *teleological* account of truth directedness. Whereas the normative account says only that it is constitutive of the concept of belief that it is governed by the truth norm, the teleological account identifies believing with both ‘(a) the conscious recognition of the basic norm of truth and (b) the intention to respect and maintain this norm in the formation of one’s beliefs.’ (Engel 2004: 84) We can see, then, that veritism will be an implication of the claim that belief aims at truth, then, only if the teleological account of truth-directedness is correct as opposed to the weaker normative account of truth-directedness. But, as it seems, the teleological account is correct only if responsiveness to the truth norm requires reflecting upon it, or having some attitude toward it. While these requirements would appear to hold for intentional acts such as ‘accepting’ or ‘judging,’ it remains that ‘the believer of an individual belief does not consider the aims of inquiry when he comes to believe a given

proposition.’ (Engel 2004: 86) To note, this is entirely compatible with the thought that the truth norm is a constitutive part of what belief is. Says Engel:

Being able, at least tacitly, to recognize the epistemic norms of beliefs is certainly a necessary condition for having the concept of a belief. But in order to be a believer, one does not need to attend reflectively to these norms. (2004: 88)

Consequently, we see how it is that the strong case for veritism as a thesis motivated by the truth-directedness of belief loses its purchase power. Engel puts this point aptly:

So the fact that belief aims at truth (in the normative sense) is one thing, and the fact that inquiry has truth as its goal is another thing... the believer in an individual belief is not someone who contemplates the general goal of aiming at truth and acts upon it. (2004: 88)

So if veritism is true and truth is the fundamental aim of inquiry, this wouldn’t simply be because belief aims at truth. If veritism is true, it must be defended some other way.

4.4 A different case for veritism

I’ve argued to this point that if veritism is true, it won’t be because beliefs aim at truth. Now I want to defend a more simple point, which is that if veritism is true, it couldn’t be a *conceptual* truth. How might one defend veritism as a conceptual truth? A tempting strategy would be to take as a starting point something like Sosa’s (2007) suggestion that ‘Evaluation is distinctively epistemic when it is concerned with truth’ (Sosa 2007: 78) and then try to show that veritism is straightforwardly implied by this apparent truism. Consider the claim:

Conceptual line: Something is evaluated epistemically iff what matters for the purposes of evaluation is just getting to the truth.

The conceptual line states a rather intuitive idea, and veritism is directly implied by it; for truth will be the fundamental epistemic aim if indeed getting to the truth is the goal against which epistemic evaluations are made. But can the veritist really defend veritism simply by offering up a claim about what is constitutive of epistemic evaluations? It should be clear that such a strategy would be ineffective and certainly incapable of gaining any traction against the opponent who endorses the pluralist view. The reason this sort of conceptual line defense of veritism would be ineffective is that ‘epistemic’ functions here in a premise that would serve to support veritism rather than pluralism only if ‘epistemic’ is *already taken to mean something* it would mean only if veritism (the conclusion) rather than pluralism is true. More plainly: the veritist could never persuade the pluralist with an argument whose premise is one that could be accepted only by one who already accepts the

veritist's conclusion. The veritist, therefore, cannot argue in a way that assumes that the epistemic point of view is just the point of view from which getting to the truth is what matters. After all, how we should characterize the epistemic point of view is specifically a point of cointention between the two positions, and so the pluralist could not in principle be moved to conviction by this sort of reasoning.

The veritist must therefore not assume in her argument that 'epistemic' means something the pluralist could not accept. What we may insist on from the veritist then is some non-questionbegging argument that establishes veritism to correctly characterize *whatever point of view it is* from which the considerations we weigh when making evaluations are just those that matter in light of our properly epistemic (understood non-committally) purposes. This is what we should insist on from the pluralist as well.

It should now be a bit clearer what a successful defense of veritism would have to look like. And this said, I think William Alston (2005) provides a helpful way of understanding how the epistemic point of view might be characterized, non-questionbeggingly, along veritist lines. Here's Alston:

We evaluate something when we dub it good, bad, or indifferent for some purpose or from some point of view.... We evaluate something epistemically (I will be mostly concerned with the evaluation of beliefs) when we judge it to be more or less good from the epistemic point of view, that is, for the attainment of epistemic purposes....The evaluative aspect of epistemology involves an attempt to identify ways in which the conduct and the products of our cognitive activities can be better or worse vis-à-vis the goals of cognition. And what are those goals? Along with many other epistemologists I suggest that the primary function of cognition in human life is to acquire true beliefs rather than false beliefs about matters that are of interest to us. (Alston 2005: 28)

Importantly, what Alston says here allows that the veritist and pluralist can agree that the epistemic point of view is the point of view from which what matters is determined by *epistemic purposes* while at the same time disagreeing about what these purposes actually are.⁵⁷ And so here we see wherein the substantive debate lies. Accordingly, then, we find ourselves in a position to assess whether veritism is correct by considering what our epistemic purposes in fact are. If they are just the purposes we have insofar as we're after the truth, then the veritist has got things right, and the pluralist goes home emptyhanded.

4.5 Kvanvig, David and the "Knowledge Objection" to Veritism

⁵⁷ Here Alston explains why the purposes that count as 'epistemic' are those directed at acquiring true rather than false beliefs. Because he thinks that a viewpoint from which things are evaluated as good or bad will be an 'epistemic' viewpoint in virtue of its being a viewpoint from which things are good or bad *with respect to our epistemic purposes*, it will follow that the epistemic point of view is viewpoint from which things are more or less good insofar as they promote the purpose of attaining true beliefs rather than false beliefs—the veritist's claim.

Our aim then is to figure out whether the veritist is right about what our epistemic purposes are. If some of our epistemic purposes are ones we have not only insofar as we're after truth, then veritism is in trouble. Against this background, I want to introduce what might seem at first to be a simple and obvious objection to veritism along these lines, but which turns out to be quite complicated and one that will deserve our sustained attention. Let's call this the *knowledge objection* to veritism; it is one that has surfaced most notably in an exchange between David (2001) and Kvanvig (2005). What I call the *knowledge objection* takes as its initial premise the conclusion we reached at the end of the previous section:

The "Knowledge Objection" to Veritism

- (1) If veritism is true, then our epistemic purposes are just the purposes we have insofar as we're after the truth (i.e. maximizing true beliefs, avoiding false beliefs). (Sec 5.4)

The next step in the objection is to offer up something that seems uncontroversial:

- (2) Our epistemic purposes consist in part in purposes we have insofar as we're after *knowledge*.

But what is jointly implied by (1) and (2) is:

- (3) Therefore, veritism is false.

This simple argument carries with it some compelling force.⁵⁸ After all, why should we think—as the veritist does—that epistemic purposes must be restricted to those purposes we have insofar as we're after something to which we would most always prefer knowledge? It should be obvious here that the Meno Assumption lends some *prima facie* credence to Premise (2). As Kvanvig (2009*b*) puts it:

When asked, “why, from a purely cognitive point of view, do you value knowledge?” we are inclined to answer...that we want to be correct, but not merely by accident, as what happens when one has a merely true belief. (Kvanvig 2009*b*: 7)

⁵⁸ Suppose, for example, that Bill truly believes that his wife is having an affair. Though he's already got a true belief, he nonetheless desires to *know*. Of Bill's situation, we might be inclined to think that the purposes Bill has insofar as he's after knowledge are purposes over and above, and distinct from, the purposes he has insofar as he's after what he's already got (true belief). And if the latter purposes matter from an epistemic point of view, then why shouldn't the former? To put the question differently:

Insofar as we want knowledge rather than mere true belief, then, we seem to have epistemic purposes distinct from the purposes we'd have if only wanting, as Kvanvig says, to 'be correct...by accident.'" (2009*b*: 7)

In sum, then: Premise (1) in the *knowledge argument* is an articulation of the substantive point the veritist must be making in order to avoid begging the question against the pluralist, and (2) captures the seemingly uncontroversial idea that our epistemic purposes, at least some times, are those we have insofar as we want to *know*, and not just merely truly believe. But (1) and (2) jointly entail the denial of veritism. Could objecting to a thesis as widely held as veritism really be this easy?

4.5.1 Marian David's "Circularity" Response to the Knowledge Objection

For obvious reasons, the knowledge objection is one to which the veritist had better have a quick and ready reply. According to Marian David (2005)⁵⁹ the best way to meet this objection is to call a circularity foul on proponents of the objection. Here's David:

Although knowledge is certainly no less desirable than true belief, the knowledge-goal is at a disadvantage here because it does not fit into this picture in any helpful manner. Invoking the knowledge-goal would insert the concept of knowledge right into the specification of the goal, which would then no longer provide an independent anchor for understanding epistemic concepts. In particular, any attempt to understand justification relative to the knowledge-goal would invert the explanatory direction and would make the whole approach circular and entirely unilluminating (David 2005: 153-154).

David's beef with the knowledge-goal objection to veritism, then, is that by placing knowledge as the epistemic goal, we would be forced to give a theory of justification that explains justification by appealing to knowledge. For example, the reliabilist—if trading the truth goal for the knowledge-goal—would have to say that a belief is justified insofar as it is produced by a process that reliably gets to knowledge. But if this is how the reliabilist understands justification, then her theory of knowledge as true belief that reliably gets to knowledge will be circular. David sums up the point:

...[That] Knowledge was supposed to be explained in terms of justification and not the other way round...does not mean that it is wrong in general to talk of knowledge as a goal, nor does it mean that epistemologists do not desire to have knowledge. However, it does mean that it is bad epistemology to invoke the knowledge-goal as part of the theory of knowledge because it is quite useless for theoretical purposes: The knowledge-goal has no theoretical role to play *within* the theory of knowledge (David 2005: 153-154)

David's point about circularity is initially compelling: a theory of knowledge that appeals to knowledge in that theory's definition *of* knowledge would be circular. But David's reply here becomes far less compelling

⁵⁹ Philip Ebert has noted in discussion a similar concern with the knowledge goal.

under closer scrutiny.

4.5.2 Kvanvig's Response to David

Now Kvanvig (2005) has offered up a rather provoking *reductio* to David's reasoning. First, Kvanvig points out that David has assumed that epistemology is the theory of knowledge (a point he makes earlier), and secondly, David takes it that 'epistemic evaluation... is most naturally understood along broadly teleological lines, as evaluating beliefs relative to the standard or goal...' (Kvanvig 2005: 10). Such a picture leaves, as Kvanvig notes, justification as the 'only epistemic concept *within* the theory of knowledge' (Kvanvig 2005: 10)—that is, within epistemology. With these assumptions in place, the only candidates for epistemic goals other than truth will be justification or knowledge itself. Between knowledge and justification, knowledge stands to be the obvious competitor. But because knowledge as the goal was thought to lead to vicious circularity within a definition of knowledge, only the truth goal remains as a candidate. And so by process-of-elimination, veritism must be true.

Taking this to be David's reasoning, Kvanvig argues, at the outset, that:

The form of argument here bears scrutiny since the particular argument in question relies on the false assumption that epistemology is the theory of knowledge. (Kvanvig 2005: 10)

Kvanvig is I think correct to take note that it was this assumption that made it seem inevitable that either knowledge or truth must be the epistemic goal, the position from which David argued against knowledge on grounds of circularity and in favor of truth. Because Kvanvig thinks that the epistemology-as-theory-of-knowledge assumption is mistaken, he says that we will need a generalization of David's argument in order to conclude that truth is the primary epistemic goal. (2005: 10) Here's Kvanvig:

Such a generalization could begin by allowing that any purely theoretical cognitive success is of value and hence a suitable epistemic goal. But for each such goal, it has no theoretical role to play within the project of theorizing about it, for such an account would make the theory 'circular and entirely unilluminating.' Moreover it would have to be claimed, any goal that would render a theory circular and unilluminating in this way cannot be an epistemic goal, leaving truth as the only standing candidate for that role. So, on this [David's] view, *nothing of epistemic value can be an epistemic goal*. Instead, only values that fall outside the domain of cognitive successes can legitimately be cited when explaining epistemic success.⁶⁰ (Kvanvig 2005: 10)

In a nutshell, then, Kvanvig is saying that once the epistemic goal—that is, whatever it is that matters primarily from the epistemic point of view—is taken to be the goal of epistemology insofar as epistemology is *just* a theory of knowledge, we would be in a position to (as David does) discount the knowledge-goal as circular

⁶⁰ *My italics.*

and be (by process of elimination) veritists only at the expense of having to grant both that (i) nothing of epistemic value can be an epistemic goal, and (ii) that only values that fall outside the domain of cognitive successes can explain epistemic successes.

If Kvanvig is right, then David is paying simply too heavy a price to pay in order to prop truth up as the epistemic goal rather than knowledge. First off, it's no good being stuck with the counterintuitive consequence that nothing of epistemic value could be an epistemic goal.⁶¹

Similarly problematic is the other implication of David's view that Kvanvig draws attention to, which is that only values that fall outside the domain of cognitive successes can explain epistemic successes. Just consider that central proponents of veritism (e.g. Goldman) commonly articulate the epistemic goal twice over: as the goal of truth, but also as the goal of maximizing true beliefs. After all, when the veritist tells us which traits count as virtues, she points to whichever ones best lead us to form true beliefs. But unlike 'truth' (full stop), 'true beliefs'⁶² are a variety of cognitive successes, and straightforwardly so. So David's circularity response to the knowledge objection to veritism would have the bizarre consequence that veritists cannot explain justification in terms of true belief—a consequence that would mean no reliabilist could be a veritist.

4.5.3 What does the knowledge objection tell us about veritism?

We must resist the temptation to infer, from the previous discussion, that the knowledge objection is sufficient for undermining veritism; all we've shown is that the knowledge objection can't be rebutted via the circularity argument David gives. Even more specifically, though, what was shown was that the veritist can't defend against the objection that knowledge is an epistemic goal over and above truth by arguing in a way that assumes that epistemology is nothing more than the theory of knowledge. It is appropriate to ask, then, whether the veritist could reply to the knowledge objection by appealing to some presumed circularity at issue with the knowledge-goal whilst at the same time *not* restricting, as David did, epistemology to simply the *theory of knowledge*.

⁶¹ For one thing, the thought that nothing of epistemic value can be an epistemic goal would be at odds with David's conception of epistemic evaluation "along broadly teleological lines." For if we think of epistemic evaluation along teleological lines, we would have to understand, for example, the epistemic value justification has as *instrumental* epistemic value—value justification has as a means to promoting some epistemic end. But it is a prominent view in value theory (i.e. Korsgaard (1983); (Rabinowicz (2000) that the instrumental epistemic value justification should be thought to have could only be understood in the context of its promoting whatever is *non-instrumentally* epistemically valuable, though epistemically valuable nonetheless. That is, whatever it is that is epistemically valuable for its own sake, and not for the sake of something else. That said, it's a problematic outcome, at best, for David to be forced to deny that truth—his epistemic goal—is epistemically valuable. It's not clear in fact how to even make sense of such a claim.

4.5.4 Ecumenical veritism

As Hillary Kornblith (2002) has observed, veritists indeed do have the tendency to, much as David has seemed to, ‘reduce epistemology to the theory of propositional knowing.’ (Axtell & Carter 2008: 3). But that’s just an accidental feature of many veritist views; they don’t have to think of epistemology this way. Let’s see how veritism might fare alongside a wider conception of what epistemology is. Toward this end, let’s think very basically: if we take metaphysics to be concerned predominantly with the way things *are*, then epistemology, in an equally broad sense, might plausibly be thought of as concerned with our cognitive account of how things are. Wayne Riggs (2004) offers a helpful way to think about our plight in the broad sense of representers of the world:

I believe that we humans, as cognitive beings, ultimately strive for a representation of the world and our place in it that is comprehensive, accurate and intelligible to us. And the cognitive apparatus we have to achieve this is fallible, limited, and largely dependent on the contributions of others. (Riggs 2004: 1)

Accordingly, it should seem that whatever else epistemology is a theory of, it should be a theory that concerns both what it is we strive for in our representations of the world and what it is that makes us more or less apt to succeed. Now Sosa (2007) might be inclined to insist on caution here by drawing the distinction as he has between the theory of knowledge and intellectual ethics and supposing that we risk here conflating the two.

He writes:

Accordingly, we do well to distinguish between two parts of epistemology: (a) theory of knowledge, and (b) intellectual ethics. The latter concerns evaluation and norms pertinent to intellectual matters generally, with sensitivity to the full span of intellectual values. It is therefore a much broader discipline than a theory of knowledge focused on the nature, conditions, and extent of human knowledge. (Sosa 2007: 89)

I think Sosa is right to point out, as he goes on to, that some evaluations about beliefs are based on considerations independent from whether the beliefs are accurate (e.g. such as whether they are morally good to have), and I’ll agree that such normative evaluations fall outside of the area of epistemic normativity; these will not be epistemic evaluations. But we can grant this while, at the same time, granting that epistemic normativity ranges over inquiry in a broader sense than David, for example, has assumed.

Hookway (2000; 2006), for one, takes epistemology—and epistemic normativity—to range over matters concerning how we are able to carry out our inquiries in a well-regulated manner. This core idea has been adopted not only by epistemologists with pragmatist leanings like Hookway, it should be noted, but also by

philosophers steeped in the areteic tradition, such as Hursthouse (1999) and Zagzebski (1996).⁶³

There's no reason the veritist can't take on board this wider purview. Let's call veritism free from the 'epistemology-as-a-theory-of-propositional-knowledge' assumption *ecumenical veritism*. The question becomes: are our (ecumenical) *epistemic* purposes—purposes we have insofar as we aim to *do what we can* to best *tell how things are*—just the purposes we have insofar as we're after the truth (i.e. maximizing true beliefs, avoiding false beliefs)?

Enter again now the knowledge objection. Ecumenical veritism would indeed avoid the problems associated with the initial 'theory of knowledge' assumption with which Kvanvig took issue, and it would do so at the cost of no longer being able to frame just truth *or* knowledge as the only candidate epistemic goals. (It was this dichotomy, recall, that led David to think truth was the only candidate goal left standing). To be clear, though, the cost of giving up this dichotomized picture of the candidate epistemic goals is that the veritist is no longer in a position to conclude that the circularity of the knowledge goal would *imply* veritism (by process of elimination). It remains open, though, for the ecumenical veritist to try to defend herself against the knowledge objection by arguing that the knowledge goal is circular.

Perhaps the best case that could be made to this effect would be to argue that *regardless of what epistemology is a theory of*, it is at least in part concerned with what knowledge is. And so if, as the objection goes, knowledge is an epistemic goal distinct from truth, then David's worry seems to reemerge: justification would be defined in terms of knowledge, and so knowledge can be defined in terms of justification only at pain of circularity.

This strategy is less ambitious than the one David himself gave us in that it doesn't attempt to establish that veritism is *true*—it aims only to establish that the knowledge objection fails to show that veritism is false. It's also more plausible than David's in that it doesn't rely on an overly restricted purview of epistemology. Despite these relative improvements David's circularity charge has with this new-and-improved ecumenical view of epistemology's subject matter we've equipped it with, the circularity charge will still miss the mark. This is because it *hits* its mark only if the pluralist proponent of the knowledge objection could explain those

⁶³ This agent-centered conception of epistemology, with a focus on epistemic practice as well as epistemic ends, chimes closely with the view endorsed by Zagzebski (1996). While Zagzebski holds the rather controversial view that epistemic success is a subspecies of moral success, she maintains the rather comparatively compelling view that the measure of epistemic success—that is, the end against which things are measured as good or bad epistemically—is the end consisting in epistemic flourishing, or alternatively, the 'epistemic good life.' Hursthouse (1999) holds a similar view here. On this note, one might think, for example, of Aristotle or Socrates as archetypal examples of epistemic flourishing. On such a picture, Hookway's end of "being a good inquirer" or Zagzebski's end of epistemic flourishing would leave us saying that for some given item of evaluation O, O has epistemic value to the extent that it promotes the fundamentally valuable epistemic end of being the sort of inquirer that we take Aristotle or Socrates to have been.

particular concepts in terms of which knowledge is defined (i.e. justification) only by appealing to knowledge—that is, only if the pluralist, when defining those concepts that knowledge is defined in terms of (let’s just use justification for simplicity), is committed to what I’ll call the *unified aim thesis*:

Unified aim thesis: If theory T recognizes E1, E2...En as epistemic goals, then whatever account of justification T gives must be one that explains justification in terms of E1, E2...En.

If a pluralist offers the knowledge objection to veritism while at the same time committed to the *unified aim thesis*, then her theory of knowledge would indeed be circular; knowledge *qua* an epistemic goal would function indispensably within an account of justification and will thus be imbedded inevitably and perniciously within a theory of knowledge.

But there is no reason that someone who offers the knowledge objection must also be committed to the *unified aim thesis*. More specifically, nothing about the mere countenancing of multiple epistemic goals would count for or against what theory we should hold about which *particular* goals are to be appealed to when evaluating something positively or negatively epistemically.⁶⁴ It is thus open to the pluralist to explain what makes a belief justified in terms of some epistemic aim distinct from knowledge while at the same time allowing that knowledge is an epistemic aim in its own right.

4.5 The ‘collapsing’ defense

I think it is only fair that we consider here briefly a line of response the veritist has available against the knowledge objection that relies *neither* on the assumption that epistemology is the theory of knowledge *nor* on the supposition that the knowledge-goal is circular. A remaining idea that the veritist could appeal to would be as follows: the purposes we have insofar as we are aiming at knowledge will be the very same purposes we have insofar as we are aiming at truth, or true belief. The idea, then, would be for the veritist to reject an implicit premise of the knowledge objection—the assumption that if our epistemic purposes consist in part in those purposes we have insofar as we are after knowledge, then our epistemic purposes are not *just* those purposes we have insofar as we’re after the truth (i.e. maximizing true beliefs, avoiding false beliefs). It was assumed in the

⁶⁴ I want to draw here from a point Riggs (2004) makes when discussing the trait of *insight*. Riggs contends that while some things are epistemically good because of their connection to the epistemic goal of truth, insight—even if it has the effect of leading to truth—is epistemically *good* because of its connection to the distinct epistemic end of understanding. And so for Riggs, while both truth and understanding are epistemic goals, insight is defined *in terms* of the latter goal but not the former. This case is a specific example of why pluralism should not be thought of as wedded to the *unified aim* thesis, and perhaps also, why pluralism would be suspect if it *were*.

argument that these purposes would be distinct enough to not collapse together. Let's call the case for challenging this idea the *collapsing argument*.

Collapsing' Argument (against the knowledge objection)

1. If we aim at knowledge if and only if we aim at the truth (i.e. true belief), then some purpose P will be a purpose we have insofar as we're after knowledge if and only if P is a purpose we have insofar as we're after truth (i.e. true belief).
 2. If some purpose P will be a purpose we have insofar as we're after knowledge if and only if P is a purpose we have insofar as we're after truth (i.e. true belief), then our *epistemic purposes* are those we have insofar as we're after knowledge if and only if they are also those we have insofar as we're after truth (i.e. true belief).
 3. We aim at knowledge if and only if we aim at the truth (i.e. true belief).
-
4. Therefore, our epistemic purposes are those we have insofar as we're after knowledge if and only if they are also those we have insofar as we're after truth (i.e. true belief).

If the *collapsing argument* is sound, then Premises (1) and (2) in the knowledge objection fail to count against veritism because simply *having certain epistemic purposes in part because we're after knowledge* would be compatible with our purposes being epistemic only insofar as we're after the truth (the veritist's claim).

The key premise of the *collapsing argument* is (3), and (3) is admittedly not without some intuitive support: because knowledge is factive, you can't have as your aim knowledge without also having as your aim truth. And additionally, it seems strange to suppose one could aim to have the truth without aiming at knowledge. After all, if you want the truth about whether or not p , you'll want your belief about p to be as well-justified as it can be; and so you aim for a true belief when you aim for a true belief that is as well justified as it can be. This looks quite a bit like aiming for knowledge.

Despite the appeal of premise (3), the biconditional is actually far too strong, and it's not hard to see on closer inspection how the two aims come apart. First off, consider that even though true belief is part of what makes up knowledge, true belief does not entail knowledge: they come apart in cases of mere true beliefs. So knowledge and mere true beliefs are distinct. If (3) is correct, then you aim for knowledge only if you are also aiming at a mere true belief. But, to revisit Kvanvig's point:

When asked, "why, from a purely cognitive point of view, do you value knowledge?" we are inclined to answer...that we

want to be correct, but not merely by accident, as what happens when one has a merely true belief. (Kvanvig 2009*b*: 7)

If (3) were correct, then you can take on the aim of knowledge only if willing to admit that you also want to be correct, just by accident ‘as what happens when one has a merely true belief’ (2009*b*: 7). Drawing from the luck platitude here, you can’t both know and be correct by accident. And so (3) is true only if you aim at knowledge only when you also aim at something you have only if you *don’t* have knowledge. In this respect, (3) is not only false, but in some sense contradictory; consequently, the veritist could rebut the knowledge argument via the collapsing argument only by reasoning through (3), and we can see that this would be a mistake. Here again, the veritist has an initially attractive response to the knowledge argument that ultimately comes unhinged, and irreparably so.

I considered the knowledge objection here in the detail that I did because I wanted to make it clear that the most obvious response to veritism is not one to which veritist has any clear and ready defense. By tracing out the possible replies, we’ve found that veritism lacks not only an obvious defense, but indeed any good defense. It should not seem at this point that veritism is nearly as attractive a position as it appears after just a superficial consideration, or after pointing out the theoretical virtues (i.e. parsimony) that made it look so promising.

Moreover, the fact that the knowledge objection draws out problems with the veritist’s proposal should lead us to think that parallel-style objections to veritism that appeal to understanding as an epistemic end will also pose problems. And indeed, Riggs (2004), Brogaard (2005), Elgin (2007) and Kvanvig (2003) among others have offered cases for thinking that understanding is both distinct from truth as an epistemic aim and also one whose epistemic value is not reducible to the value of truth. I’ll not consider these objections here, but rather, I think it will be more useful to move on now to a different kind of worry, one that raises a different set of questions.

4.6 Veritism and Epistemic Responsibility

I want to take a stronger line against the veritist now; in this section, I shall show by example that veritism commits itself to an absurdity. The absurdity itself is designed to reveal something to us about how the epistemic point of view must be characterized. In that respect, the objective here will not simply be a negative verdict for veritism, but a positive endorsement of the view it will naturally give way to.

I want to begin now with a thought experiment. Imagine that two old men, Gray and White lie upon

their deathbeds, praying to God. Watching from above, God gets ready to evaluate Gray and White morally to determine whether they are fit for heaven. Quickly bored, God changes his mind and decides that he wants Gray and White to be evaluated from a purely epistemic, rather than a moral, point of view. He sends an angel, Gabriel, to visit Gray and White, with precisely these instructions. At the bedside, Gabriel is quickly impressed by Gray, who he notices has acquired twice as many truths as White; but upon closer investigation, Gabriel realizes that the majority of Gray's truths were ones he held on the basis of scant evidence, stubbornness and wishful thinking. Gabriel scrawls down his mixed assessment of Gray and next turns to White only to find quite the opposite case. White had, in his life, acquired just half the truths Gray had. 'Not good,' thinks Gabriel, who scratches the number of White's truths in his angel pad. But upon closer look, Gabriel is stunned: as it turns out, nearly all of White's beliefs were based on well-grounded reasons and on the basis of carefully weighed evidence. It was just a matter of bad fortune that so many of these beliefs turned out false. Gabriel scrawls this information in his angel pad. Seeing Gabriel to have finished, God asks: 'Who was the better epistemic agent? Gray or White?'

It is important that we resist the temptation to smuggle moral considerations into what is here a purely epistemic evaluation; all God cares about here is who the better epistemic agent is. What we're evaluating epistemically then is agency, not belief. Gabriel, suppose, needs a bit of help here; he takes his angel pad to the local veritist, shows the veritist the raw data for both Gray and White and asks who the better epistemic agent was.

Before Gabriel can finish showing the veritist all his notes and explain again the important fact that God wants to know which *agent* was epistemically better, the veritist interrupts: 'An agent,' says the veritist slowly 'is subject to the same criteria of epistemic evaluation as beliefs, belief-forming processes, and anything else. There is a simple formula that we veritists apply across the board: for any X, X is epistemically valuable to the extent that it maximizes truth and avoids error.' With this, the veritist quickly circles Gray's name and says, 'There's your winner. That agent is epistemically better than the other one.'

Now, I contend here that Gabriel asked the wrong person; the veritist's conclusion is not the right one, and I want to now go about saying why.

When we considered how it is that beliefs aim at truth, we saw that this aim should be understood as characterizing a constitutive fact about belief: part of what it is to be a belief is to be the sort of thing that is correct only when true. Agency, on the other hand, is not like belief in this respect: agency is not (epistemically) correct only when true, and that is so even though truth is something at which agents can aim. If an agent aims

at the truth, then, and tries to bring it about: what are the conditions under which her agency is correct? We might be tempted here to say that the epistemic norm that governs agency is truth nonetheless, even though agency itself isn't correct only when true. The idea might be that agency is epistemically correct insofar as it brings about the truth. This is the line the veritist would have us take. But it is misguided. When truth makes beliefs correct, it is because beliefs fit the world. When does truth make agency correct? When agency 'fits the world'? Neither the intentional aims of agents nor agency itself is something that aspires to fit the world. The direction of fit is the other way around. Perhaps then agency is epistemically correct when the world fits the agent's intentional aims? Even if one is guided by the goal of truth, truth cannot be an intentional aim. As Lynch (2008) says:

One does not simply will oneself to believe the truth. Rather, we pursue truth indirectly, by pursuing those beliefs backed by reasons and supported by the evidence. (Lynch 2008: 4)

It stands to reason then that agency is epistemically correct just when the world fits the agent's intentional aims of having beliefs backed by reasons, supported by evidence, etc. The world fits these aims, and thus agency will be correct, just when the agent succeeds in her aim of believing responsibly.

The picture this gives us then is one according to which beliefs aim at truth and agents aim at responsible belief. The fundamental aim of belief is truth, the fundamental aim of epistemic agency is responsible belief. On this proposal, we can make much better sense of the case of Gray and White than the veritist could; our view says that if it is *beliefs* that are being evaluated from the epistemic point of view, then Gray's are more epistemically valuable than White's. This seemed to be the right result, as Gray had far more true beliefs. If it's agency that is being evaluated, however, our view clearly rules that White, rather than Gray, is the better epistemic agent. The fundamental aim of epistemic agency is responsible belief, and White had far more responsible beliefs than Gray. The present view is, thus, a variety of *epistemic value pluralism*, although not in the standard sense: unlike pluralists such as Riggs, I am not allowing for manifold sources of epistemic value—instead, there will be just one epistemic value to govern belief, and another to govern agency.

4.7 Knowledge and the Epistemic Point of View

But before moving on, there is a final conclusion that we are in a position to draw—a conclusion about what sort of thing knowledge is. Here it will be helpful to retrace our steps: our overarching project has been to develop an anti-luck theory of knowledge out of our value platitude. Up to now, our aim has been a modest one—simply to find out what exactly the Meno Requirement demands of us, and that has been a task in itself.

What we found in Chapter 4 was that the Meno Requirement consists in two distinct criteria, one pertaining to knowledge and true belief qua types and the other pertaining to knowledge and true belief qua tokens. But because we saw that the Meno Requirement is one that we can't purport to address without a theory of epistemic axiology, we needed to specify one that would work. We concluded in this chapter that veritism is inadequate, and that we should understand the epistemic point of view as broadly one on which both the good by way of belief and by way of agency matter. We know then that a theory of knowledge that preserves the two criteria of the Meno Requirement is going to have to do so in light of the fact that both the good by way of belief and the good by way of agency are fundamental epistemic values. A precedent for this sort of thinking is found in Sosa's (2007) recent work on epistemic normativity, whereby Sosa supposes a response to the value problem for knowledge turns on the thought that knowledge is connected to distinct though equally fundamental epistemic values. Here's Sosa:

One part at least of the solution to the value problem lies in a point central to virtue epistemology: namely, that the value of apt belief is no less epistemically fundamental than that of true belief. For this imports a way in which epistemic virtues enter *constitutively* in the attainment of fundamental value, not just instrumentally. (Sosa 2007: 87)

Sosa here recognizes that the epistemic value of virtuously formed belief should play a role in accounting for the value of knowledge not *just* instrumentally as a means to truth, but as a fundamental epistemic value in its own right. Does this sort of dual account of fundamental epistemic value tell us anything about what a theory of knowledge will look like? I think it does. By learning all that is entailed in satisfying the Meno Requirement, and determining at least roughly what the value sources are (good belief and good agency) within which the Meno Requirement is framed, we're in a position to see what the Meno Requirement implies that knowledge *couldn't* be. For one thing, knowledge can't be:

K1: A true belief that is not held responsibly.

Because the epistemic aims of belief and agency are distinct, the epistemic value of K1 reduces to the epistemic value that beliefs have simply in virtue of being true. K1 derives no additional epistemic value in virtue of its connection to the value that governs agency. Therefore, we learn from the Meno Requirement that knowledge must *at least* be:

K2: A true belief that is held responsibly.

The Meno Requirement implies that knowledge must at least consist in K2; knowledge must be epistemically valuable not only insofar as true beliefs are valuable, and this will be so only if knowledge consists in a belief that is epistemically valuable in virtue of being true as well as epistemically valuable in virtue of its relation to some other source of epistemic value: on our view, that other source of epistemic value is the epistemic aim of agency, which is responsible or intellectually virtuously formed belief. So knowledge is more valuable than true belief only if knowledge consists in true belief held responsibly—that is—held as a result of correct or virtuous agency. The Meno Requirement is thus not only one that restricts what we can say about the value of knowledge, it also implies necessary conditions on the nature of knowledge.

4.8 Concluding Remarks

By beginning with platitudes about the value of knowledge, we now seem to have uncovered something positive about the nature of knowledge—that knowledge must be, loosely, something like *true belief that is connected to intellectually virtuous agency*, or as Sosa (2007) calls it, *apt belief* (2007: 92). This is of course only a template at present, but having derived a template for analyzing knowledge straight out of the Meno Requirement is significant: it is a step forward that places the remaining goals of our project more clearly in sight. Dialectically, we are left with two master objectives that will occupy the remainder of the wider project.

(1) Our working idea is that the *true belief connected to virtuous agency* template theory of knowledge is the sort of theory that would be *necessary* for satisfying the Meno Requirement. A first objective will be to put this idea on surer footing: it is needed that we demonstrate that such a theory would also be *sufficient* for satisfying the Meno Requirement, both its type and token criteria. That way, we can vouchsafe that whatever counts as knowledge on a theory developed out of this template will be better, epistemically, than true belief, and in just the sense the Meno Requirement demands. Once this sufficiency condition is argued for, the next task dialectically will be:

(2) To develop this template into a defensible account of knowledge. Because knowledge excludes luck, the core task in doing so will be to show how the connection between true belief and the cognitive ability it must arise from will suffice to block the relevant sorts of epistemic luck.

The consequence of satisfying these objectives, then, will be that we'll have shown how the connection between true belief and cognitive ability essential to knowing is one that both seals knowledge from undermining luck and accounts for the distinctive value we take knowledge to have. This will be, ultimately

then, a theory of knowledge that is not simply developed independently from the luck and value problems and only later tested against them; it will be a theory that builds a satisfaction of the conditions arising from these platitudes into an analysis of what knowledge is.

5

Evidentialism, Reliabilism and the Value of Knowledge

5.1 Introduction

In Chapter 4, we saw that the Meno Requirement must be divorced from a veritist axiology. The epistemic point of view, it was argued, is not one from which *just* true belief matters. What we saw was that what makes beliefs good and what makes agency good come apart in epistemic evaluation, and that this much counts in favour of a theory that understands knowledge as connected to both of these ends. Our reasoning led us to think, then, that in order to satisfy the Meno Requirement, knowledge must be conceived of as a kind of cognitive success through intellectually virtuous agency. At least, this was taken to be the most obvious candidate.

It would be too quick, though, to move optimistically forward before considering the prospects that some rival theories of knowledge have for addressing the value problem. After all, we should not stake too much weight in what was, in the previous chapter, only a strong presumption—and not a guarantee—that virtue-theoretic account is needed to satisfy the Meno Requirement. In the present chapter, then, my aim will be to see how well some competing views measure up to the task. The two theories I'll consider in detail are evidentialism and reliabilism—theories that capture instincts about knowledge that are strongly internalist and externalist, respectively.

5.2. Evidentialism

A prominent-yet-simple (at least on the surface) view in epistemology, and one with quite a bit of intuitive pull, is *evidentialism*. Evidentialism is generally presented as a theory of justification; according to evidentialism, whether S 's belief that p is justified is entirely a matter of S 's occurrent evidence for p . Evidentialism is also understood to be a theory about not only belief, but any doxastic attitude. Earl Conee and Richard Feldman (2004) capture evidentialism as follows:

Evidentialism: Doxastic attitude, d , toward p is justified for one at t if and only if one's evidence at t supports one's taking d towards p (Conee & Feldman: 2004: 15)

Evidentialism is an *internalist* theory of justification, which is to say that the theory is committed to saying that whatever serves to justify S 's belief that p must be reflectively accessible to S at the time she believes p . For the evidentialist, then, facts about the reliability of the processes that bring about p won't necessarily be part of what it is that makes p justified. What matters is just the evidence S has for p at the time she believes it.

According to evidentialism, one *knows* just in case her evidence fits her belief to a satisfactorily high degree, and her belief is true. Accordingly, evidentialist knowledge can be thought of as true belief justified to a high degree rather than as justified true belief that meets some independently specified condition over and above evidentialist justification.⁶⁵

5.3. Mentalist Internalism

Evidentialism entails *internalism*, and so if internalism is false, then so is evidentialism. As Duncan Pritchard (*forthcoming* b) points out, there are a variety of nuanced positions opponents take themselves to be arguing against when they purport to be arguing against internalism. The different nuances of internalism come to surface on the matter of what would be a sufficient condition for two agents to possess the same degree of justification for some belief. One variety of internalism, *accessibilism*, takes it that two agents have the same degree of justification so long as the reflectively accessible facts each is able to take as evidence for a belief are

⁶⁵ Trent Dougherty (2007) has recently argued that evidentialism, as a theory of knowledge, is perfectly compatible with our template theory of knowledge as *true-belief-through-ability*. He takes a rather strong view here and claims the two positions to be extensionally equivalent. If Dougherty is right about this, then the *true-belief-through-ability* template will satisfy the Meno Requirement if and only if an evidentialist theory of knowledge does also. Dougherty's extensional equivalence claim is motivated in part by his thought that one cannot inquire virtuously without proportioning her beliefs to the evidence, and that proportioning one's belief to the evidence entails that one has believed virtuously. This idea might strike us as compelling. I argue that it is premature, however, to draw the conclusion Dougherty does.

the same. Pritchard characterizes a slightly different variety of internalism as trading not on sameness of reflectively accessible facts, but on sameness of experiences, a view he calls *discernibilism*.

The sort of internalism Conee, Feldman and Dougherty understand evidentialism to be taking on board is one that Duncan Pritchard (*forthcoming b*) calls *mentalism*.

(MENT) If *S* and *S** do not differ in their mental states then they will not differ in the degree of epistemic justification that they have for their beliefs. (Pritchard 2008c: 4)

Unlike evidentialism, Mentalism is not itself a theory of justification. It is instead an articulation of what it is that a theory of justification will be committed to if that theory allows—as the standard line of evidentialism does—mental states to exhaustively justify beliefs.

If *mentalist evidentialism* is correct, then because evidentialist knowledge is understood as a true belief held with a sufficiently high degree of evidentialist justification, the following will also be true:

Ment-Knowledge I (MKI): If *S* and *S** do not differ in their mental states, then *S* has knowledge if and only if *S** has knowledge.

From this, it follows naturally that:

Ment-Knowledge II (MKII): If *S* and *S** do not differ in their mental states, then if *S* and *S** know *p*, *S*'s knowledge that *p* is more epistemically valuable than were *S* to have a mere true belief that *p* *only if* *S**'s knowledge that *p* is more epistemically valuable than were *S** to have a mere true belief that *p*.

MKII is a consequence of the prominent sort of evidentialist position that Conee, Feldman and Dougherty endorse, and I shall argue now why it is that MKII is false.

Consider on this note Donald Davidson's (1984) 'Swampman' example, which Sosa has popularized in epistemology. On the Swampman example, we imagine that a man (Bill) is walking near a swamp, and a bolt of lightning strikes in a very peculiar way: the bolt of lightning causes molecules by the swamp to form together to produce an identical replica of Bill—Swampman—who, when formed, is identical to Bill not only physically, but also mentally.

Suppose that, immediately after the creation of Swampman, Bill and Swampman form the same true

belief p about the swamp. If MKII is correct, then if Bill knows p , then Bill's knowledge that p would be better, from an epistemic point of view, than were he to have merely truly believed p only if what Swampman has would *also* be better, from an epistemic point of view, than were he to have a mere true belief.

Because the epistemic point of view is one from which correctness of belief and correctness of agency matter, MKII will be correct only if Bill's knowledge that p is precisely as good as what Swampman has, from the epistemic point of view. It's a mistake, though, to think that what Swampman has is any better epistemically than a mere true belief. What Swampman has is good by way of correctness of belief—he's got the truth—but not as good as what Bill has by way of correctness of agency, *even though* Bill and Swampman have, *ex hypothesi* formed the belief in the same way. This is because *even if* Bill formed the true belief through virtuous agency, Swampman could not have.⁶⁶

Given that we should resist MKII, it should be obvious that *accessibilism* and *discernibilism* would fare no better. Bill and Swampman are not only in the same mental states, they are *ex hypothesi* also going to have the same available facts and experiences to use as evidence.

We are in a position to reject two things here: (i) Trent Dougherty's recent suggestion that the true-belief-through ability template is extensionally equivalent to evidentialism (supplemented by either of the three considered varieties of internalism)⁶⁷, and (ii) that evidentialism could explain why knowledge is more valuable, from an epistemic point of view, than mere true belief, when it is. We must reject both (i) and (ii) even if Dougherty is right about the intimate connection between the exercise of intellectual virtue and the evidentialist requirement that we believe in proportion to the evidence.

5.4 Reliabilism

While internalist approaches such as evidentialism will have to account no less than any other approach for the value of knowledge, externalist accounts have been placed much more frequently on the hot-seat. In fact, it has become commonplace to explain *the very idea* that a theory must preserve that knowledge is epistemically better than mere true belief by using reliabilism as an example of how this idea is *not* preserved. As a consequence,

⁶⁶ Epistemic agency, as John Greco (2009a) points out, is something we exhibit by exhibiting those traits that make up our cognitive character, traits that must be (unlike Plantinga's "brain legion" or perhaps a fleeting True-Temp moment) *stable* features of our cognitive makeup. Even if Bill and Swampman believed p truly by going through all the same cognitive motions, Bill's forming of his belief that p is something good by way of belief *and* agency, while Swampman's forming of his belief that p will not be. Whatever good things we want to say about the way Swampman came to form his belief, we can't say that it's something good from the point of view in which good agency matters. What led to Swampman's true belief was a bolt of lightning just as much as it was anything we could rightly call agency.

⁶⁷ Dougherty argued for this point in a paper given at the 2008 *Epistemic Agency* Conference at the University of Geneva.

reliabilism has been targeted on this score not only by the arch-progenitors of the critique, Kvanvig (2003) and Zagzebski (2002), but by most all writers who have addressed the Swamping Problem.

Unsurprisingly, defenders of reliabilism have come to its defence. The most thorough recent defence of reliabilism against this objection has been given by Olsson (2007). I'll consider now the viability of Olsson's attempt to rescue reliabilism—and I shall in doing so be adhering to a requirement established in Chapter 3, which implies for us that Olsson's defence must not be combated in a way that relies on the precarious assumptions behind the traditional construal of the Swamping Problem—assumptions that are most always employed by those opponents Olsson takes himself to be defending against.

5.4.1 Olsson, Reliability and Stability

Olsson (2007) proposes that Swamping critiques against reliabilism are ill-founded—not because the underlying premises behind the critique are misguided—but because reliabilist knowledge is more valuable than critics of the view had thought.

According to Olsson, 'true beliefs that are reliably acquired are more stable and therefore more valuable for the purposes of guiding practical action over time' (Olsson 2007: 1). This is his central idea. We can break this idea into two parts: the first part of his claim is a claim about stability: that true beliefs reliably acquired are more stable than otherwise. This idea he calls the Reliability Stability Thesis (RST):

Reliability Stability Thesis (RST): $P(\text{S's belief that } p \text{ will stay put} \mid \text{S believes truly that } p \text{ due to a reliable process}) > P(\text{S's belief that } p \text{ will stay put} \mid \text{S believes truly that } p \text{ due to an unreliable process})$. (Olsson 2007)

The second part of his claim is a claim that takes the stability of reliably formed beliefs to entail that such beliefs will be more valuable than otherwise for the purpose of guiding practical action over time. He calls the second part of the claim the Stability Action Thesis (SAT), a thesis that takes the Larissa case to be illustrative:

Stability Action Thesis: $P(\text{S will get to Larissa} \mid \text{S's true belief as to where Larissa is stays put}) > P(\text{S will get to Larissa} \mid \text{S's belief as to where Larissa is will be lost})$. (Olsson 2007)

RST and SAT are taken by Olsson to imply what he calls the Reliability Action Thesis (RAT):

Reliability Action Thesis (RAT): $P(S \text{ will get to Larissa} \mid S \text{ has a reliably acquired belief as to where Larissa is}) > P(S \text{ will get to Larissa} \mid S \text{ has a mere true belief as to where Larissa is})$. (Olsson 2007)

Olsson takes it that RAT constitutes an explanation for why reliably formed true belief is better than mere true belief: it's better because reliably formed belief is a better guide to action. If Olsson's right, then Meno was wrong to think that knowledge and mere true belief will get you to Larissa just the same; the probability you'll get there if you have reliabilist knowledge is greater, according to Olsson, than is the probability that you'll get there if you have a mere true belief—one that is not reliably formed.

One might be tempted to object at the outset that Olsson has given a defense of the value of reliabilist knowledge orthogonal to the actual critique—which was a critique along the lines of value from an epistemic, and not from a practical, point of view. This observation is not without cause. At present, though, it will be good to investigate whether the sort of argument he gives might go through otherwise, and whether a case could be made from what he's claimed that could vindicate reliability as making a belief more epistemically valuable as well.

On that note, let's consider Olsson's first step of the argument (RST), which says:

RST: $P(S\text{'s belief that } p \text{ will stay put} \mid S \text{ believes truly that } p \text{ due to a reliable process}) > P(S\text{'s belief that } p \text{ will stay put} \mid S \text{ believes truly that } p \text{ due to an unreliable process})$.

RST is very much akin to a view that Timothy Williamson (2000) endorses, which is that knowledge is less susceptible to rational undermining by future evidence than is mere true belief. (2000: 77-79) Williamson thus takes it that the value knowledge has over and above mere true belief rests in its comparative stability, the same idea captured by Olsson in RST.

Kvanvig has objected on several grounds to this sort of proposal, and as Olsson admits, the points Kvanvig (2003) has raised for Williamson will apply as well to RST. The first point Kvanvig makes is that knowledge, no less than mere true belief, can be lost. This observation is closely related to his second point, which is that RST would be 'undermined if the true beliefs [not reliably formed] are thoroughly dogmatic ones' (Olsson 2007: 14) Both the first and second point Kvanvig makes hit at something true. The first claim, in particular, is straightforwardly true: knowledge, not unlike true belief, can be lost. But as Olsson points out, this doesn't imply that knowledge is not *less likely* to be lost than mere true belief—and so the first point is

compatible with RST.

Kvanvig's second point captures the idea that was discussed in Chapter 2 in the example of Brother Jed: one can dogmatically hold an unreliably formed true belief, and hold it just as steadfastly as it would be held if known. Olsson grants this point as well but argues that *even if* dogmatic beliefs can be as stable as knowledge, this would not change the fact that the probability your belief will stay put given that it was reliably produced is higher than the probability it would stay put if unreliably produced—and that's so even though some unreliably produced beliefs will be held dogmatically and in a way that leads them to be as stable as they would be if known.

Kvanvig offers a third point, however, which unlike the first two, leaves the defender of RST without any quick reply. His third point is that it isn't the property of being reliably produced that itself makes a true belief more stable; rather, it's the property of being held with *conviction* that makes a belief likely to stay put. As such, the stability that attaches to reliability is 'highly contingent.' (Kvanvig 2003: 13)

The point being made here is that 'being held with conviction' would be a merely accidental property rather than an essential property of beliefs reliably formed. But crucially, because it is the conviction the belief is held with and not the reliability *per se* that is the actual *stability-maker* for true beliefs, we can replace 'reliably produced' with 'held with conviction' in RST, a move we're entitled to make and which would have the effect of leaving the concept of reliability otiose for the purpose of explaining the value of knowledge.

Suppose though for the sake of argument that RST were correct. Even then, Olsson's argument is problematic for another reason, which is that RAT (RST in conjunction with the Stability Action Thesis (SAT)) betrays a certain 'jumping ship' and abandoning of the sort of externalist position that reliabilism should be taken as endorsing. SAT, recall, connects stability to action by claiming that: $P(S \text{ will get to Larissa} \mid S\text{'s true belief as to where Larissa is stays put}) > P(S \text{ will get to Larissa} \mid S\text{'s belief as to where Larissa is will be lost})$. And from RST and SAT, Olsson gets:

Reliability Action Thesis (RAT): $P(S \text{ will get to Larissa} \mid S \text{ has a reliably acquired belief as to where Larissa is}) > P(S \text{ will get to Larissa} \mid S \text{ has a mere true belief as to where Larissa is})$.

RAT falls prey to a very serious sort of objection. The probability that *S* will get to Larissa given that *S* has a reliably formed belief about where Larissa is will be greater than the probability *S* will get to Larissa if her true belief is unreliably produced only because Olsson takes it that reliably formed beliefs are more stable than

unreliably formed beliefs. We saw though that reliably formed beliefs wouldn't *be* more stable than unreliably formed beliefs if it weren't for the fact that reliably held beliefs are generally held with more conviction. We can grant Olsson that reliably formed beliefs are usually held with more conviction than unreliably formed beliefs while emphasizing that it is the *conviction*, specifically, and not the reliability itself that makes a belief stable.

Now if S and S^* have all the exact same beliefs, then it will follow that S could not hold some belief p with more conviction than S^* does. S could hold p with more conviction than S^* only if S had some additional belief, which S^* doesn't have, in virtue of which S is more convinced that p is true than S^* is. Conviction that p is thus the sort of thing that carries with it some second-order belief *about* p . Thus:

Conviction: For S , Φ, φ , S will hold Φ with more conviction than φ only if S believes Φ in conjunction with some further belief Ψ , where Ψ is the belief that Φ is more probable given S 's evidence than φ .

Because reliabilism is an externalist theory, it is going to imply that the difference between a knower and a mere true believer is explained by a difference in the reliability of the respective belief forming processes and not a difference in respective beliefs. Thus, reliabilism implies that a knower and a mere true believer cannot be distinguished on the basis of their respective conviction. And so the consequence for Olsson is not good: his argument implies that what makes knowledge more valuable than mere true belief is something entirely orthogonal to what it is that reliabilism says makes knowledge *different* from mere true belief.

Although Olsson does not appreciate the extent of this problem, he does seem to be aware that reliable true belief, in order to be more stable than unreliable true belief, must be supplemented with some, as he puts it, 'modest internalist requirement' which he takes to be the requirement of 'track-keeping' (Olsson 2007: 18-19). Here's Olsson:

The thesis presupposes... the holding of some identifiable empirical conditions. One of these conditions is track-keeping, stating that the person maintains a record of how a given belief was arrived at, i.e., of the type of belief-acquisition process that terminated in the belief in question. Only then can the subsequent discovery of the unreliability of a given fixation method lead to the discrediting of other beliefs previously fixed using that same method or process. Without track-keeping this is hardly possible (Olsson 2007: 18-19).

But whatever the addition of this modest internalist requirement helps to save, it's not reliabilism, at least where I come from. Reliabilism is supposed to be an externalist theory, through and through. That's precisely why reliabilism has the resources it does for dealing with the sort of radical skeptical challenges that are thought to face internalist views. If reliabilist knowledge gets its value then because of an internalist condition rather than

an externalist condition, reliabilism would be inadvertently chucking the baby out with the bathwater.

5.4.2 Goldman & Olsson on the value of reliability: two further defences

As the critique of Olsson's stability argument has shown, the reliabilist must vindicate the value of reliably formed true beliefs over unreliably formed beliefs by locating the value-making property somewhere, loosely speaking, in the *reliability* rather than in somebody's head.

Goldman & Olsson (2009) have, in a recent and interesting paper, offered up two ways to do this. Let's look now at the first.

5.4.2 (a) Reliability, Type- and Token-Instrumental Value

Goldman & Olsson (2009) highlight what they take to be a key piece of reasoning that lies behind the standard Swamping Critique of reliabilism offered by Kvanvig and Zagzebski. They call this the *simple derivation thesis*:

(Simple Derivation Thesis) If the value of one state is entirely derivative from the value of a second state, the value of the composite state consisting of the two components does not exceed the value of the second component. (Goldman & Olsson 2009: 12)

Once it is thought that the value reliability has is entirely derivative from the value of truth, then by reference to the Simple Derivation Thesis, the value of reliably formed true belief does not exceed the value of true belief. Goldman & Olsson take this, and I think rightly so, to capture the basic structure of what I challenged under the name of the Swamping Thesis in Chapter 3.

Goldman & Olsson argue that the Simple Derivation Thesis helps to get the result Kvanvig and Zagzebski want only if a certain assumption about instrumental value is in place. They call this *token instrumentalism*:

(Token Instrumentalism) One token event (or state of affairs) inherits value from any other token event (or state of affairs) that it causes if and only if the resulting token event has fundamental value (Goldman & Olsson 2009: 13)

On this view, a true belief inherits value from some reliable source just in case the reliable source is fundamentally (epistemically) valuable and serves as the cause of the belief.

Goldman & Olsson take it that although *token instrumentalism* represents the way most epistemologists

conceive of the sort of instrumental value that would be at issue between a belief and a reliable process, there is another viable option on the table. They call this *type instrumentalism*.

(Type instrumentalism) A type or class of events C inherits value from another type or class of events E if instances of C tend to cause instances of E and instances of E (invariably or usually) have fundamental value. Derivatively, each instance or token of such a type C has instrumental value by virtue of belonging to such a type. (Goldman & Olsson 2009: 14).

Whereas the derivation thesis will, on token instrumentalism, render reliably formed true belief no more valuable than true belief, it will not have the same effect, they argue, if the instrumental value relation between the belief and source is understood along the lines of *type instrumentalism*. They reason as follows:

Consider a type of belief-forming process that generally produces true beliefs, say, the reasoning process of modus ponens... If modus ponens reasoning as a type is generally reliable, this means that tokens of that type usually produce true belief results... Since true belief is assumed to have fundamental epistemic value, the modus ponens type of reasoning should be deemed to have instrumental epistemic value. Moreover, any particular *token* of this type also has instrumental epistemic value, under the type-instrumentalist conception. Notice, however, that this doesn't depend on each particular token actually *causing* a true belief (Goldman & Olsson 2009: 14, *my italics*).

The idea here is that my reliably produced true belief will be (epistemically) more valuable than an unreliably produced true belief not because a reliable process caused my true belief, but because the process I employed is instrumentally epistemically valuable (as a type of process that usually produces true beliefs) and any token of this type of process—that is, whatever belief is produced by it—also has instrumental epistemic value.

Though the appeal to type instrumentalism rather than token instrumentalism seems promising, there is a problem that prevents the initial promise from panning out. Consider that Goldman & Olsson would have us think that the value of the reliably produced true belief is a value it has not because a reliable belief forming process causes the belief, but because the belief constitutes a token of a type of process that is reliable (and instrumentally valuable). But wait a minute—the belief would constitute a token of *that* particular type of process rather than some other type of process only if *that* process is what causes the belief rather than some other type of process that is not instrumentally epistemically valuable. And so unless a token belief can be caused by some process other than, and not by, the type of process of which it is said to be a token, Goldman & Olsson should maintain that value a reliably produced true belief has over and above an unreliably produced true belief is a value it has at least in part because type of process that causes the belief is reliable. But that's

precisely the conclusion they employed the machinery of type-instrumentalism to avoid. So this response doesn't work.

5.4.2 (b) The conditional probability solution and the value of *future* true beliefs

Goldman & Olsson offer an alternative way to defend the value of reliabilist knowledge over mere true belief; this second defense is a bit simpler than the one what we've just considered. They call it the *conditional probability solution*. They express the idea as follows:

What is this extra valuable property that distinguishes knowledge from true belief? It is the property of *making it likely* that one's future beliefs of a similar kind will also be true. More precisely, under reliabilism, the probability of having more true belief (of a similar kind) in the future is greater conditional on S's *knowing* that p than conditional on S's *merely truly believing* that p (Goldman & Olsson 2009: 16).

The key idea of this argument, a variation of which has also been advanced by Kristoffer Ahlstrom (2006), is that the epistemic value that a belief has in virtue of being reliably produced is a value it has not just because it makes the true belief it causes likely to be *what it already is* (true), but because it causes more true beliefs to be likely. The 'source' of the value reliability adds to a true belief is thus *future true beliefs* and not any actual true belief.

This is not only more promising initially than Goldman & Olsson's first attempt, but also more promising as a way to defend the *epistemic* value of reliabilist knowledge than an attempt like what we saw in Olsson's stability argument, which was an argument that appealed straightforwardly to pragmatic value.

From an epistemic point of view, it's hard to deny that a lot of true beliefs—even if only 'earmarked' by our reliably formed true belief—are more valuable than a belief that doesn't carry with it such an earmark. But what must specifically be more epistemically valuable than a mere true belief is not a bunch of true beliefs, but a single belief that supposedly carries with it a claim to the epistemically valuable true beliefs that according to Goldman & Olsson a reliably formed true belief makes it probable that we'll have. Even understood this way, it's somewhat plausible to think Goldman & Olsson have a fair case to make.⁶⁸

But this plausibility wanes once the implications of the proposal are drawn out. One implication would be that reliabilism is, for lack of a better term, *agist*—locating the value of knowledge over true belief entirely in

⁶⁸ After all, from the point of view in which money matters, a single ticket worth a dollar is not as valuable as another ticket that is not only worth a dollar but also likely to bring you five more. Presumably, the same would hold true when comparing epistemic value of two true beliefs, each epistemically valuable insofar as they're true, but one more valuable than the other by making it *more probable* than the other that we'll get more epistemically valuable true beliefs in the future.

what it promises in the future implies unambiguously that what you know when you're old is less valuable epistemically than what you know when you're young. While it's true that *pragmatically* speaking, knowledge isn't especially valuable for older people because they have less opportunity to *use* it, the same argument does not hold for epistemic value. From the point of view in which right belief and right agency matter, a well-formed true belief from the deathbed can be just as good as a one similarly formed from the crib. While such considerations might sound either non-philosophical or sophomoric, the point stands that the view Goldman & Olsson give us has no way of explaining why, for example, (apologies for the appeal to drama) a dying mother's belief on her death bed that her son fighting in the war is alive is something that, from the epistemic point of view, and regardless of what it can get her pragmatically (i.e. in this case, nothing, as she's minutes from kicking the bucket), better to have than a mere true belief that her son is still alive, one that—before she dies—she believes (truly) while only hoping it to be true.⁶⁹

5.5 Concluding Remarks

At the end of the day, neither the internalist nor the externalist contender had at its disposal a viable strategy to account for the distinctive value of knowledge. It was instructive to consider just where each ran into its problems and what implications this might have for the view we've been developing. In the next chapter, we'll see that underlying the surface problems faced by reliabilism and evidentialism considered in this chapter is a shared structure that contributed to stacking the decks against their Meno-resolving prospects—and we'll see why a virtue-theoretic approach stands to be promising in the absence of such limitations.

⁶⁹ As sure as I am that this reply to Goldman & Olsson lacks much philosophical sophistication, I'm also sure that what they've claimed about the value of reliably formed true belief invites it. And while the notion of ageism seems absurd, it is no more absurd that the thought that, for an old man, true belief is just as good as knowledge. There is an additional problem, though, with Goldman & Olsson's second solution—a problem much like the one that was raised to Olsson's stability response. The problem is that a reliably formed true belief would only *be* indicative of future true beliefs against a background assumption that the agent has the ability to recognize that the process employed in the given case was successful enough to merit using again in the future. Otherwise, there's no way to explain why a reliably formed true belief actually is one that would make probable future true beliefs moreso than an unreliably formed true belief would. But once this sort of 'tracking' requirement becomes an essential part of an explanation for why reliabilist knowledge is better than mere true belief, the value of reliabilist knowledge can only be vindicated by appealing to the value of something not made valuable by the belief's being formed reliably, but rather by the satisfaction of an internalist condition; once this much is ceded, we're no longer giving an explanation for why meeting the *reliabilist's conditions for knowledge* would give us anything epistemically better than what we have when we have only a mere true belief.

6

Knowledge, Ability and Final Value

6.1 Introduction

My intention in the previous chapter was not to suggest that the true-belief-through-ability template should be thought to constitute a viable way of accommodating the Meno Requirement simply by process-of-elimination. Our detour into evidentialism and reliabilism was instead with the more direct purpose of making clear the sorts of problems our own view must avoid. Despite their obvious differences, one thing evidentialism and reliabilism have in common is a shared commitment to analyzing knowledge straightforwardly in terms of properties of beliefs—a commitment that carries with it the following consequences:

(1) The value of knowledge must be explained in terms of the value of certain properties of beliefs. (e.g. Reliabilism and Evidentialism)

And secondly,

(2) What makes an agent good (i.e. intellectually virtuous) must be explained in terms of what makes a *belief* good to have, and not the other way around.

(1) and (2) together imply that:

(3) What makes knowledge better to have than true belief⁷⁰ cannot be explained in terms of valuable properties of the knowing *agents* themselves (otherwise, the account will be circular).

⁷⁰ In the sense that the Meno Requirement demands that a theory preserve. Note that in some references that I leave without this qualification, my aim will be brevity only. It is of course only that sense that we're interested in.

This sort of conceptual commitment⁷¹, according to which good-making properties of beliefs will be logically prior to good-making properties of agents, is directly analogous to what Bernard Williams (1985, 1995) has challenged under the name of *centralism* in ethics. What Williams specifically rejected was that, within a theory, conceptually ‘thin’ concepts such as ‘right’ and ‘justified’ should be logically prior to ‘thick’ concepts such as those that pick out virtues of the agents themselves. (Williams 1995: 184) *Anti-centralist* views in ethics simply deny that thin concepts hold a right to any sort of logical priority over thick concepts. Thus, on anti-centralist approaches, ‘right’ or ‘justified’ might be defined themselves in terms of conceptually thick concepts such as virtues.

The epistemic analogue of this sort of *anti-centralist* approach in ethics will be one that allows intellectual virtues to function more centrally within a theory of knowledge; intellectual virtues will be the sorts of things we can account for without having to appeal to some antecedent concept of justified belief. By rejecting the logical priority of the ‘thin’ over the ‘thick’, an epistemological theory will be within its rights also to explain what makes knowledge *valuable* by appealing directly to considerations of what make good *agency* valuable, rather than exclusively to what it is that makes justified *beliefs* valuable. (Axtell & Carter 2008: § 2)

Because we’ve said that the epistemic point of view is one from which *both* the good by way of beliefs *and* by way of intellectual agency matter, it’s going to be the case that a theory that can explain how knowledge derives its value from the goodness of both the belief *and* the agency giving rise to it has a clear advantage over a theory that can’t. This idea was merely suggested at the outset, but now stands on a stronger footing. For it should be more obvious now that a theory that rejects *veritism* in favor of our sort of belief-agent *pluralism* will be capable of accounting for the value of knowledge only if it in turn rejects *centralism* in favor of *anti-centralism*.⁷² Put more precisely, when the problem presented is that of accounting for the epistemic value knowledge has

⁷¹ To make these consequences clearer, just consider that for the evidentialist, what would make an agent intellectually virtuous is that she believe in accordance with the evidence. And so knowledge—true belief that fits the evidence—can’t then be explained to be more valuable than true belief in a way that appeals, in the explanation, to the value a true belief would have by arising from the sort of agency *the evidentialist would count as virtuous*; that is, what makes a true belief more valuable than otherwise because it fits the evidence couldn’t be explained by appealing to the value of some trait *that would itself be valuable for a belief to arise from* only insofar as believing in accordance with the evidence is valuable. Similarly for reliabilism: what makes a true belief more valuable than otherwise because it is reliably produced can’t be explained by appealing to the value of some trait that would itself be valuable for a belief to arise from only insofar as reliably formed beliefs are valuable. Evidentialism and reliabilism, thus, not only make the concept of “justified belief” logically prior to the concept of “virtuous agency” but also, in doing so, block out the possibility of explaining why knowledge is better epistemically than mere true belief by appealing to facts about what make the sort of *agency* constitutive of knowers good.

⁷² And a *pluralist, anti-centralist* view, it might be added, looks pretty good against the competition considering that targets of the Swamping Problem have almost exhaustively been theories that analyze knowledge in terms of properties of beliefs (*centralism*) and then try to explain why true belief isn’t as valuable as knowledge by arguing the only way they can—by trying to show that the value these properties derive from the value of truth (*veritism*) is enough to add value to a true belief.

over true belief—what the Meno Requirement demands we account for—centralism limits what sorts of valuable properties a theory can appeal to, and veritism limits what the sources are that could serve to make those properties valuable.

These considerations make the *pluralist, anti-centralist* view on which we conceive of knowledge as *true-belief-through-ability* quite attractive, although we shall see that there are several tricky issues that must be addressed in order to actually show how true beliefs arising through ability have the value we want them to in light of the particular variety of belief/agent pluralism I've taken on board. The motivation is thus well in place. We just need to work out the details.

6.2 Epistemic Value Pluralism: Belief and Agency

Let's not forget that even if knowledge requires that a true belief be through ability, it doesn't change the fact that knowledge picks out a belief with certain properties, and not an intellectual trait or ability with certain properties.⁷³

Consequently, the true-belief-through-ability template could satisfy the Meno Requirement only if it can explain how true beliefs *themselves* would be more epistemically valuable if they arise from intellectually virtuous agency, than they would be otherwise. And it will not suffice for us simply to point out that there is something other than truth that explains why the virtuous agency that gives rise to known true beliefs is epistemically valuable. We still would need to explain why the true beliefs themselves gain some additional value for having been formed in this way.

Zagzebski's machine-product analogy lurks: it seems like, for the purpose of achieving the aim consisting in correct belief, the property of having arisen from virtuous agency is a valuable property for a *belief* to have insofar as it makes the belief likely to be true. How then, can we explain why a true belief gains an additional 'boost' of value in virtue of having arisen from intellectual virtue?

It will be tempting to try to dodge the machine-product objection by pointing out that, unlike the reliabilist, who is committed to veritism, we don't have to concede that epistemically virtuous agency is valuable only insofar as truth is valuable. That is indeed true. *But nonetheless*, the machine-product argument that

⁷³ And so, even if intellectually virtuous agency is valuable in its own right, a true belief that arises out of intellectual virtuous agency will be more valuable than otherwise only if it can be explained why true beliefs *themselves* are more epistemically valuable when produced this way than they would be otherwise. Our *true-belief-through-ability* template therefore has some explaining left to do. Even though our pluralistic picture of the epistemic point of view recognizes distinct aims of correct belief and correct agency, the specific implication for the *true-belief-through-ability* template, as a template for a theory of knowledge, will be that (i) the agency required for knowledge is valuable insofar as good agency is an epistemic end; and (ii) knowledge is a valuable sort of belief insofar as true belief is an epistemic end.

undergirds Zagzebski's (2002) espresso analogy doesn't immediately go away as we might expect. This is because the sort of epistemic value pluralism I'm endorsing *actually does accept* veritism (albeit, only) as a measure of correct *belief*. The view is a pluralist one because it maintains that, over and above truth, there is a distinct epistemic end of virtuous agency. Virtuous agency, on my view, is *not* epistemically valuable only insofar as truth is valuable, even though truth is the measure for what makes *beliefs* (as opposed to agents) epistemically valuable.

This picture might strike as somewhat complicated, and so I want to make explicit and clear an important implication of the sort of pluralism I've endorsed. It is this implication that should bring to light why it is that, *even though* intellectually virtuous agency isn't valuable only insofar as the truth is valuable, Zagzebski's machine-product argument doesn't automatically go away. Call this implication *veritistic value pluralism*:

Veritistic value pluralism. Even though intellectually virtuous agency itself is not epistemically valuable only insofar as truth is valuable, the property of 'having arisen out of intellectual virtue' is a valuable property *for a belief* to have only insofar as it makes *the belief* likely to be true. (And similarly, 'having true beliefs' is something epistemically valuable *for an agent* only insofar as it makes it likely that the *agent* is *intellectually virtuous*.)

The *veritistic value pluralism* I'm advancing implies the following: if knowledge is valuable only insofar as correct (true) belief is valuable, then no matter how epistemically valuable the virtuous agency is that knowledge requires a true belief arise from, *knowledge* (which is a variety of belief, not agency) will fail to be something our view can vindicate as more epistemically valuable than mere true belief. What our *true-belief-through-ability* template had better explain then is:

Burden of explanation: It is necessary that we explain how *knowledge*—a kind of successful belief—is valuable insofar as:

- (i) Correct belief is valuable AND
- (ii) Correct agency is valuable

Further, to avoid the machine-product objection:

- (iii) What makes knowledge valuable insofar as correct agency is valuable must *not be explained* by appealing to the idea that correct agency makes correct belief likely (even if it does).

Given these conditions, it will be useful to consider whether appealing, as Goldman & Olsson did, to *type instrumentalism* (as opposed to token instrumentalism) will be of any use. Recall that type instrumentalism says:

(Type instrumentalism) A type or class of events C inherits value from another type or class of events E if instances of C tend to cause instances of E and instances of E (invariably or usually) have fundamental value. Derivatively, each instance or token of such a type C has instrumental value by virtue of belonging to such a type. (Goldman & Olsson 2009: 14)

Because we are pluralists, there are two ways we can read this. On one reading, we could say that the type of events consisting in intellectually virtuous belief formation tends to cause instances of true beliefs, which have fundamental epistemic value (in that true belief is the fundamental epistemic aim of beliefs), and that therefore, token beliefs of the class of virtuously formed beliefs have instrumental epistemic value insofar as they belong to the class of virtuously formed beliefs. Problematically, this first route fails to meet requirement (iii) of our explanatory burden: it is a strategy that explains the value of being virtuously formed in terms of the value of true beliefs, and thus, is open to the machine-product objection.

Our pluralist view allows for a second reading: we could say that the type of events consisting in intellectually virtuous belief formation tends (albeit trivially) to cause instances of *intellectually virtuous agency*, which have on our view fundamental epistemic value, and therefore, token beliefs of the class of virtuously formed beliefs have instrumental epistemic value insofar as they belong to the class of virtuously formed beliefs.

Even though, as was mentioned in the critique offered against Goldman & Olsson, a belief counts as a token of the type consisting in virtuously produced beliefs only because virtuous agency (rather than some other type of agency) ‘caused’ the belief, this isn’t any obvious problem for our view. That is, it’s no problem so long as a true belief caused by intellectually virtuous agency *actually would*, as Goldman & Olsson say, ‘derivatively’ acquire additional epistemic value in virtue of being ‘instances’ of that type of belief. (2009: 12-14) The question then becomes: how can a true belief be a more valuable sort of *belief* to have, epistemically, just because its source, virtuous agency, is itself fundamentally epistemically valuable? More generally, how can a true belief be more valuable than otherwise because it stands in a certain relation—the relation of being produced by—something valuable?

6.3 Rabinowicz and Rønnow-Rasmussen on Final Value

In answering this question, it becomes relevant to consider the distinction between relational and non-relational properties. This is so given that we’re interested in how true belief might derive value from standing in a certain

relation to intellectually virtuous agency. According to Wlodek Rabinowicz and Toni Ronnow-Rasmussen (1999), a relational property can be either internally relational or externally relational.

Internally relational property: A property that an object possesses in virtue of the relations it has to its own *parts*. (components, elements, constituents, etc.). (Rabinowicz & Ronnow-Rasmussen 1999: 2)

Externally Relational Property: a property that an object has in virtue of its relation to something that is not one of its parts. (1999: 2)

Because intellectually virtuous agency doesn't, by giving rise to a belief, thereby become an element or a constituent of the belief *itself*, intellectually virtuous agency stands to be an external relational property (hereafter, a relational property) of a true belief so produced. Hence, 'having arisen through virtuous agency' is a property a true belief would have in virtue of its relation to something that is not one of its parts.

That said, could a true belief derive value from a property it has in virtue of its relation to something that is not one of its parts? Rabinowicz & Ronni-Rasmussen (1999) think so, and they hold an additional fact to be true in cases in which we value something *because of* a relational property it has. In such cases, it will be true that we value such a thing for its *own sake*. Such things will be, accordingly, 'finally' valuable. Borrowing from O'Neill (1992), Rabinowicz and Ronnow-Rasmussen give several examples that support the idea:

...A wilderness may be valuable...[more so than otherwise] because of its being untouched by humans, which is clearly a relational property. Another class of examples involves cases in which a thing is valued for its own sake in virtue of its special relationship to a particular object, event, or person. An original, say, an original work of art, may be valued for its own sake precisely because it has the relational property of being an original rather than a copy. Its final value supervenes, in part, on its special causal relation to the artist. Princess Diana's dress may be another case in point. The dress is valuable just because it has belonged to Diana. This is what we value it for. (Rabinowicz and Ronnow-Rasmussen 1999: 8)⁷⁴

⁷⁴ *Ibid.* p. 8. The notion of *supervenience base* that Rabinowicz & Ronni-Rasmussen use here is one they contrast in several papers with the related notion of *constitutive ground*. As I've explained in Chapter 4, the constitutive grounds and supervenience base of an object's value come apart, and just how they do is I think best understood by considering that what counts as the 'source' of something's value (which is, without further qualification, ambiguous.) Suppose we're talking about Diana's dress. In one sense, the source of the value of Diana's dress could be understood as located, for example, in subjects attitudes toward the dress. In this sense, the 'source' of the value of the dress is referred to by Rabinowicz and Ronni-Rasmussen as the *constitutive grounds* of value. Subjectivists and objectivists about value can disagree about whether the constitutive grounds of value lie in subjects attitudes or in the objects themselves, while agreeing about what it is that value *supervenes* on. The supervenience base of an object's value is just those properties of that object on which value supervenes, regardless of whether value is *constituted* by subjects or the objects themselves. Keeping with the case of Diana's dress, the supervenience base of the value of Diana's dress will be properties that are both internal and relational. After all, the properties of Diana's dress that make it her *dress*, rather than say, a grain of sand that she owned, are part of what it is that makes her dress have the value it does. In this respect, value supervenes on properties internal to the dress. But also, value supervenes on *relational* properties of the dress, a fact that is both interesting, and quite important for our purposes. The properties Diana's dress has in virtue of being *her* dress make it more valuable than it would be if it had all the same intrinsic properties, but wasn't actually hers.

Now for a thought experiment from Rabinowicz & Ronnow-Rasmussen: suppose that there are two dresses, one that belonged to Diana and another created by lasers; the latter is an intrinsic duplicate of Diana's dress was never hers. The two dresses have, *ex hypothesi*, all the same intrinsic properties, however Diana's dress is more valuable to have. The extra value we take Diana's dress to have is value it has in virtue of its *relational properties*, properties the dress has in virtue of its being worn by Diana. That we value Diana's dress over an intrinsic duplicate—in virtue of its non-intrinsic, *relational* properties—accounts for why we take her dress, and not the duplicate, to be something we value *for its own sake*—we'd continue to prefer it to other things that are intrinsically identical. Therefore, Diana's dress is *finally valuable*—much in the same way that, as Wayne Riggs argues, an Olympic gold medal that you've won fair and square is more valuable, because of its relational properties, than it would be were you to have simply had an identical replica made.

6.4 Knowledge and Final Value

The picture of final value that Rabinowicz and Ronni-Rasmussen give us is one that we may apply, as Duncan Pritchard has, to the specific question of how a true belief might gain value—value it wouldn't have otherwise—in virtue of its relation to something else.

Following in Pritchard's footsteps, let's return then to the puzzle we were left with at the end of the previous section. The question at hand was whether true belief would gain additional epistemic value, value it wouldn't have had otherwise, by standing in a certain relation to epistemically valuable virtuous agency. Our reasoning was that *unless* it does, then it's unclear why true-belief-through-ability (knowledge) would be more epistemically valuable than a mere true belief, and that's so *even if* intellectually virtuous agency is epistemically valuable in its own right, over and above its connection to truth.

We can begin by asking then if knowledge is like Diana's dress; we valued Diana's dress over an intrinsically identical copy because of *relational properties* of her dress that an intrinsic duplicate lacked. Might we similarly value knowledge over mere true belief because of the *relational properties* a true belief has when known, but would lack otherwise? Put another way, is the property of 'having been a product of virtuous agency' enough to make a true belief epistemically valuable for its own sake?

While this property sounds comparatively less seductive than the property of 'being something of Diana's (alas, any given belief might be thought additionally valuable if Diana held it!), a better comparison is made

when considering Riggs' case of the Olympic medal.⁷⁵ An Olympic medal earned is something more valuable than the same Olympic medal would be if stolen, and this is not due obviously to any intrinsic property of the medal, but rather, to the fact that in the former case but not the latter, the medal is the product of what your athletic achievements have earned you. The Olympic medal would accordingly be more valuable than an intrinsic replica because *athletic achievements* are valuable, and the medal earned rather than stolen has the relational property of being a product of athletic achievement.

We should not be too quick to draw just yet the tempting analogy from the value of athletic achievements of which Olympic medals are a product to the value of intellectual virtues of which we take knowledge to be a product. It should first be noted that in Riggs' example of the Olympic medal, it's the athletic *achievement* and not merely athletic virtues alone that is doing the work. After all, if Riggs' hero had exhibited all the same athletic virtues but finished outside the top three, then any Olympic medal he might acquire (i.e. steal) would not be as valuable as it would be were it a product of an achievement.

By parity of reasoning, then, a *known* true belief would not be more epistemically valuable than otherwise just because intellectually virtuous agency is epistemically valuable and known true beliefs (unlike mere true beliefs) have the relational property of being the product of epistemically valuable intellectual agency. Rather, known true beliefs would be more epistemically valuable than not-known true beliefs because *cognitive achievements* are epistemically valuable, and known true beliefs (unlike mere true beliefs) have the relational property of being the product of finally valuable *cognitive achievements*. Moreover, the epistemic value of a cognitive achievement is a value that derives from both components of our *veritistic value pluralism*: cognitive achievements are epistemically valuable in part because correct belief is valuable and in part because correct agency is valuable. Sosa's (2007) view of the epistemic value of cognitive achievements, or as he calls them, apt beliefs, falls squarely in line with this suggestion. Sosa claims that manifestation of competences, or intellectual virtues, can 'enter into the constitution of something with fundamental epistemic worth: namely, the apt belief (i.e cognitive achievement), true because competent.' (Sosa 2007: 88)

On the present view, thus, the value that *knowledge* has, over and above mere true belief, is a value derived in part from the value of correct belief and in part from the value of correct agency.

If it is not already clear, then it should be made clear why our picture sidesteps Zagzebski's machine-product argument: we have not tried to explain how a true belief produced through virtuous agency is valuable

⁷⁵ Sosa (2007) in his lecture "Epistemic Normativity" gives an example of a ballerina that I think illustrates this point just as well.

by relying *in any way* on the thought that known true beliefs are more valuable *because* they're the product of virtuous agency. On our view, the story is that the additional value known true beliefs have is instead a value they have because *achievements* are valuable for their own sake—and not simply because *what produced* the belief (i.e. intellectual virtue) is valuable. This exonerates us from Zagzebski's critique of the reliabilist.

That said, as John Greco (2009)⁷⁶ points out, a cognitive achievement is not implied by any of the following combinations:

- (i) Cognitive success (true belief) in the absence of cognitive ability
- (ii) Cognitive ability in the absence of cognitive success (true belief)
- (iii) Cognitive ability compresent with cognitive success.

A cognitive achievement requires, in addition, that the cognitive success be *because of* the cognitive ability, and not merely compresent with ability. Sosa (2007) offers an archery example that captures this point, that achievement is distinct from mere compresence of success and ability. Here's Sosa:

An archer might manifest sublime skill in a shot that does hit the bull's-eye. This shot is then both accurate and adroit. But it could still fail to be accurate *because* adroit. The arrow might be diverted by some wind, for example, so that, if conditions remained normal thereafter, it would miss the target altogether. However, shifting winds might then ease it back on track towards the bull's-eye. Though accurate and adroit, this shot would still fail to be "apt," that is to say, accurate *because* adroit. (Sosa 2007: 79)

Consider now the epistemic case: in the original Gettier case, for example, it is Smith's judicious weighing of the evidence that the man who will get the job has ten coins in his pocket that causes him to believe this; Smith's belief is true, however his cognitive success is attributable not to his own efforts, but to the sheer chance that Smith was unexpectedly hired while unwittingly carrying ten coins. Thus, his cognitive success, even though an indirect product of his having formed the belief virtuously, does not constitute a cognitive achievement. The idea was:

Cognitive achievement: For S , p , S 's true belief that p is a cognitive achievement only if S 's coming to believe p truly is attributable to S 's abilities.⁷⁷

We may now sketch more formally the argument I've been developing here:

1. The *true-belief-through-ability* template theory of knowledge requires, by definition, that known true beliefs

⁷⁶ Sosa (2007) makes a similar case in his discussion of aptness.

⁷⁷ In Chapter 11, I'll refine this claim, but it will suffice to leave it in this general form for present purposes.

are ones gained through ability.

2. A true belief is gained through ability just when the cognitive success consisting in the true belief is attributable to the agent's abilities.
3. Cognitive success attributable to an agent's abilities constitutes a cognitive achievement.
4. A true belief that is the result of a cognitive achievement is epistemically better than a true belief that is not.
5. Therefore, on the *true-belief-through-ability* theory of knowledge, knowledge will be more epistemically valuable than mere true belief.

6.5 Addressing the Type Criterion and the Token Criterion

I think some perspective is in order before moving forward, so that our current position is located more clearly within the broader dialectical argument. In Chapter 4, I argued that a theory will satisfy the Meno Requirement only if it accommodates (put here roughly) the thought that, from the *epistemic point of view*:

- (1) *Tokens* of knowledge are more epistemically valuable than corresponding tokens of mere true belief to an extent proportionate to the non-cognitive value that would be generated by believing that truth.
- (2) Knowledge, as a *type* of epistemic standing, is more valuable than mere true belief, as a type of epistemic standing.

It was argued that, as a prerequisite for satisfying these requirements, we must have some background theory that tells us just what the epistemic point of view is supposed to be. The purpose of Chapter 5 was to defend such a theory, and the conclusion was that the epistemic point of view is one from which both the good by way of agency and belief matter for the purposes of epistemic evaluation.

With this being our account of the epistemic point of view, I offered at the end of Chapter 5 some reasons for thinking that the sort of theory most promising for the purpose of satisfying the Meno Requirement would be one that builds the good by way of belief (truth) and the good by way of agency (intellectual) virtue into an analysis of what knowledge is.

In the present chapter, I've attempted to make good on this promissory note. In showing why evidentialism and reliabilism lack the resources needed to account for explaining the value of knowledge, I suggested that—underlying the surface explanation for why neither theory is up for the task—was a shared commitment to *epistemic centralism* analogous to the sort of *ethical centralism* that has been a target of Bernard

Williams. The idea I advanced here was that, against the background of having rejected *veritism* in favor of *pluralism*, our prospects for accounting for the value of knowledge will be better if we also reject the sort of *epistemic centralist* framework underpinning evidentialism and reliabilism in favor of an *anti-centralist* framework, one on which the value of virtuous agency could function within an explanation for the value of knowledge. In having shown the need for an anti-centralist theory of knowledge, I in doing so made a stronger case (than what was merely suggested at the close of Chapter 5) for thinking the true-belief-through-ability template theory of knowledge would be *necessary* for preserving the distinctive value of knowledge.

My next move was to argue for *sufficiency*. I've argued here that if knowledge is understood along the template of *true-belief-through-ability*, then knowledge will in fact be more epistemically valuable than true beliefs that are unknown. The argument I've given relied not on the thought that the epistemically valuable virtuous agency that gives rise to a known true belief makes the known true belief more valuable, but rather on the idea that *cognitive achievements are epistemically (finally) valuable*, and that a true belief that is the product of cognitive achievement is more valuable, epistemically, than a true belief that is not. Of course, this will be so just so long as knowledge and cognitive achievements do not come apart. (I'll be defending in later chapters against some challenges aimed to show that they do come apart).

This argument in the previous section, however, was meant just to show the sufficiency of the *true-belief-through-ability* template for accommodating the Meno Requirement in broad terms. Now that we've shown that our template is one on which a true belief would count as something epistemically better if known, we can move to the final task of this chapter, which will be to show how this fact will be preserved when extended to both knowledge types and knowledge tokens (Chapter 4).

Recall what it was we've said to be the First of the Two Criteria that must be met in order to satisfy the Meno Requirement.

Type Criterion (Meno Requirement): The Meno Requirement holds a theory of knowledge to preserve that, from a *purely epistemic point* of view, knowledge *qua type* of epistemic standing (on the theory) has a purely epistemic value that exceeds that of mere true belief *qua type* of epistemic standing.

Given that a theory of knowledge specifies conditions under which any epistemic standing counts as knowledge, the Type Criterion would be satisfied so long a true belief's being the product of a cognitive achievement would make it invariably more valuable, from the epistemic point of view, than were it not the

product of cognitive achievement. No matter how important or trivial the proposition at hand, the epistemic standing we have with respect to the proposition counts as knowledge only if the truth of the belief is a product of a cognitive achievement; it will be the product of a cognitive achievement if and only if the true belief is attributable to the knowing agent. Because our template theory defines knowledge as *true-belief-through-ability*, an agent knows only if his true belief is attributable to his manifestation of ability, and thus, our theory of knowledge implies that one knows only if her true belief has a certain property—the property of being the product of an achievement—which would *always* make it more valuable to have than if it weren't known (and thus not a product of an achievement). Thus, knowledge understood as *true-belief-through-ability* satisfies the First Criterion of the Meno Requirement.

Let's turn now to the Token Criterion:

Token Criterion: (Meno Requirement): The Meno Requirement holds a theory of knowledge to preserve that, from a *purely epistemic point* of view, the extent to which a given knowledge token will be more (epistemically) valuable than a corresponding true belief token will be directly proportionate to how important (as opposed to pointless) it is that the truth is known.

We turn now from the value cognitive achievements have in virtue of being cognitive achievements to the value they would have in virtue of being *particular* cognitive achievements. Some cognitive achievements, for example—figuring out correctly, after weighing carefully testimony and evidence—which criminal to convict is more important, from the epistemic point of view, than an achievement consisting in believing truly that it is 4:00 as a result of carefully consulting a clock. It follows straightforwardly then that our theory of knowledge will always make it possible that a given true belief that satisfies the condition of 'being through ability' will be more epistemically valuable than a corresponding true belief that fails to satisfy the condition of 'being through ability' *on the basis of, and proportionate to*, how important it is that that belief is known.

This is because any true belief that counts as knowledge will, by being such that it's produced through ability, also be one that is the product of a cognitive achievement. And what makes *any given cognitive achievement* epistemically valuable owes both to:

- (i) the fact that it consists in a true belief (no matter what it is) that is a result of intellectually virtuous agency (something that makes knowledge as a *type* of standing invariably more valuable than mere true belief), and *additionally*,
- (ii) to how epistemically important it is that the belief that *that particular* achievement affords us be known. The more important it is, from the epistemic point of view, that that particular belief is known, the more

epistemically valuable the cognitive achievement will be—something that makes our theory of knowledge capable of accommodating what the Second Criterion requires: that any given token of knowledge will be more epistemically valuable than a corresponding true belief token on the basis of how important it is from the epistemic point of view that that true belief is known.

7

Knowledge, Credit and Testimony

7.1 Introduction

In the previous chapter, we saw that the Meno Requirement not only motivates a theory of knowledge on which there must be an essential connection between true belief and intellectual virtue or ability, but also, that if such a theory could be defended, then it *really would* be one on which we could potentially account for the distinctive value we take knowledge to have. Indeed, we saw that the *cognitive achievements* that occur when an agent's getting to the truth owes to her cognitive abilities are themselves of final value, and further, that this final value matches what the Meno Requirement demands a theory of knowledge account for. And so if knowledge inherits the final value of cognitive achievements, then a theory of knowledge can accommodate the Meno Requirement so long as it can show that a true belief connected in the right way to ability is what's *essential* to knowledge.

If, however, we can't defend an account of the nature of knowledge in terms of this connection between true belief and ability, then the final value solution defended in the previous chapter would have to be given up—after all, it was a solution that would be useful to satisfying the Meno Requirement only if we can defend an account of knowledge on which knowledge plausibly *has* this value. We can see, then, the important sense in which our project would be threatened if it turned out that knowledge and cognitive achievements come apart.

Duncan Pritchard (2009), while in agreement with the conclusion drawn from the last chapter that cognitive achievements are finally valuable, has suggested that knowledge and cognitive achievements *do* come

apart, and that this much should lead us to reject the thought that knowledge is *finally valuable*.⁷⁸ We now have some very good reasons to try to resist his conclusion: after all, his conclusion requires we abandon what I've shown to be a compelling (and presently, the only clear) way to address the Meno Requirement, and also one that squarely motivates what Pritchard (2007*a*; 2007*b*; 2008*d*; 2009), Greco (2009*a*) and Sosa (2007) have either explicitly or implicitly endorsed as *ability intuition*—the widely-shared thought that knowledge is something you get through ability.⁷⁹

There is, thus, a symmetry between the value intuition and the ability intuition about knowledge: each, in its own way, requires that knowledge be understood along the lines of a true belief appropriately connected to ability. If Pritchard is right that knowledge and cognitive achievements come apart, then knowledge can't *essentially* be understood along these lines—and if it can't, we must part ways with what now seems like a rather simple and elegant bridge between our insights about the value of knowledge and nature of knowledge.

7.2 Lackey's Argument

Pritchard reasons that knowledge and cognitive achievements come apart by supposing that not *all* instances of knowledge require cognitive achievement. Now, as cognitive achievements are ones where the cognitive success at issue is something we credit to the agent's abilities, a natural extension of identifying knowledge with cognitive achievement will be to identify knowledge as true belief for which the agent is *creditworthy*.⁸⁰ Jennifer Lackey (2003; 2006) provides a well-noted case designed to challenge the thought that knowledge has such a condition. What Lackey rejects is what she calls the *credit thesis*:

CREDIT: For S, p, if s knows p, then it is to S's credit that S believes p truly.⁸¹

⁷⁸ Pritchard (2009) in "Chapter 3: Anti-Luck Virtue Epistemology" argues on these grounds that the tertiary value problem (see Chapter 1) can't be solved by simply appealing to the final value of cognitive achievements.

⁷⁹ This is Pritchard's term of phrase; also see Carter (2007) "Value Holism and Epistemic Parts" for a defense of this intuition.

⁸⁰ The notion of credit here is meant only in the weak sense of attribution—it is not "round of applause credit" (thanks to Duncan Pritchard for making this distinction in conversation).

⁸¹ A standard way to think about intellectual or cognitive credit is offered by John Greco (*forthcoming*), who takes as a model Joel Feinberg's view of moral credit attributions. Greco gives the following general template for *intellectual credit*—the sort of credit that a variety of virtue epistemologists think must be attributable to an agent who has knowledge:

Intellectual credit (Greco): S deserves intellectual credit for believing the truth regarding p only if:

- (i) Believing the truth regarding p has intellectual value
- (ii) Believing the truth regarding p can be ascribed to S, and
- (iii) Believing the truth regarding p reveals S's virtuous cognitive character.

Lackey rejects CREDIT by arguing that satisfying the CREDIT condition is not necessary for one to have knowledge. Lackey argues that if CREDIT is false (which she thinks it is), then it follows naturally that knowledge and cognitive achievements must come apart, a result that would threaten directly our present approach.

Now, Pritchard (2008*c*; 2008*d*; 2009)—though sympathetic to the general line taken here by Lackey—is right to note that her conclusion is ambiguous. For purposes of clarification, it’s important to highlight a subtle difference between what Lackey calls CREDIT on the one hand, and the sort of credit thesis that actually falls out of the claim that knowledge consists in cognitive achievement on the other. Pritchard (2009) understands the difference here as follows:

... We need to make a distinction between a true belief being *of credit* to an agent, in the sense that the agent is deserving of some sort of praise for holding this true belief, and the true belief being *primarily creditable* to the agent, in the sense that it is to some substantive degree down to her agency that she holds a true belief. Lackey’s focus when employing this example is on the former claim, and this is not surprising since a number of commentators...have expressed their view in such a way that it seems to straightforwardly support this thesis. That said, strictly speaking the robust virtue-theoretic proposal is the latter claim. (Pritchard 2009, Ch. 2: 39)

Pritchard here notes that CREDIT, as formulated, is a very weak thesis; after all, he thinks, it’s always at least of *some* credit to you that you have the true beliefs you do. Those who endorse the view that knowledge consists in cognitive achievement want to say something stronger—that a knower’s true belief must be *primarily* creditable to her agency. Lackey’s objection wouldn’t work against the weak CREDIT, then, and gains its philosophical import only if understood as one that challenges the subtly distinct yet stronger claim that falls out of a commitment to the view that knowledge consists in cognitive achievement. Therefore, the way to engage with Lackey’s argument will be to understand it as a challenge to the controversial (rather than uncontroversial) thesis endorsed by her opponent:

CREDIT*: For S , p , if S knows p , then S ’s having believed p truly is *primarily* creditable to S ’s cognitive abilities.

For the present purposes, I’ll refer to theories that accept CREDIT* as ‘credit theories.’ These will include all views according to which knowledge consists in cognitive achievements which themselves are generated just when one’s cognitive success is primarily creditable—in the sense of of being attributable to—the agent’s cognitive abilities. For the sake of convenience and style, I’ll use a variety of locutions such as

‘creditable’ ‘deserving of credit’, ‘creditworthy’ etc. to mean *primarily creditable* (to an agent’s abilities) in the sense captured by CREDIT*; we are now leaving the idea expressed by CREDIT behind.

That said, let’s consider just how, specifically, Lackey attempts to demonstrate that credit theory is false. The argument, modified in a way that respects the subtle distinction just mentioned, goes as follows:

1. If CREDIT* is necessary for knowledge, then it cannot be the case both that (i) one knows some proposition *p*, and (ii) one’s believing *p* truly is not creditable to her cognitive abilities.⁸²
 2. In a wide variety of cases in which one comes to gain knowledge through testimony, her having believed truly is not creditable to her abilities.
 3. Therefore, CREDIT* is not necessary for knowledge. And *ipso facto*—
-
4. The ‘credit theory’ of knowledge is false.

Lackey supports (2) in the argument by offering a testimony case, and testimony cases indeed have a structure that appears to challenge our sort of proposal. Here Sosa (2007) confesses:

Any belief that is knowledge must be correct, but must it be correct due to an epistemic competence? That seems strained at best for knowledge derived from testimony. (Sosa 2007: 93)

Lackey’s much-discussed case highlights this strain:⁸³

Lackey’s Sears Tower Case: Having just arrived at the train station in Chicago, Morris wishes to obtain directions to the Sears Tower. He looks around, approaches the first adult passerby that he sees, and asks how to get to his desired destination. The passerby, who happens to be a Chicago resident who knows the city extraordinarily well, provides Morris with impeccable directions to the Sears Tower by telling him that it is located two blocks east of the train station. Morris unhesitatingly forms the corresponding true belief. (Lackey 2003: 10)

Lackey adds also that:

There is nothing that is particularly unusual about [*vis.* this case], and it is nearly universally accepted that a situation such as Morris’s not only can but often does result in testimonial knowledge. (2003: 10)

And so what Lackey wants to say is that in circumstances like these, Morris can come to know where the Sears Tower is even though it is false that the true belief Morris receives from the informant is one we’d credit to his abilities. As premise (3) suggests, credit-theories of knowledge, of the sort defended by Greco

⁸² Here and elsewhere I am using the language of credit loosely to avoid repetition; in all instances, though, I’m not implying any more robust notion of creditworthiness than the perfunctory notion of ‘attributable (primarily) to one’s abilities’, as stated in CREDIT.

⁸³ See Wayne Riggs, John Greco (2009), Ernest Sosa (2007) and Duncan Pritchard (2008, 2009).

(*forthcoming*) and Sosa (2007) as well as myself, hold that knowledge is a type of cognitive achievement creditable to the agent; by reasoning from her purported counterexample to the necessity of CREDIT* in (2), Lackey infers the conclusion (4) that:

...It is precisely this sort of case that shows CREDIT [credit theories of knowledge] to be false. (Lackey 2006: 10)

In order to avoid Lackey's conclusion, we can go one of two ways, both of which take aim at (2). The first way is to reject that testimonial cases such as the one she gives are actually cases of knowing, a move that would make it unproblematic for her hero, Morris, not to deserve credit. The second way is to grant with her that knowledge can be gained through testimony in these sorts of cases, but to then affirm further what she denies—that the agent deserves credit. While the second strategy is more popular amongst defenders of what I've called credit theory (Greco 2009*a*; Sosa 2007), I think there is a very strong case that can be made for going the first route. And so I want to 'bite the bullet' as it were; knowledge and cognitive achievements don't come apart in Lackey's case for the reason that it's not a case of knowledge.

7.3 First line of response

Circularity threatens here: any argument I might be able to offer in favor of thinking that the agent isn't a knower in Lackey's example cannot *itself* appeal to credit theory in order to explain why the agent doesn't know.⁸⁴

With this dialectical restriction in the background, I want to start by making explicit an assumption that serves to drive much of the force behind Lackey's case. The assumption is that we are permitted to generalize from what credit theory's verdict is in the Sears Tower case, to what credit theory must say about the wider class of knowledge thought to be gained through testimony.

If we *are* permitted to make this generalized inference, then it's going to be a really bad thing for credit theory⁸⁵ if it rules that Morris doesn't know. Duncan Pritchard (2009) articulates the worry straightforwardly. Here's Pritchard:

⁸⁴ This is because Lackey's argument is one that puts the matter of whether credit theory is true on trial; thus, it would be questionbegging on my part to attempt to reject a premise of her argument that credit theory is false by appealing to a theory that would be apt to demonstrate this only if Lackey's conclusion is false. To avoid begging the question against Lackey, then, any case I might make for thinking her counterexample to credit theory isn't really a case of knowledge must appeal to considerations that would count against her hero (Morris) knowing *regardless* of whether CREDIT is correct.

⁸⁵ I'm using this here as a blanket reference for the sort of views that, like the template I'm advancing, take knowledge to consist essentially in successes through abilities that are themselves 'cognitive achievements' for which the agent's abilities are creditable.

We would naturally describe such a case [as Lackey's] as one in which the informant's knowledge was straightforwardly communicated to our agent; indeed, if we don't allow knowledge in cases like this then it seems that [were the credit theory, which denies knowledge in these cases, true,] quite a lot of our putative testimony-based knowledge is called into question. (Pritchard 2009: 9)

Let's consider even more explicitly the sort of assumptions that generate this worry. Let this set of assumptions be called the *Case-Source Argument*.

Case-Source Argument

1. Testimony is a common method by which we can come to have knowledge.
 2. Morris' epistemic situation in the Sears Tower case is at least as good as the epistemic situations we take ourselves to have in most cases whereby knowledge is thought to be acquired through testimony.
-
3. Therefore, *most of what we take ourselves to know by testimony* can be preserved by credit theory only if credit theory counts the true belief Morris comes to have about the Sears Tower as knowledge.

Everyone should grant that (1) is correct—testimony is certainly a method by which we can come to have knowledge, and even more, we should grant that testimony is a method upon which a *good deal* of our knowledge is acquired. Given that this is so, if (2) in the *case-source* argument is true, then what follows—(3)—would prevent our first strategy for replying to Lackey's argument from ever getting off the ground. This is because (3) implies that the price for denying that Morris is a knower will be more than any theory aiming to preserve testimony as a viable *source* of knowledge could afford.

I have no intention of paying that price. Instead, I will in what follows see our first strategy through and argue, in a way that doesn't already assume credit theory, that Lackey is wrong to think that the Sears Tower Case is one where knowledge is acquired. First though, it will be needed that I defend how, by maintaining this about the case she gives, we do not stand at any tension with the idea that a good deal of what we know is gained through the method of relying on testimony.

7.3.1 Testimony and the Epistemic Situation

Importantly, premises (1) and (2) in the *case-source argument* imply (3) only if a certain theoretical consistency requirement is being assumed, a requirement that functions as an implicit premise in the argument:

Consistency Requirement: If a theory fails to count a given case *C* as a case of knowledge, then it must also fail to count as knowledge the set of all cases relevantly similar to *C*.

If Morris' epistemic situation is 'relevantly similar' to the wider set of cases that includes most of the knowledge we gain through testimony, then in what should this similarity be said to consist? One answer would be to say that the similarity consists in a shared method. Let's call this the 'method view':

Method view: What makes Morris' case relevantly similar to the class of cases that includes our putative testimonial knowledge is a similarity consisting in a shared method (testimony).

If the *consistency requirement* is underpinned by the *method view* about what constitutes the relevant similarity between (i) some case a theory can deny counts as knowledge and (ii) the wider set of cases it must deny count as knowledge, then a theory would in trouble if we didn't count Morris as having knowledge. The consequence would be that other cases in which beliefs are formed *via* a relevantly similar method—in this case, the method of believing on the basis of an informant—fail to count as knowledge.

But the *consistency requirement* should not be underpinned by the notion of similarity captured by the *method view*. To see why, consider Goldman's barn façade case; if the *method view* captures the sort of similarity the consistency requirement should be thought sensitive to, then a theory could fail to count Henry's belief that 'there is a barn' as knowledge only if it also fails to count as knowledge the wider class of beliefs we hold *via* perception, including those of the Moorean sort (e.g. "I have a hand). While the method through which Harry comes to believe there is a barn is relevantly similar to the *method* he would use to ascertain that he has a hand, it would be absurd to think that a theory is committed to skepticism about perceptual knowledge if it fails to count fake barn cases as knowledge.⁸⁶

So the consistency requirement implied in the argument must be thought to appeal to a different notion of similarity. The most obvious choice appeals to the idea that if a theory fails to count some case *C* as knowledge, it must also fail to count as knowledge the wider class of cases in which the epistemic situations in those cases are relevantly similar to the epistemic situation in *C*. The *method view* of similarity should then be distinguished from the *epistemic view*, which tells us this:

Epistemic view: The relevant similarity (that the consistency requirement should be sensitive to) between Morris' case and the wider class that includes putative testimonial knowledge is a similarity that holds in virtue of the *epistemic situations* that make up these cases.

⁸⁶ One who might balk at my confidence that this result would be an absurdity is Sosa, but I should clarify why. If barn façade cases are relevantly similar to Sosa's case of the kaleidoscope perceiver, then he might make the case that you fail to count barn façade cases as knowledge only if prepared to fail to count cases of ordinary perception as knowledge given the closeness of the dreaming possibility. Here I demur; because I do buy into the thought that outright safety is required for knowledge (a claim Sosa rejects), I'll offer that barn façade cases aren't ones of safe belief even though ordinary perception cases are safe (the safety of ordinary perceptual beliefs is not threatened by the nearness of the dreaming possibility.)

We can then give a more precise articulation of the consistency requirement implicit in the move from (2) to (3):

Consistency Requirement (Revised): If a theory fails to count a given case C as a case of knowledge, then it must also fail to count as knowledge the set of all cases similar to C in virtue of the *epistemic situations* that make up these cases.

The consistency requirement understood this way demands that credit theory give the same verdict in cases that are similar to Morris' in terms of the sort of epistemic situation Morris has; a consequence then is that the wider class of cases credit theory must fail to count as knowledge if it fails to count Morris as knowing is a class of cases that will not be restricted to those in which *testimony* is the method used.⁸⁷

We are in a position now to see the important sense in which it is an open question whether (2) in the argument is true—whether Morris' epistemic situation in the Sears Tower case is at least as good as the epistemic situations we take ourselves to have in most cases whereby knowledge is thought to be acquired through testimony. Likewise, we are in a position to be able to determine what sorts of considerations would make (2) true: it is true just in case most of the cases in which we take ourselves to have knowledge by testimony are ones in which the true beliefs at issue are epistemically justified to the same (or similar) degree that Morris' true belief is epistemically justified.

7.3.2 Lackey's case revisited

What has been suggested then is a broader way of conceiving of how the consistency requirement will determine *what else* credit theory would have to give up if it implies that Morris doesn't know. Put simply, the idea is that the better Morris' epistemic situation is, the more *any theory* yields to the sceptic by ruling that Morris

⁸⁷ After all, the extent to which two cases are similar with respect to the epistemic situations they consist is something we should ascertain by comparing the extent to which the beliefs held in the respective cases are (for example) justified to a similar degree. And so if, by hypothesis, I'm equally justified in believing on the basis of perception that there is a dog in the room as I am in believing that the liquid I'm about to drink is poison on the basis of my mother's testimony, then if both beliefs are true, credit theory can deny I know there's a dog in front of me only if it also denies that I know there is poison in the glass, and vice versa. Given that we're supposing by hypothesis that these beliefs are equally justified, then that a theory can count one as knowledge only if counting the other is quite obvious. Thus, what credit theory must deny counts as knowledge if it implies that Morris is not a knower is a class of cases *wider than the class of cases in which beliefs are formed through the method of testimony*. But by the very same token, what credit theory must deny counts as knowledge if it implies that Morris is not a knower is *not as wide* as the set of *all testimonial cases*, a set of cases. After all, within the set including all of our true beliefs reached through the method of perception, some will be more justified than others, and the same holds true for the beliefs we've arrived at through testimony.

doesn't know, and conversely, the less satisfactory Morris' epistemic situation is, the less a theory has to give up. And if it turns out that Morris does not know, then what is yielded to the sceptic would be exactly what a correct theory of knowledge *should* yield to the sceptic, a bulk of propositions that we really don't know.⁸⁸

Now that we see what credit theory (or any theory) would *actually* be committed to implicitly if it denies that Morris knows, let's examine first how good Morris' epistemic situation actually is—a matter from which we could then conclude whether credit theory would be in hot water by failing to count him as knowing.

7.3.4 The structure of knowledge transmission

Without relying on what credit theory itself implies must hold for one to gain knowledge from an informant, I shall investigate the phenomenon of testimonial knowledge by examining different conceptions of it defended in some classic and more recent literature on the epistemology of testimony. I want to start by introducing a rather traditional and widely-shared view, according to which an informant, who knows *p*, can by telling someone else, *transmit* his knowledge to her. There are a variety of views about what has to occur for knowledge to transmit. Let's start with a simple, somewhat flat-footed view that has been defended by (among others⁸⁹) Hintikka (1962), Adler (2006) and Audi (1997).

Knowledge Transmission (KT): If S knows that *p* and S asserts that *p* to H, and H accepts *p* on the basis of S's testimony, then H knows that *p*.⁹⁰

If KT is true, then so long as Morris' informant knows where the Sears Tower is, then so does Morris. But KT is false; transmitting knowledge just can't be this easy.

To see why this is so, just suppose that S knows *p*, and asserts *p* to H, and H accepts *p* on the basis of S's testimony, but should not have. H, suppose, has good reasons for thinking that it is in S's interest to lie to H about whether or not *p*, but H's 'wishful thinking' that S would 'not dare lie to *him*' makes him insensitive to his good reasons for doubting *p* on the basis of S's testimony that *p*. KT rules that H knows *p* just because S knew it, asserted it to H, and H believed it on the basis of S asserting it; but clearly, H does not know *p*. After all, H is

⁸⁸ All of these consequences are outputs of the doubting by analogy principle, a principle that will require that what must be called into doubt will depend on how good the epistemic standing is toward the proposition that functions in the antecedent of the principle.

⁸⁹ Proponents of KT include Hintikka (1962), Welbourne (1979; 1981; 1993) Evans 1982, Ross 1986, Adler 1996, Audi 1997,⁸⁹ see Adler (2006).

⁹⁰ Adler (2006).

not even justified in believing p on the basis of S 's testimony, given that he has good reasons for thinking S would say p when $\sim p$.⁹¹

Testimonial knowledge, thus, is not a phenomenon in which knowledge is transmitted *just on the basis* of our believing a knower's testimony. If knowledge were transmitted so easily, then in a world in which, unbeknownst to us, everyone is an omniscient truth teller, the best way to gain knowledge would be to be maximally naïve, and to believe everything everyone tells you just because they said it. After all, we're defending knowledge as the sort of thing one has when her true belief arises from intellectual virtue⁹² and there's nothing intellectually virtuous about naivety. Testimonial knowledge, quite obviously, is not gained *just because* a truth-telling knower tells us something that we thereby believe.

7.3.5 Testimony and the structure of reasons

The problem with KT, then, is that it overlooks the fact that you can't gain knowledge from an informant unless certain background conditions are in place. After all, whether some knower A 's telling S p counts as a reason for S to believe p depends at least in part on what background, non-testimonial evidence S has for thinking that A would tell S p only if p , and $\sim p$ only if $\sim p$.

Crispin Wright (2007) has gone to some lengths to establish just what these sort of background conditions would have to consist in. Wright characterizes this point as follows: for propositions a , b , whether we are justified, or 'warranted'⁹³, in believing a on the basis of an informant telling us b is not a matter that would be settled independent of what our epistemic standing to some further c would be, where c is some

⁹¹ And even more problematically, given H 's good reasons to think that S would lie to him about whether or not p , S 's testimony that p constitutes a good reason for H to believe *not* p . KT, then, mistakenly allows that you can know some proposition that you're justified in believing to be false. That KT is mistaken does not undermine the idea that knowledge transmission occurs, or that testimonial knowledge requires knowledge transmission. What we should include, rather is that a view about the conditions under which knowledge transmits had better not allow that you can know p in the absence of any good reason to believe p , or alternatively, when the reason you have for believing p is a bad one. To stress, in the presence of positive reasons to believe S would lie to us, for example, there simply isn't any *good reason* to accept p on the basis that S said p . Failing to require that an agent have a reason to believe what he knows is an common objection to, more broadly, any view that allows that one can know just by satisfying some externalist condition which itself does not require any evidence on behalf of the knower.

⁹² It might be said here that our naïve epistemic agent who, in the world of omniscient truth-tellers, gains knowledge maximally, would thereby be intellectually virtuous because, in such a world, naivety is a reliable belief forming process, and thus, an intellectual virtue. However, the absurdity of ruling the maximally naïve intellectually virtuous stands as evidence for denying that intellectual virtue is solely a function of reliable truth-conduciveness and nothing else.

⁹³ Of warrant, Wright says: "Let a warrant for a belief be, roughly, an all-things-considered mandate for it: to possess a warrant for P is to be in a state wherein it is, all things considered, epistemically appropriate to believe P " (p. 6). Although Wright says later that he does not intend his account of warrant to be committed to either internalism or externalism, his discussion of warrant—at least in this debate—suggests strongly that warrant is not divorceable from some form of justification internalistically construed.

proposition that if true would legitimize us in taking *b* to count as evidence for *a*, and if false, would fail to do so. Here's Wright:

Whenever the belief that P is formed for reasons, we will characteristically be able to find a number of *justificational triads*, each of the following form

- 1 A kind of evidence that P that constitutes the reasons for believing it
- 2 The proposition that P itself
- 3 An *authenticity-condition* (Wright 2007: 5)

Wright offers this view of the structure of reasons to apply straightforwardly to cases of testimony; in such cases, he says (1) would be the evidence consisting the informant's telling *p* to the receiver; (2) would be the proposition the informant comes to believe (i.e. that the Sears Tower is located in a certain place). And (3), the *authenticity condition* is a condition that, were it not to hold, the evidence would not indicate the truth of the proposition believed.⁹⁴ Of this condition, Wright says:

... the third element of the triad—the authenticity-condition— is related to the first two like this: a thinker who doubted 3 could not rationally believe 2 just on the basis of evidence 1. (Wright 2005: 6)

Wright makes a fair point here: if you *doubted* your informant was either honest or sincere (what he takes to be the salient authenticity condition in cases in which beliefs are held on the basis of an informant's testimony), then it wouldn't be rational for you to believe what your informant tells you just on the basis that he said it.

Because you obviously can't know some claim *p* on the basis of reasons from which you can't rationally infer *p*, you can't gain knowledge on the basis of testimony *T* if you doubt whether the conditions hold under which *T* would indicate the truth of the proposition your informant tells you.

That said, let us bring into focus the authenticity conditions at play in the case of Morris. Following in Wright's footsteps, I offer that the following conditions could not be doubted to hold if Morris' informant warrants him in believing that the Sears Tower is where he says it is.

(3): Morris' informant is neither untruthful nor likely to be mistaken (*Authenticity Condition*)

Given that Morris' evidence and belief then are:

⁹⁴ This is the way Wright presents it in the case he calls 'The Stranger' case, one which, for all he says, could just *be* Lackey's Sears Tower Case insofar as there is no fact relevant to the receivers respective epistemic standing that applies to one case but not the other.

(1) The informant's telling Morris that the Sears Tower is at location *L*.

(2) Morris' belief that the Sears tower is at *L*.

Morris would lack any good reason, and indeed as Wright suggests, be irrational, if he believed (2) on the basis of (1) whilst nonetheless doubting (3): that the informant is a truth-teller and/or not likely to be mistaken. And so we have established one way in which Morris' knowledge in the Sears Tower case might be defeated. Let's call this, then, the *defeat scenario*:

Defeat scenario: Morris doubts that the informant is truthful and/or unlikely to be mistaken.

Nothing about the case as Lackey presents it, though, would entitle us to think that Morris is in the *defeat scenario*, for the scenario supposes that Morris actually doubts the honesty of competence of the informant; but Morris by stipulation doubts neither of these.

Lackey's case could do the work it's meant to do if Morris should be thought to not only *not* be in the defeat scenario, but also to actually be in a scenario in which he clearly knows. Doubting the authenticity condition is sufficient for undermining Morris' knowledge. But the relevant question becomes: what must Morris' standing be with respect to the authenticity condition in order for him to come to know on the basis of his informant's testimony?

One response to this question would be to take the line taken by Pryor (2004) and Davies (2004), which I'll call the *dogmatist* response.

Dogmatist's response: A thinker's having the evidence (2) suffices to give her warrant for (3) provided she *has no reason to doubt* (1) (i.e. the authenticity condition). (Wright 2007: 10)

If the dogmatist's response is correct, then so long as Morris doesn't have a reason to doubt the honesty and competence of the informant, he is warranted in believing where the Sears tower is on the basis of the informant's telling him.

Upon further examination, the prospects are not promising that we could vindicate Morris as a clear knower by supposing that all he has is a testimonially-based true belief coupled with dogmatist's warrant. This is

because one could have no reason to doubt that the informant is honest and competent, and nonetheless quite clearly *not* know the belief he forms.

We can easily construct such a case by imagining a type of testimonial ‘fake barn’ scenario⁹⁵, in which A has no reason to doubt what B tells him, but B happens to be the lone truth-teller in a crowd of liars. To be clear, it’s not being suggested here that we build in to Morris’ case that his informant happens to be standing amongst a crowd of liars. We’re taking Lackey’s case to be as straightforward as she wants it to be. The relevance of the testimonial fake barn case is that it shows us that *whatever conditions Morris has to satisfy* in order for his epistemic situation to be as good as those ruled as knowledge, they must require more than what that *dogmatist’s* warrant-conditions require.

If Morris would have to meet more stringent conditions in order to know than simply *having no reason to doubt* that his testifier is truthful, then this would count against a prominent view in the epistemology of testimony that Lackey herself rejects as too weak. The view is an implication of *non-reductionism*:

Non-reductionism: testimonial knowledge can be acquired merely on the basis of a speaker’s testimony—so long as the hearer does not have any relevant defeaters. (Lackey 2006: 11)

Lackey, by rejecting non-reductionism, is committed to the very idea our own reasoning has led to, which is that simply believing truly on the basis of the informant’s testimony and having no relevant defeaters isn’t enough to make Morris a knower. And so Lackey must think that Morris has satisfied some condition over and above what non-reductionism requires for knowledge. What Lackey herself rejects non-reductionism in favor of is a stronger position implied by *reductionism*: (Lackey 2003: 23)

Reductionism: The hearer [in order to gain testimonial knowledge] must always have some appropriate non-testimonial positive reasons for accepting a speaker’s report. (Lackey 2006: 11)

If *reductionism* is true then, as Lackey says, ‘testimonial knowledge can be acquired only when the hearer has relevant background information for accepting the report in question.’ (Lackey 2006: 11) And so,

⁹⁵ One might object here for the reason that I am ‘building in’ to the Morris case what was not originally there. It is true that I am here, however, this is orthogonal to the point being made, which is that Lackey cannot explain Morris’ knowledge within the assumption that the *dogmatist’s* principle accurately captures the requirements one must meet in order to know. The dogmatist’s principle, if correct, would be meant to work for *any* case. The fact that it doesn’t (i.e. by making knowledge too ‘easy’ to achieve) disqualifies it as an acceptable explanation that can be given to explain Morris’ presumed knowledge; the result of course is just that for Morris to know in the *original* case, it will be needed that he has a stronger standing to the authenticity principle than merely failing to doubt it, the requirement of the dogmatist.

reductionism implies that you don't get testimonial knowledge cheaply, *simply* on the basis of someone telling you something, and in the absence of positive independent reasons for accepting what the informant says.

Reductionism is, in all relevant respects, a specific view about testimony in alignment with what Crispin Wright calls *conservatism*, which requires for warrant a stronger standing with respect to authenticity conditions than merely having no reason to doubt them. With regard to warrant we might have for believing a proposition on the basis of some evidence, Wright says that *conservatism* 'represents these warrants as hostage to the possession of collateral information about authenticity-conditions,' (Wright 2007: 10) collateral information that the dogmatist doesn't ask us for, and that the non-reductionist denies that testimonial knowledge requires.⁹⁶

Therefore, in the absence of the sort of positive reasons reductionism takes to be necessary for testimonial knowledge, Wright denies that you will possess warrant. Because on Wright's view warrant is not factive, knowledge is stronger than what Wright takes warrant to be. And thus, if a lack of 'collateral information' robs a testimonial recipient of warrant, it's going to by implication also rob us of knowledge. That said, let's consider the reductionist's commitment more carefully:

The hearer [in order to gain testimonial knowledge] must always have some appropriate non-testimonial positive reasons for accepting a speaker's report. (Lackey 2006: 11)

What would make the relevant sort of non-testimonial positive reasons 'appropriate' ones for the purpose of making the hearer, by having them, a knower? Lackey gives us a hint. In a more recent paper, Lackey (2006), with the intention of addressing critical response, tries to get clearer about what we should assume to be 'built in' to the her Sears Tower case—facts about what we are entitled to conclude about Morris' background beliefs about his informant. These facts weren't originally made explicit, but (I think) they could have been reasonably inferred. Here's Lackey:

I certainly never presented CHICAGO VISITOR as involving a hearer who is "simply opening his brain and putting into it whatever some random stranger has to say," much like a robot would. On my view, testimonial knowledge requires that Morris exercise the capacity to be appropriately sensitive to defeaters—if, for instance, he would accept the passerby's testimony even if he appeared highly intoxicated or told him that the Sears Tower was in France, then he would clearly not acquire the knowledge in question. In addition, I require the presence of minimal positive reasons for rational acceptance of testimony—if Morris had no relevant beliefs about humans' general testifying habits, or about the reliability of humans when offering directions, or about Chicago, and so on, then, once again, I would deny that testimonial knowledge has been acquired. (Lackey 2006: 5-6)

⁹⁶ We might think of reductionism as a sort of 'testimonial extension' of Wright's conservatism because Wright's conservatism *implies* reductionism. That is, Wright's conservatism implies that you are not warranted or justified in believing some proposition *p* on the basis of evidence consisting in testimony *T* unless you have positive independent reasons for (and not just an absence of reason to doubt) the authenticity condition, which in testimonial cases, will be that the informant is neither mistaken nor insincere.

If Lackey is right that Morris clearly knows, then the background beliefs about human's general testifying habits and the reliability of humans when offering directions, which are made explicit here by Lackey, need to be good enough to give Morris positive reason to believe that his informant is neither mistaken nor insincere in what he's saying. Otherwise, by the implication of the reductionist view that Lackey herself endorses, Morris could not be said to know .

7.3.6 Morris's evidence: a closer look

To be clear: Lackey *must* be assuming that Morris actually does have good enough positive, non-testimonial reasons for accepting what his informer says: good enough, she must think, to give him testimonial knowledge. These non-testimonial reasons, she says, consist in Morris' 'beliefs about humans' general testifying habits...[and] about the reliability of humans when offering directions.' (Lackey 2006: 5-6) Now beliefs about (presumed sincerity that marks people's) testimonial habits and about the reliability of people in general are ones Morris can take to count in favor of his *own* informant being neither insincere nor mistaken only if he relies on an inference from what people are usually like to the conclusion that his informant will be similar. Morris is of course going to be aware that at least a handful of the millions of people he would be inclined to ask for directions (i.e. people who don't appear drunk, confused, etc.) will offer false testimony. Lackey surely doesn't want to deny that.

Are Morris' reasons for believing what his informant tells him good enough to count him as knowing that what the informant says is true?

No. To make the case for doubt here, I shall appeal to a thought experiment by which it should be clear that the agent in question lacks knowledge; the case will be one in which the evidential support structure will be the same as it is in Morris' case.

Consider the following case:

UNDERCOVER DETECTIVE: Campbell is dropped off by some unknown thugs in the middle of Moscow. He finds out that he is wanted by the KGB. Fortunately for Campbell, the KGB don't know what he looks like, however there are five highly-trained KGB agents who know exactly what Campbell's voice sounds like. In fact, they have studied a recording of his voice so carefully that, were they to hear him speak even a few words, they would immediately know it was him—and would accordingly apprehend him. Now Campbell doesn't know what any of these five KGB agents look like, but as he scans over the people he encounters, he considers that KGB agents would likely not be elderly, especially young, mentally handicapped or drunk. And so he rules all of these people out. Suddenly, a middle-aged man walks up to him and tries to casually strike up a conversation. Nothing about the man suggests he is *not* one of the five KGB agents who could identify him by his voice; after

all, the man is neither elderly, young, mentally handicapped or visibly drunk. Campbell hesitates before speaking, but then considers that there are about 10 million people in Moscow, and so the odds are verly low that the man who approached him is one of the five agents. With the odds on his side, Campbell takes a deep breath and speaks to the man, who chats with him a bit and then goes on his way. The man wasn't an agent afterall.

Did Campbell know that the man he spoke to wouldn't recognize his voice? No, he didn't, anymore than you can know your lottery ticket will lose. But Campbell's evidential support for his belief that the man he spoke to wouldn't recognize his voice is, on closer inspection, precisely the same as Morris's evidential support for believing what his informant tells him. Because the class of folk who Campbell suspects wouldn't be KGB (children, visibly drunk, mentally handicapped, etc.) is approximately the same as the class of folk Morris rules out of the class of folk from whom he would ask directions, Morris and Campbell find themselves in the same situation: Morris trusts the informant's testimony even though he will grant that at least five people out of the millions who would look trustworthy will be mistaken or deceitful. Campbell trusts that the man he speaks to will not recognize his voice even though he grants that five people out of the millions who look like they could be trained to recognize his voice actually would recognize it.

For epistemic considerations that don't already presuppose that credit theory is true, I've tried to give a compelling case for thinking that the counterexample relied upon to undermine credit theory is not an example of knowledge; in arguing for this, I've aimed to draw attention to some overlooked though epistemically significant features of Lackey's case that should compel us to find it no better a candidate for knowledge than cases that would, I think by a reasonably wide consensus, not count as knowledge.

7.4 The second line of response

So I've made a case for thinking that Morris doesn't know. So then what? There is a question from Section 1 that has yet to be addressed: is a theory that denies that Morris knows also committed to denying knowledge in the wider class of testimonial cases? We've said that it would *only if* Morris' epistemic situation is relevantly similar to the epistemic situations at issue in the class of putative testimonial cases.

Again, without arguing the point in a way that presupposes that the epistemic criteria that will be relevant are the criteria countenanced by credit theory, I contend that the Morris case is not relevantly similar to cases of putative, everyday testimonial knowledge. After all, in most cases in which we take ourselves to gain knowledge by testimony, our informant is someone who—in the absence of our beliefs about the testimonial

habits of humans in general and the general reliability of humans—we would *still* have some positive, non-testimonial reason to believe. (Whereas Morris, in the absence of such beliefs, would have not).⁹⁷

Before moving to a consideration of everyday testimonial cases—ones where the epistemic situation is better than Morris', I want to consider briefly an objection to the verdict I've given Morris. It might be objected that in cases like Morris', where the source of the testimony is a complete stranger, the recipient's belief might be good enough for knowledge so long as he actually *is* in a good environment—e.g. an environment where there are very few misinformed or devious testifiers, and so he wouldn't easily have been misled. In such an environment, (it might be argued) Morris could have a safe belief *even though* his epistemic situation is analogous to the KGB case, where Campbell appears to lack knowledge. Even if we grant the point that Morris' belief would be safe in a good environment, we're well within our rights to point out that good environment cases will be only a class of cases within the wider class of cases that includes those where testimonial belief is acquired in conditions like the ones Lackey offers. And so even if Morris' belief would be safe in good environments, this would not suffice to show that the wider class of testimonial beliefs acquired by asking competent-appearing strangers will count as knowledge if true. And it's that stronger claim that Lackey wants to make: that cases like *these* (i.e. ones that meet the conditions met in the Morris case, without any further stipulation that the environment is good) that clearly count as knowledge.

A different line of response to the safety-based objection I'm considering though would insist on a hard-nosed denial that Morris' belief *is* safe, even though the environment is good. This hard-nosed line would not be without some precedent: lottery cases, after all, are ones where—as Pritchard (2007a) suggests—beliefs are not safe even though they are based on excellent statistical grounds. If I'm right that the positive, non-testimonial evidence Morris has for believing his informant is neither deceitful nor mistaken is no better than the evidence Campbell had for believing the man who approached him wouldn't recognize his voice, then Morris' evidence is (like Campbell's) structurally similar to the evidence we have that our ticket in a fair lottery will lose, evidence we have wholly from reasoning from what we take to be true of tickets in general to what we

⁹⁷ For example, regardless of what I believe about humans in general, I can come to know something about France that my history teacher tells me because I have positive reasons for thinking *she* is both knowledgeable about France and has never deceived me in the past. Likewise, I can come to know that my father is mad at on the basis of what my mother tells me, and my reasons for believing her include, among other things, that she cares about my well-being—a reason I hold not because I've inferred that, given that most mothers care about their children's well-being, it's probable that mine does too. The positive, non-testimonial reasons we have for believing what our informants tell us in the wide class of cases a theory actually should count as knowledge are thus quite a bit better than the sort of reasons taken to ground knowledge in Lackey's counterexample to credit theory. And this is true *even though* Morris employ the same method (belief on testimony) that is employed in cases of genuine testimonial knowledge.

infer will be true of that particular ticket. And so if lottery propositions aren't safe, then neither are the beliefs you form on the basis of asking strangers.

7.4.1 Recasting the Objection

Although Lackey has stood adamantly by CHICAGO VISITOR as being 'precisely this sort of case that shows CREDIT [credit theories of knowledge] to be false,' (2006: 5-6) it is worthwhile to consider whether credit theory would be in trouble if the testimonial case used were a better one—i.e. one in which the positive, non-testimonial reasons we have for believing our informant don't owe entirely to lottery-style evidence. Let's take then the case of a trusted History teacher's testimony (about history) and run Lackey's line using this situation instead. Her argument this time begins with a claim we can accept—that someone (say Bill)—comes to know what his history teacher tells him about France, and further, that he has a true belief isn't something he deserves credit for.

Since in this case, the knowledge at issue really is something we could challenge only if calling into question most any other piece of putative testimonial knowledge, it becomes the burden of credit theory to show how Bill *does* deserve credit of the relevant kind for his true belief.

Lackey will have the same story to tell about why Bill doesn't deserve credit as the one she gives for why Morris didn't deserve credit—an explanation that serves as her support for Premise (2) of the original argument (as opposed to the *case source* argument) I've attributed to her. The explanation, meant to hold for all those who gain knowledge by testimony, will be—to put the point bluntly—that they haven't done enough to deserve the sort of credit that credit theory insists must be attributable to their abilities. Her idea is that in a given case in which knowledge is transmitted from informant to recipient, it is the informant's abilities—not the recipient's—to which credit is attributable. In our case, then, Bill's true belief about France isn't something for which we credit his own abilities—after all, all Bill presumably did was take his professor for his word—and it was the professor who actually did the epistemic leg work; Bill, on this picture, owes to the professor's intellectual work and not his own for the fact that he believes truly what he does about France.⁹⁸

⁹⁸ Consider that the same thing can be said about *any* piece of testimonial knowledge, even those that we take to be the clearest examples of it. The idea is that the *method* used—the method of accepting someone's testimony—is not a method by which the cognitive successes gained through the method constitute any cognitive achievement on the part of the receiving agent, and credit is—as was suggested in Section 1—something an agent gets only if the cognitive success is attributable *to her*. So if knowledge is something we get credit for, and credit is something an agent gets only if her cognitive success is something attributable to her own intellectual efforts—then if true beliefs acquired by testimony are not something we can attribute to the intellectual efforts of the individual who acquires them by testimony, then credit is not *necessary* for knowledge, and Lackey's original conclusion stands.

Against this challenge, there are two separate lines of argument that can be made. The first line rejects a certain assumption about credit attributions that loads the argument in favor of the conclusion Lackey reaches. The second line I'll advance builds upon the first, and shows why it is a mistake to appeal to facts about the informant's intellectual efforts in concluding whether a testimonial recipient deserves credit for believing truly what she is told.

7.4.2 Against the disjunctive picture of testimonial credit attributions

The assumption that I think loads the deck in Lackey's favor, and leave credit theory with literally *no* good defense for why a knowing recipient deserves credit, is captured by what I'll call here the *disjunctive view* of testimonial credit attributions.

Disjunctive view: If recipient *R* comes to know *p* on the basis of informant *P*'s testimony, then the fact that *R* believes *p* truly is attributable either to *R*'s intellectual abilities or to *P*'s intellectual abilities.

It should be obvious why the disjunctive view stacks the deck in Lackey's favor. Because we take it that the informant expended more intellectual effort than the recipient, the disjunctive view implies that it will be the informant, and not the recipient, who will always be the one deserving of the credit.

The disjunctive view, however, overlooks something important about what it is for which credit is being awarded. It's not just a 'single item' consisting in a true belief, even though it's the *same thing* believed by informant and recipient. Credit, after all, doesn't attach to beliefs themselves, but to agents' abilities. Whether the informant deserves credit for how she came to hold the belief truly is one matter, and whether the recipient deserves credit for how she came to hold her belief is another. Just because the informant deserves credit for how he came to hold the true belief doesn't mean, as the disjunctive view implies, that the recipient thereby does not. The disjunctive view, therefore, should be resisted for the reason that it mistakenly supposes that there is only one 'nugget of credit' to go around.⁹⁹

In Bill's case, then, no matter how much we want to shower with credit Bill's professor for knowing what he did about France, it's going to be an open question whether Bill himself deserves credit for the true belief *he* acquired by believing him.

⁹⁹ By rejecting the disjunctive view, we thus take a lot of the force behind Lackey's argument; in particular, we've undercut entirely any reason to think that the recipient doesn't deserve credit for her true belief that appeals to facts about the informant having been comparatively *more* creditworthy.

7.4.3 Credit attributions

I've argued that the second line I take builds off the first line, which itself targeted the disjunctive assumption responsible for thinking that a receiver couldn't *in principle* deserve credit, so long as the informant does. To take this idea a step further, I want to now draw from some ideas that are developed in Sosa's (2007) discussion of testimonial knowledge acquisition. Sosa essentially concedes a premise of the opponent's argument—that the informant's efforts in a testimonial exchange are an essential part of the explanation for why the recipient has the belief she does—but Sosa then shows how this concession need not be threatening whatsoever to credit view, and indeed, this concession yields only a rather mundane fact about the social aspect of knowledge acquisition.

An example he gives is one involving a Volvo, which we may suppose was built in some factory F and also contains some defect D. Says Sosa:

It is helpful here to compare the fact that what explains why there is this defective car in existence may not even partially explain why this car is defective. (Sosa 2007: 96)

Similarly, he claims, will be the case for beliefs that we acquire through testimony: 'The error resides...in supposing that what explains there being my true belief is what explains why my belief is true.' (Sosa 2007: 96) This all seems right, and so it wouldn't follow then that the success of Bill's belief is not primarily attributable to his abilities *simply because* his abilities only partially explain why he formed *that* belief. As Sosa points out further, competences are in a way 'socially seated'—our knowledge owes in part to others with whom we are interconnected. He gives a football analogy that is I think useful here:

Something similar holds good of socially seated competence generally. A quarterback may throw a touchdown pass, for example, thus exercising a competence. But this individual competence is only one part of a broader competence, seated in the whole offensive team, that more fully explains the successful touchdown pass, the apt performance of that quarterback. The pass receiver's competence may be crucial, for example, along with the individual competences of the offensive linemen, and so on. (Sosa 2007: 94)

Even though there is a sense in which the quarterback's competences are only partially creditable for his throwing a touchdown (as opposed to an incomplete pass), his touchdown pass is nonetheless an achievement for which we may rightly credit him. And the same holds true in the cognitive arena. Sosa sums up the point as follows:

Testimonial knowledge can therefore take the form of a belief whose correctness is attributable to a complex social competence only partially seated in that individual believer. The account of animal knowledge as apt, creditable belief can

thus explain how testimony-derived knowledge might count as apt, creditable belief, despite how little of the credit for the belief's correctness may belong to the believer individually (Sosa 2007: 97)

In Chapter 11, I'll address in some detail the relationship between cognitive achievements and luck. A careful discussion of this would be needed in order to give a positive defence for the claim that testimony from trusted sources (as with the case of Bill) suffices for cognitive achievement. A perfectly rigorous defense against Lackey's charge needn't include such a defence, it should be noted: our rejection of the disjunctive account of credit attributions along with what Sosa points out about partial credit is sufficient for undercutting the charge we're considering here: that cognitive successes gained through testimony wouldn't be attributable to the recipient's competences. And so if the testimonial case is strengthened to put the recipient in a better epistemic position than Morris—a position whereby we should say knowledge is gained through testimony—it wouldn't follow that the success of the recipient's belief is, in virtue of being formed via testimony, not creditable to her abilities.

And so just as knowledge and cognitive achievements wouldn't come apart in cases where we gain truths by a stranger's testimony, neither would they come apart when the truth gained by testimony is from an expert or trusted source.

8

John Greco on Credit, Saliency and Causal Explanation

8.1 Introduction

In Chapter 7, we saw that the Meno Requirement motivates a template account of knowledge on which knowledge is thought to consist in some essential connection between true belief and ability, and in Chapter 8, I aimed to diffuse a prominent objection to the thought that such a connection really is *essential* to knowing. That said, in order to give a theory of knowledge that satisfies the Meno Requirement, we'll need to show how the true-belief-through-ability template account of knowledge is defensible in its own right. Here, it is useful that we turn to the recent work of John Greco (2006; 2009a) who has provided a comprehensive attempt at making good on such an account. If Greco's account works, then this bodes well for our own attempt to give an account of knowledge that satisfies the requirements imposed by our insights about the value of knowledge. If his account doesn't work, then this will make it clearer what mistakes must be avoided if the true-belief-through-ability template is to be defended as a *bona fide* theory of knowledge.

8.2 Greco's New Proposal

An idea at the forefront of John Greco's (2006; 2009a) recent thinking is that knowers don't just get to the truth any old way, or by some way that is merely reliable, but by coming to believe truly in a way that is primarily creditable to her cognitive abilities. One of his projects has been to develop this general platitude about

knowledge into a more substantive position that functions as an essential component of, rather than a mere fig leaf to¹⁰⁰, a plausible account of knowledge.

Greco accordingly develops, by appealing closely to Joel Feinberg's theory of moral blame, an account of credit attributions meant to provide necessary conditions for any sort of credit attribution, intellectual or otherwise. His suggestion on this score is that, pace Feinberg:

Credit Attribution (Feinberg) S deserves credit for some action Φ only if:

- (i) Φ has a value of kind K
- (ii) Φ can be ascribed to S, and
- (iii) Φ reveals S's K-relevant character. (Greco 2009a, Ch. 6: 15).

With this being Greco's working account of credit attributions construed generally, his next move is to apply it to the cognitive domain, where we may shift our focus from (for example) actions bearing moral value to true beliefs bearing epistemic, or intellectual value. The resulting picture becomes:

Intellectual credit (Greco): S deserves intellectual credit for believing the truth regarding p only if:

- (i) Believing the truth regarding p has intellectual value
- (ii) Believing the truth regarding p can be ascribed to S, and
- (iii) Believing the truth regarding p reveals S's virtuous cognitive character. Alternatively: S's cognitive character is an important necessary part of the total set of causal factors that give rise to S's believing the truth regarding p. (Greco 2009a, Ch. 6: 15)

Now, because Greco thinks knowers deserve credit for getting to the truth the right way, conditions necessary for deserving intellectual credit will function also as necessary conditions within an analysis of knowledge. A note of caution: we must be careful not to make Lackey's (2003; 2006) mistake of reading Greco here as claiming simply that knowledge requires that it *be to the credit of the agent* that her belief is true. This, recall, was misleading: without qualifying further what Greco claims this to consist in, it would appear as though he is claiming something far weaker than he really is—something that is satisfied so long it is at least of *some* credit to an agent that she's come to believe truly. This, however, is not Greco's point: his account of intellectual credit reveals a commitment that we shall see turns out to a, if not *the*, most crucial feature of his account of knowledge: the idea that an agent deserves intellectual credit for the truth of some belief only if the agent's

¹⁰⁰ I borrow this locution from Blackburn's (2001) own critique of virtue epistemology.

abilities are an ‘important’ part of the factors giving rise to her believing p truly. And so insofar as we are to think of Greco’s credit claim as a binding condition on knowledge, we should think of an agent’s believing truly as something we attribute not simply to the agent herself in some wide sense, but rather as something we are attributing specifically, and in some important way, to the agent’s cognitive *abilities*; for an agent deserves intellectual credit only when her *abilities* are sufficiently causally important in giving rise to her true belief. It is only through this step (which Lackey missed) that we can understand, in a way that isn’t misleading, the sense in which knowledge is supposed to be connected to credit. That said, here’s how Greco takes his account of intellectual credit to actually bind an account of knowledge:

Knowledge (Greco): S knows p only if S’s cognitive abilities are an important necessary part of the total set of causal factors that give rise to S’s believing the truth regarding p . (*Ibid.* 16)

For the purposes of getting clear about the connection between cognitive ability and cognitive success, we are in a position to point fingers at what looks to be the *multum in parvo* in desperate need of careful unpacking—Greco’s requirement that cognitive abilities function as an ‘important’ necessary part of the total set of causal factors giving rise to the agent’s true belief.

8.3 Salience and Causal Explanation

What Greco argues is that the importance of cognitive abilities as necessary parts of the total set of causal factors giving rise to a true belief is itself a matter of explanatory salience. That is, for some causal factor F , F ’s importance as a cause of Φ is something we would determine by assessing whether F is comparatively more salient within a correct causal explanation for Φ than other factors within the total set of causes that give rise to Φ .

To make this idea clearer, consider an example: suppose that your dear friend Martin, who you’ve always known to be compassionate, surprises you by, for no apparent reason, pulling out a revolver and shooting your neighbour’s dog. You demand that he explain what on earth caused him to do this, and he tells you: ‘I shot Old Yeller because I kept my shooting arm very steady.’ His response obviously fails to satisfy you, and it would remain unsatisfying even if you were to learn that, had he not kept his shooting arm as steady as he did, he would have missed. After all, when asking him for an explanation, what you were after was not just any causal factor among the set of those that explain the dog-killing act, but the important causal factor—the one

that would best answer the question ‘why?’ As it so happens, the causal factor that would be most salient in such an explanation wasn’t that Martin kept his hand steady, but rather, that he had been slipped a psychedelic pill an hour ago and believed Old Yeller was evil. How does this causal condition relate to luck? Greco seems to understand the relationship between ability and luck by appealing to what I’ll call the principle of comparative salience.

Principle of comparative salience: If S knows p , then S 's abilities are more salient in explaining S 's believing p truly than is luck.

The principle of comparative salience implies that a salience of abilities in bringing about a true belief will be incompatible with a salience of luck, an idea that seems wedded to a further principle which I call *dependence incompatibilism*.

Dependence Incompatibilism (DI): S 's coming to believe p truly because of ability is incompatible with S 's having come to believe p truly because of luck.

Greco attempts to block luck, then, by suggesting that in explaining why a knower came to form a true belief, the salience of an agent’s abilities implies that luck will not be salient, an idea couched in an incompatibilist assumption about dependence on ability as opposed to dependence on luck. We can see that the plausibility of Greco’s account of the cognitive ability and cognitive success connection, as a connection meant to capture what’s essential to knowing, hangs crucially on his doctrine of explanatory salience, and somewhat tenuously so when we consider that different features of an explanation will stand out as salient for different people, with different purposes in mind when an explanation is sought. Salience for Greco is too important to be left up for grabs. What’s needed here is some principled way to preserve salience as being more than something that should just be intuitive.¹⁰¹

8.4. Causal Explanations and Context Dependence

Salience, it turns out, is not taken as basic in Greco’s theory, although his attempt to flesh it out comes at some cost. In particular, his appeal to explanatory salience is part of a package in which his virtue-theoretic account of

¹⁰¹ After all, a failure to make salience something more than a mere hostage of the intuitions would leave Greco's account of knowledge defensible only by means of brow-beating.

knowledge is endorsed alongside a corresponding commitment to a version of interest-dependent contextualism. For the present purposes, I will focus on just those aspects of his virtue-friendly contextualism that illuminate salience. With this said, I turn to Greco's first key thesis, which is that causal explanations are context sensitive. A precise articulation of this idea would be:

Causal Explanations: Causal explanations are context sensitive insofar as: (i) (Φ because φ) implies that φ is a salient part of the causal explanation for Φ , and (ii) φ saliently explains Φ just in case the interest and purposes that make up the explanatory context render φ salient.

I will set aside the relevance of distinguishing (as Greco unfortunately doesn't) between causal relations of the form A caused B and causal explanations, which take the form B because of A , a distinction that functions importantly in theories of the metaphysics of causation, such as those advanced by Davidson (1967), Beebe (2000) and Lewis (1970).

Interpreting Greco charitably we should suppose he has the latter in mind. His claim is that causal explanations, and in particular, whether one thing or another will count as a salient within an explanation, is sensitive to the interests and purposes that frame the relevant context in which the explanation is given. This is a general thesis meant to hold for all reports of causal explanation. So why must causal explanations answer to interests and purposes that make up the context in which the explanation is offered? Greco explains here through an example:

Consider... the case of a car crash at a major intersection. The police at the scene deem that the crash was caused by excessive speed, and accordingly they issue the driver a summons. Later in the year, city planners consider the crash along with several others that occurred at the same location. They determine that the crash was caused by difficult traffic patterns and recommend changes to the Board of Transportation. Who is right, the police or the planners? (Greco 2006: 5)

According to Greco, they're both right. Given the interests and purposes of the city police, their explanation is correct, and given the interests and purposes of the city planners, their explanation is correct. If we grant him this, then salience is no longer wedded at the hip to brute intuition; rather, we'd have a method that is meant to yield an answer to whether the agent's abilities would, in a given case, saliently explain her true belief. Given his commitment to the comparative salience principle, the appeal to context¹⁰² in picking out whether the agent's

¹⁰² It would be good to consider one further expository note concerning how it is that Greco's breed of contextualism differs from those of classic attributor contextualists such as Cohen and DeRose. A prominent point of unison between Greco and attributor contextualists is that all endorse the maxim that, construed broadly, whether S knows p is a matter that is determined in part by the various interests and purposes that make up the relevant epistemic context. Ways become parted when we ask just what it is that context has a hand in determining. On the

abilities are salient will at the same time imply that luck is not salient, and presumably, given dependence incompatibilism, that the true belief would not depend on luck, and thus, could not easily have been false.

Greco, in sum, has given us a virtue-theoretic account of knowledge that (i) appeals to causal-explanatory salience to determine whether the cognitive success was relevantly because of cognitive ability and not luck; (ii) understands causal explanations to be sensitive to the practical interests and purposes that make up the relevant context by determining what will count as salient, and (iii) allows his virtue-theoretic approach to ‘inherit the context-sensitivity of causal explanations’ (Greco 2006: 5) while not inheriting the more traditional contextualist thesis that contextual considerations determine directly the appropriateness of raising and lowering epistemic standards.¹⁰³

In what follows I will outline what I think are actually three lines of defence to be found in Greco’s work for explaining this facet of his contextualism: Considered first will be his appeal to the example of the gambler as an illustration of the connection between interests and the comparative salience of abilities and luck. Second, will be his argument that a consideration of the dependence on certain kinds of informants within a social epistemic environment motivates a default salience of cognitive abilities in explaining true beliefs. Finally, I will consider his suggestion for understanding interests and purposes in a way that upholds two guiding theses about the concept of knowledge advanced by Hawthorne (2004) and Craig (1990).

8.5 Salience, contextualism and interests: three lines of defence

8.5.1 Luck and the Gambler

Greco (2006) offers a case of a gambler that is meant to illustrate how it is that purposes and interests would designate whether abilities are more salient than luck in explaining a true belief. The gist of the example is that a victorious gambler’s claim that he knew a certain horse would win—when he did indeed have unusually good evidence before the race—could be true in a context comprised by the purposes and interests of himself and

classical attributor account, the interests and purposes undergirding the context will determine how high or low the standards for knowing will be. However on Greco’s account, the interests and purposes circumscribe not epistemic standards directly, but rather, ‘the salience of various contributing causal factors, thus determining the responsibility for a true belief.’

¹⁰³ There are plausible reasons for thinking that Greco’s appeal to contextualism, despite its appearances, ultimately just lifts the buck from what would otherwise be an up-for-grabs notion of salience only to pass it to what his commitment to contextualism leaves waiting at the end of the line—that is, to the matter of getting clear about the role interests and purposes are supposed to play so as to dictate that some causal factor Φ would be salient in explaining outcome O while ϕ , on the basis of these same considerations, would not be.

other like-minded gamblers, whilst false when told to his wife, when it is the couple's own purposes and interests—marked by financial hardship and a history of bad luck—that frame the context. (Greco 2006: 6)

The argument Greco gives is that the context comprised by like-minded gamblers is one in which the sagacity of betting on the horse that he did is recognized, and becomes more salient than does the fact that he was (at least to some perhaps small degree) lucky. In the context in which the gambler's claim to have known the horse would win is offered to his wife, rather than to the other gamblers, Greco says: 'Their financial situation and his gambling history make salient the role of luck in picking winning horses, while at the same time undermining the salience of the gambler's abilities.' (Greco 2006: 6) Two points jump out:

(1) In a context made up of other gamblers who understood the unusually high probability that the gambler would win on the basis of the information he had before the race, Greco stresses that the gambler's own abilities were more salient in explaining his success than was whatever luck Greco supposes was at play.

(2). Given this, the most plausible reason to think that, in the gambler-and-wife context, the small bit of luck the gambler enjoyed would stand out as a more salient cause of his success than his abilities would be because the wife's *perceived* interests, and not actual ones, lead her to believe that luck played a larger role than it actually did in explaining why the gambler won.

The thought in (2) is that actual interests are ones that would be sensitive to actual riskiness of the bet (which was very low), rather than merely perceived riskiness that is salient on the basis of previous bad luck.

And so:

Diagnosis: It's unlikely that with respect to couple's *actual* interests luck would be more salient than the abnormally good evidence that made possible such an apt bet. If my suspicion about this is right, and that it is *perceived* and not *actual* interests that lead the gambler's luck to be more salient than his ability (in the context when it is), then the case Greco intends to be illustrative of the way interests frame salience is at tension with what he wants to (and indeed should) say about the role actual versus perceived interests should play within a theory of knowledge. He writes:

It is more plausible that the truth-values of knowledge claims depend on the actual interests of some relevant group, than it is that they depend on the perceived interest of some individual. The most plausible interest-dependent theories will be formulated accordingly. (Greco 2006: 20)

If there is to be found in Greco's account of the connection between cognitive ability and cognitive success a principled way of understanding how purposes and interests dictate whether an agent's abilities would or would not be salient in explaining the agent's true belief, it will have to at the minimum provide a plausible

explanation for how perceived versus actual interests hook up with salience. Such an account would have to reconcile two competing intuitions, one concerning objectivity and the other concerning salience.

Objectivity intuition: The objectivity we frequently ascribe to knowledge would seem to require that, insofar as the salience of cognitive abilities in getting to the truth in Greco's account is determined by interests, it must be determined by actual and not merely perceived interests.

Salience intuition: In a wide range of contexts our perceived interests are not our actual interests, and in these contexts it is our perceived interests on which we base our judgments about what is salient in an explanation and what is not.

Lets turn now to whether either of Greco's two other suggestions can satisfactorily reconcile the this tension.

8.5.2 Default salience

If an argument can be given for why abilities will normally be salient in explaining true beliefs unless trumped by some abnormal feature of the explanation, then perhaps Greco could then defend, in light of the previous concerns, a further link between abnormalities and actual (rather than perceived) interests. Greco (2009a) makes a case along these lines. Drawing from Sosa's idea that dependence on cognitive character is of fundamental social importance, he supposes that there will be what he calls a *default salience*¹⁰⁴ of cognitive abilities as explaining agents' true beliefs. He claims that:

Unless something else is present that trumps the salience of *S*'s cognitive abilities, those will be the most important part of the story, or at least a very important part of the story. (Greco 2009, Ch. 6: 23)

Abnormalities, he thinks are, when part of the necessary set of causes of something, generally salient within corresponding explanatory stories; further, the salience of the abnormality trumps the default salience of

¹⁰⁴ Sosa (2007) has also employed the notion of 'default' assumptions in his own virtue epistemology, although Sosa's appeal on this score is less contentious. Sosa's idea is that for apt belief to be apt belief 'duly noted' (i.e. reflective knowledge) an agent has to have apt belief that his first-order belief is apt. Now this won't always require actually forming a positive belief about first-order aptness; rather, the second-order condition can be satisfied so long as you take yourself to have believed in normal conditions, and this is something for which we can (for example, in cases of perceptual knowledge) have a default disposition to affirm. Of course, sometimes, our first-order belief will be apt but our second-order belief won't, because the default disposition to accept normal belief-forming conditions fails. An example here would be the case of the kaleidoscope perceiver. But even though this default assumption fails in some cases, the default position itself is only a position that assumes normal conditions, and my point here is that this is a less contentious default position than Greco's, which takes as default that abilities will be most salient in explaining beliefs. Given his contextualist appeal, abilities will unavoidably be competing with a variety of other candidates for the role of salient explainer of beliefs. And so his appeal to contextualism makes his 'default' salience of abilities a more contentious claim than Sosa's use of a 'default' assumption; I mention this because it would be easy to assume mistakenly that both appeals would, on the matter of plausibility, stand or fall together.

S's cognitive abilities. (2009*a*, Ch. 6: 23) Setting aside whether abnormalities could be understood in terms of actual rather than perceived interests, this idea has initial promise when we consider, as Greco does, that Gettier cases are ones in which there is some 'odd twist': plausibly, it is the abnormality of the odd twist that trumps the default salience of the agent's abilities in getting to the truth.

That said, there stands a serious problem that lies internal to this suggestion. It is a mistake to infer from, as he says, the fundamental social importance of our cognitive abilities being dependable that these abilities deserve, in the space of explaining true beliefs, a position of default salience. Even if we grant that within the social practice of information sharing there exists a presupposition that the truths informants share with us are ones whose acquisition would be best explained by their abilities than anything else, it must be kept in mind that in addition to the true beliefs we have and, in our capacity as information sharers, would offer to others, we hold perhaps an equally large stock of true beliefs that the social practice of information sharing depends on us *not* offering to inquirers—that is, those true beliefs that for all we know, could easily be false.

Because our stock of true beliefs includes an array of both sorts, those most saliently explained by abilities and those most saliently explained by the world's happening to align with what we thought, it is incorrect to ignore the latter and assume that our abilities will be a default explanation for them all. This thought is entirely consistent with Greco's appeal to the social importance of cognitive abilities being dependable, a claim I don't dispute; what precisely is of social importance, though, is that the true beliefs we let loose into our social environments are ones we've gotten by our abilities and not ones that just happen to be true. And so, the prominent space in our collection of true beliefs filled by ones that, for all we know, could easily be false, makes illegitimate the assumption that our abilities would by default best explain for any given one why we have it.

8.5.3 Interests and purposes as indexed to a practical reasoning environment

A third proposal Greco (2006) offers for understanding how salience of abilities could be thought of as determined by contextually-based considerations takes as a starting point ideas advanced by Hawthorne (2004) and Craig (1990). The ideas here are, first *pace* Hawthorne, that knowledge is the norm of practical reasoning, a thesis that legitimizes what Greco takes to be the platitude that, all things equal, we should act only on what we know. And secondly, *pace* Craig, our interest in knowledge is tied importantly to our need to distinguish good

sources of information from bad ones by, as Craig puts it, seeking to ‘flag approved sources of information.’ (Craig 1990: 11) Putting these two ideas together, Greco takes the following to be plausible: that an:

...important function of our concept of knowledge and our knowledge language, perhaps its primary function, is to flag information and sources of information for use in practical reasoning. (Greco 2006: 22).

His idea is that, with regards to establishing explanatory salience, the relevant purposes and interests would sometimes be those in the subject’s practical reasoning environment, and other times, in the practical reasoner’s own practical reasoning environment. This motivates a specific way of thinking about the purposes and interests that his view proposes will frame the context within which we may assess whether abilities, or something else, stands out as salient. Greco’s conclusion:

These considerations suggest the following general rule: the truth-value of knowledge attributions (and the like) depends on the interests and purposes operative in the relevant practical reasoning environment. Sometimes this will be the practical environment of attributor, sometimes that of the subject, and sometimes that of some third party. (Greco 2006: 23)

This position constitutes a form of what Greco calls *interest-dependent, subject-sensitive* contextualism. (Greco 2006: 23)¹⁰⁵ In what follows I will defend an argument that will show why the specifics of his contextualist approach discussed here are in principle inept for the purposes of reconciling the tensions outlined in the previous section concerning objectivity, salience and actual versus perceived interests.

8.6. Knowledge and objectivity

Ever since the emergence of the very first contextualist proposals, marked by their attractive resources for preserving—almost too easily—Moorean instincts about everyday knowledge against the threats posed by closure-based skeptical arguments, epistemological contextualism has inspired worries that the shifting of epistemic standards that contextualists employ is at odds with some of the most fundamental assumptions we have about knowledge.

Now we’ve seen that Greco’s special kind of contextualism wouldn’t appear committed to radical shifts in epistemic standards, and this is for the reason that what directly depends on context in his view is not

¹⁰⁵ As a self-standing proposal, the position should be interpreted charitably in light of Greco’s admission that he offers it as merely suggestive. And insofar as it suggests a way to reconcile the rather compelling and distinct platitudes offered by Hawthorne and Craig, the proposal would seem to stand up quite well to competing contextualist views. I’ve argued that if there is to be found in Greco’s view a principled way of understanding how purposes and interests dictate what is salient and what is not, it will have to at the minimum provide a plausible explanation for how perceived versus actual interests hook up with salience. Greco’s suggestion that interests and purposes be linked to groups that make up practical reasoning environments, despite what might be its comparative merits among other contextualist approaches, doesn’t wear any obvious signs of whether it implies or could be compatible with such an explanation.

epistemic standards, *per se*, but the salience of the agent's abilities in acquiring true beliefs. A consequence Greco is comfortable with is that, as with the case of comparatively more radical versions of contextualism, the truth-value of knowledge claims is interest-dependent. That is, whether one knows for Greco is partially determined by the interests that frame the practical environment in which their abilities will be more or less salient in explaining their true beliefs.

Epistemologists have pointed out what he takes to be two central worries about interest-dependent accounts of knowledge insofar as they seem to threaten deeply-rooted assumptions about knowledge. First, they have been thought to, as Greco says, 'rob epistemology of its proper subject matter' by way of allowing 'know' to pick out different things in different contexts. (Greco 2006: 19) Secondly, such theories have been thought to rob knowledge of its objectivity.¹⁰⁶ With a commitment to perserving objectivity in mind, Greco writes:

...the most plausible interest-dependent theories will severely restrict the ways in which interests affect the truth-values of the moral and/or epistemic claims in question. (Greco 2006: 19).

Given the overarching thought of knowledge as importantly objective, this seems right. And correspondingly, a theory that depends on actual interests would be far more restrictive in this sense than a theory that depends on perceived interests which themselves could be based on the wildest of beliefs—beliefs of the sort that knowledge should in no way depend on. Against this background, I offer the following argument.

First Argument

1. In order to not rob knowledge of its objectivity, an interest-dependent theory must severely restrict the ways in which interests affect the truth-values of knowledge claims.
2. If an interest-dependent theory allows that the truth-value of knowledge claims depend on perceived interests, then the theory does not appropriately restrict the ways interests affect the truth values of knowledge claims.
3. Greco's proposal implies that the truth value of knowledge claims depends on perceived interests.
4. Therefore, Greco's proposal robs knowledge of its objectivity.

Greco positively endorses premises (1) and (2) of the argument. I turn attention now to what is the crucial premise, (3). Despite Greco's claim that his theory is meant to appeal to actual interests, it is in fact tied down to perceived interests. His commitments, I'll argue, will appeal to actual interests only when actual and

¹⁰⁶ Regardless of what features of a theory we think would qualify as proper robbers of objectivity, it seems fair to think that varying accounts of knowledge might be guilty of this crime (supposing it is one) to greater or lesser degrees. Even the most radical interest-dependent version of attributor contextualism would be, for example, more objective in nature than a theory claiming that knowledge can be defined however you like, and plausibly less objective than more traditional, invariantist approaches to knowledge.

perceived interests happen to align; otherwise, the direct appeal will be to perceived interests. The Gambler case suggested that it was perceived, rather than actual, interests that led the Gambler's luck to appear explanatorily salient in the context that it did. With this in mind, I want to move now to an argument about perceived interests and explanatory salience more generally.

8.7 Perceived Interests

Greco understands *perceived* interests in terms of what they are not. He describes them as 'opposed to...actual interests, whether or not perceived to be such.' (Greco 2006: 20) This distinction implies that, when our interests are merely perceived, it will be because we are mistaken about what our actual interests really are. Given the addictive tendencies and negative effects on overall well being associated with heroin use, it should be fair to say heroin use is opposed to our actual interests if anything is. This is so no matter what a naive junkie or his dealer might think. That said, consider the following case:

[DRUG ADDICTS]: Just as a terrible storm was coming through, Rodney and Damon get a call from their heroin dealer, Dave, informing them that the shipment is in. Fiending addicts, Rodney and Damon leave their flat in the middle of dinner, hop in the car and navigate through lightning and heavy rain to Dave's house. They return to find Jared, a fellow addict, standing in their flat looking shocked. 'What caused you to guys to leave in the middle of that storm?' he asks them. Rodney and Damon reply: 'We're a couple of drug addicts.' Jared, having been accustomed to shooting up with Rodney and Damon almost daily, is confused by the answer. 'Well yeah you're drug addicts, we all are. But why would you guys leave when you knew that storm was coming? Why didn't you just stay here and feed your addiction where it's dry, like you usually do?' Rodney and Damon then explain that Dave's shipment had come in, an explanation that immediately resolves Jared's confusion. 'Aha!' Jared says, no longer confused.

Now we would be right to imagine that in a context of doctors and drug counselors who have in mind the *actual* interests of the group—for example, to adopt a healthy, non-self-destructive lifestyle—what counts as salient in explaining their sudden departure in the storm is not that Dave's shipment came in, but *that they are addicts*. But this explanation, even though salient in the context of their actual interests, we saw was not salient in the context framed by the junkie's perceived interests—a context in which Jared's question is best answered by explaining that Dave's shipment came in (an explanation that, correspondingly, would be of little use to the doctors and counselors with the thought of health and well-being in mind).

Greco, then, is right to distinguish between actual and perceived interests as well as to think that a plausible interest-dependent account of knowledge must depend on the former and not the latter, but has failed to recognize that actual interests are not always those that determine what's salient in a given explanation. In light of this, I offer the following argument:

Second Argument:

1. When actual interests oppose perceived interests, what counts as salient in an explanation will align with perceived interests.
2. According to Greco, whether one knows will depend on whether one's abilities are salient in explaining her true belief.

3. Whether one's abilities are salient in explaining her true belief will depend, when actual interests oppose perceived interests, on perceived interests.

In combining the First Argument and Second Argument, we get the complete picture for why it is that a failure for his theory to preserve the objective nature of knowledge is unavoidable.

Combined Argument

1. In order to not rob knowledge of its objectivity, an interest-dependent theory must appropriately restrict the ways in which interests affect the truth-values of knowledge claims.
2. If an interest-dependent theory allows that the truth-values of knowledge claims depend on perceived interests, then the theory does not appropriately restrict the ways interests affect the truth values of knowledge claims.
3. According to Greco, whether one knows will depend on whether one's abilities are salient in explaining her true belief.
4. Whether one's abilities are salient in explaining her true belief will depend, when actual interests oppose perceived interests, on perceived interests.

5. Therefore, Greco's view robs knowledge of its objectivity.

We are now in position to see why Greco's appeal to the Hawthorne-Craig ideas—an appeal which qualifies the relevant interests at play as those operant in the relevant practical reasoning environment—would not be capable of reconciling the objectivity of knowledge claims with the subjectivity that marks the perceived interests to which explanatory salience is tied. At the end of the day, then, it seems as though understanding the connection between agency and true belief in a way that relies on salience to explain the role of cognitive abilities is a road that leads to two dead ends: either (i) what counts as salient is something we should 'just be able to tell,' a move licensing our brute instincts the role of final adjudicators in the court of knowledge

ascriptions; or (ii) what counts as salient would be established in a matter that fails to appropriately restrict the way that knowledge depends on interests.

8.8 Concluding Remarks

The conclusion drawn then is that Greco's appeal to explanatory salience within causal explanations for true beliefs was ultimately untenable. By holding knowledge hostage to explanatory salience, it was shown that we would also hold knowledge hostage to not only actual, but also perceived interests that inevitably encroach the bounds of what counts as salient in a given explanation—a consequence that compromises the sort of objectivity knowledge should be thought to have.

This fault of Greco's view suggests a presumption in favor of thinking that the relationship between cognitive success and cognitive ability should not be thought of in terms of purely causal explanations, although there is more to be said on this score. In the next chapter, however, I shall be considering a much deeper problem that I think is responsible for why Greco's view was ultimately untenable; the problem generalizes beyond just Greco's view, though it is explicable partly in terms of what Greco has overlooked about the relationship between knowledge and luck.

9

On the Problem of Epistemic Luck

9.1 Introduction

Although several chapters now have had as their collective focus problems arising out of an insight about the value of knowledge, it was argued from the beginning that there are *two* master intuitions that should guide a theory of knowledge. One was an intuition about the value of knowledge and the other an intuition about the relationship between knowledge and luck. We've traced out some of the far-reaching implications of the insight about the value of knowledge; an upshot is that we now know what a satisfactory response to the requirement arising from our value intuition would look like—that is, we know it must take the form of a theory of knowledge on which an agent's true belief is connected in some close way to her manifestation of epistemic virtue. This sort of connection, we saw, is one that a theory of knowledge must appeal to in accounting for the relevant insight that the epistemic value of knowledge exceeds that of true belief. What we learned about the value of knowledge, thus, told us something about what a satisfactory theory of knowledge will look like.

But to this point the connection knowledge requires between success and ability has not been developed in any form of sophistication, and consequently, we aren't yet in a position to say, for any given belief, whether it qualifies as knowledge on the 'proto'-theory of knowledge that the Meno Requirement has restricted us to. In order to develop our proto-theory into defensible theory of knowledge, it will be needed that we make precise the connection between truth and ability that is needed for knowledge. This is what the task of giving an analysis of knowledge requires of us.

By beginning our project with the value intuition and showing how we could account for it only with an analysis of knowledge on which there is a close link between truth and ability, we've proceeded so as to develop a theory of knowledge that is answerable to the value intuition. But an analysis of knowledge, we've said, must

also answer to the luck intuition. Naturally, then, the proto-theory that the value intuition has restricted us to is one that we'll need to appeal to the insight about luck in order to flesh out. The ultimate goal in place from the start has been an analysis of knowledge on which a belief *counts* as knowledge only if it satisfies conditions that will ensure that it is both (i) epistemically more valuable than a mere true belief (in the sense that the Meno Requirement has imposed); and (ii) such that it is not 'luckily' true, in the sense that, by investigating the luck insight, we'll take it to be requiring.

In the previous chapter, it was shown that we must not take Greco's lead and suppose that the connection between ability and truth essential to knowing could be rightly understood as a *causal explanation* relation. For one thing, we saw that this approach ran in to problems insofar as the notion of salience that is needed to specify the conditions under which it is *ability* that causally explains an agent's true belief is one that poses problems for the prospects of appropriately restricting how it is that knowledge depends on our interests. Given our task at present, it is worth noting another aspect of Greco's (2009*a*) approach that is troubling: his account of luck was parasitic upon his account of salience in a way that leaves it worryingly vacuous with regards to what epistemic luck actually is. A true belief is undermined by luck, on Greco's view, just in case luck rather than ability more saliently explains the true belief.¹⁰⁷

We need a better account of luck than this, and I'll argue that a place to start will be to highlight some features that Greco's own account of epistemic luck shares with some other recent accounts that have been defended. I'll then argue that these proposals share a common flaw.

9.2 Lack of Control Accounts of Luck

In having understood epistemic luck as what it is present when our coming to have a true belief is not a matter for which our cognitive abilities are an important or salient part of the relevant causal explanation, Greco reveals the sense in which he thinks epistemic luck should be understood in terms of what stands within our cognitive control, the thought his 1995 view made somewhat more explicit. After all, his present account implies that if the competences we exercise in forming a belief have little to do with why we got the truth, then

¹⁰⁷ His descriptions of what goes wrong in Gettier cases (which he takes to be paradigm cases where luck undermines knowledge) is simply that there is some 'abnormality' in the case that prevents the agent's abilities from being salient in causal explanation for the true belief. While his new view does not define luck independently from other notions (e.g. ability, salience) central to his definition of his knowledge, we can get perhaps some theory-independent insight into his conception of luck by looking at his 1995 view, on which a lucky event was just one that was 'out of our control.' Plausibly, he now takes the abnormalities that trump salience in explaining true belief as themselves indicating a relevant lack of control on the part of the agent. Perhaps this is why he thinks that when an agent has a belief that is true but not because of abilities, it is "luck" that he claims to be explaining the true belief more saliently than abilities; it is these cases where knowledge is thought by Greco to be *undermined* by luck.

our having the truth is a matter of luck. A similar idea can be found in Riggs's (2009) recent thinking, in which control is replaced by the related notion of responsibility. According to Riggs (2007): If an event E is lucky for S , then S was not sufficiently responsible for bringing about E . These proposals are grounded in essentially the same guiding idea. Jennifer Lackey has, in a recent paper, described the idea captured by Greco's and Riggs's accounts of luck—and which she attributes also to Stapp (1991) and Zimmerman (1993)—as a *lack of control account of luck*. Lackey captures the common thread of these accounts as follows:

Lack of Control Accounts of Luck (Lackey) (LCAL): An event is lucky for a given agent, S , if and only if the occurrence of such an event is beyond—or at least significantly beyond— S 's control. (Lackey 2006: 2)

LCAL stands as a good starting point for two reasons: first, it gets at one of Greco's core ideas but without the baggage that comes with his appeal to salience. Secondly, LCAL—in defining lucky events in terms of what is within our control—should strike us as especially suiting to the template account of knowledge as a true belief connected to ability that was motivated in previous chapters. After all, control seems to be related in some obvious way to ability—and so if the luck insight is understood in terms of abilities, then our account which must also understand *knowledge* in terms of abilities would seem uniquely tailored to this particular account of the sort of luck it must exclude.

But as a first point of order, consider that in order to answer to the luck intuition, it stands clear that we'll need to get straight on three key matters—matters which are ones Pritchard (2007*b*) takes to reflect 'three-stages' essential to the anti-luck project. He says:

First, we offer an account of luck. Second, we specify the sense in which knowledge is incompatible with luck. Finally, third, we put all this together to offer an anti-luck analysis of knowledge. (Pritchard 2007*b*: 2)

Though our own project is not a *robust* anti-luck project in the sense that we are not motivating an account of knowledge primarily out of the anti-luck insight—the project Pritchard associates with *anti-luck epistemology*—we no less want to require that our ability-based account block luck in the relevant way. Therefore, we'll need no less than those driven by the *robust anti-luck* project to engage in Pritchard's three dialectical tasks:

Pritchard's Three Anti-Luck Tasks

- (1) Give an account of luck
- (2) Specify the sense in which knowledge is incompatible with luck.
- (3) Show what conditions must be satisfied in order to block the kind of luck incompatible with knowledge and incorporate them within a theory of knowledge.

In the next section, I'll take to the first of these three tasks—a task that invites us to ask 'What is luck?' Let's begin by considering just how the *lack-of-control account of luck* we've introduced stands up to the challenge of answering this.

9.3. What does the LCAL account tell us about epistemic luck?

The LCAL account, as Lackey specifies it, is an account tailored to lucky events generally. If we extend this account of lucky events to specifically *epistemically* lucky events, then the *lack-of-control account* of *epistemic* luck will be:

[EPISTEMIC LUCK] (*Lack-of-control account*): S's true belief is lucky if and only if S's coming to have this true belief is beyond-or at least significantly beyond—S's control.

Lucky events do seem, intuitively, to be out of our control. The *lack-of-control* account (hereafter, LCAL) tries to accommodate this intuitive idea by supposing that lucky events can be *defined* in terms of control, and in a way that trades exclusively on this contrastive relationship. This might seem a promising strategy given the intuitive pull of the related ideas that: (i) to say that an event for me is lucky seems to carry with it the implication that it stood outside of my control; and (ii) to say that some event is within my control seems to carry with it the implication that it was not lucky.

But the promise here is specious insofar as we might think we can rely exhaustively on this two-way contrast in addressing Pritchard's first key task—the task that asks us to give an *account* of luck. To generalize this point: we shouldn't be duped into thinking that the fact that *X* events are ones we describe as 'not *Y*' events (and vice versa) would imply that *X* events are ones we can define correctly and informatively in terms of *Y*.

In seeing why LCAL goes wrong, it's important to get this point clear, and here I think the following example will be helpful. Consider that just as we explain lucky events by contrasting them with events that stand within our control, we might in a similar fashion explain (for example, to a curious child in a science class) explosions that are non-toxic by contrasting them with explosions that are harmful. Now though we might explain non-toxic explosions in a way that appeals to the contrast between what's non-toxic and what's harmful, we clearly *aren't* in doing so providing an account of 'non-toxic explosions' that is itself either informative or for

that matter correct. After all, the explanation of a non-toxic explosion we give the curious child¹⁰⁸ counts as an informative, correct definition of non-toxic explosions only if it can be supported with an *independent account of what's harmful* that itself neither (i) appeals to the notion of 'toxic' or (ii) fails to apply to all and only explosions that are non-toxic.

Because some explosions might be harmful but not toxic, this poses a problem: A definition of non-toxic explosions as those explosions that do not emit anything harmful fails to apply to the class of non-toxic explosions that emit things harmful but not-toxic (i.e. those that cause bricks to fly dangerously.) We could avoid this implication by simply defining harmful as 'toxic' and non-harmful as 'non-toxic'—but this would (along with other obvious problems) make the account of non-toxic explosions circular: Thus we see that, at the very least, definitions that trade exhaustively on intuitive contrasts seem to be on very shaky grounds insofar as we should want (as we want our account of luck to be) them to be informative and correct. With this cautionary note in mind, let's see if LCAL can be defended as an account of luck in a way that avoids the problems that we saw faced a similar attempt to extract a definition exclusively from an apparent contrast.

As a first point, we may say that LCAL accounts of luck can satisfactorily define luck in terms of its contrastive relationship to events that are within our control only if they avoid both sorts of objections we raised to the account of non-toxic explosions that traded exclusively on the contrast between what's non-toxic and what's harmful. What this means for LCAL is just this:

Requirement for (LCAL): The account LCAL offers of lucky events must be grounded in some *independent account of control* that itself:

- (i) Does not *itself* appeal to the notion of luck (*LCAL must be non-circular*)
- (ii) Is one that applies to all and only events that are lucky (*LCAL must not be open to obvious counterexamples*)

I think LCAL fails both these requirements. First, if an event's being lucky for *S* just means that it's not within *S*'s control, the conditions under which an event is within one's control must at least be ones we can specify without appealing to luck—a requirement for a non-circular definition. Here it might seem promising to give an account of control that appeals not to luck, but to *abilities*. On closer inspection, this move runs into trouble once we consider cases where control and ability come apart, and come apart in a way that could be 'brought together' only by invoking *luck* (something off limits to LCAL insofar as it wants to avoid a circular account of luck).

¹⁰⁸ *via* appealing to the connection between "toxic" and "harmful" and the related contrast between "non-toxic" and "non-harmful"

Take, for example, Pritchard's case of the *forcefield* archer: An archer (call him Archie) is surrounded by a dozen bullseye-targets. Unbeknownst to him, 11 of the twelve are surrounded by forcefields that would have blocked his arrow. Luckily for Archie, he chooses the one target out of the dozen without a forcefield around it and skillfully hits that bullseye. Was the event consisting his Archie's successful shot within his control? If we hold fixed what *was* in Archie's control—his intentions of hitting a bullseye and his ability to hit what he aims at—it's clear that he had little control over whether the success he had was brought about. Thus, we see how control and ability can come apart—and for LCAL, this counts against the prospects for supposing an account of events under our control could be provided squarely in terms of ability.¹⁰⁹

A proponent of LCAL might contest this point and note that the Archie case, in showing ability and control to come apart, does *not* show that the account of control that grounds LCAL's account of luck *cannot* be one that gives an account of control in terms of ability, but only that it cannot define control in terms of ability *exclusively*. And so it remains open that an account of control could appeal to ability *as well as to some other notion*—a notion that picks out what, in addition to ability, must be present in all cases in which events are within our control.

Here is where it becomes obvious why LCAL is doomed. As we've said, an account of luck in terms of control will be non-circular *only* if the account of control it provides doesn't *require* any appeal to luck. Problematically for LCAL, if events within one's control are defined in terms of ability in addition to some other notion needed to prevent control and ability from coming apart, it becomes quite obvious that luck is the needed 'other notion.' After all, the straightforward explanation for why Archie's success wasn't in his control was that, when we hold fixed all those things that *were* within his control (e.g. his ability to shoot accurately at that at which he's decided to aim), it was just a matter of *luck* that his success came about. Indeed, even if one were to try to argue that luck could be replaced with here with related notions like chance or low-probability, we could remind her that the precise thing LCAL is trying to give us an account of is the very thing that stands as the culprit for why control and ability come apart when they do in such cases. And so the point stands that LCAL will inevitably incur the charge of circularity if it defines lucky events in terms of events within our control, and then tries to give an account of *these* events in terms of ability.

This problem points to an even deeper one for LCAL: An account of luck in terms of control, as we can see, must be one that (i) cannot give an account of control in terms of luck; (ii) cannot give an account of

¹⁰⁹ We suggested this as an option for how LCAL might provide us an account of control in terms of which their account of lucky events would be plausible, informative and non-circular.

control in terms of *ability*, given that an account of control in terms of ability will be one that must appeal also to luck.

But what else is control if not something we have over events that stand in a certain balance with respect to what we have the ability to bring about and what remains left to luck? I contend here that there's *nothing* wrong (and indeed much right) about an account of control in terms of ability and luck—and so all the worse for the thought that LCAL could escape circularity.

An important point to gather from this, along with the fact that (as we now see) we can't give an account of luck in terms of control, is that as a concept, luck is more basic than control. Some understanding of luck is necessary in order to understand what it is for events to be in our *control*—something we grasp by understanding them (events within our control) as they lie in a certain balance between those intended outcomes that owe their manifestation to our abilities and those that don't.

Now, for the sake of argument, suppose that some form of LCAL could be non-circularly defended. Even if this were the case, then there is still no way LCAL could plausibly defend its answer to Pritchard's first question, and give us an account of luck. This is because, as we said, an account of luck must not only be non-circular, but it must avoid obvious counterexamples. That said, I want to highlight an easy way to generate them: just suppose that a golden-winged phoenix picks me up and carries me against my will to China, where I can't help but to believe (and rightly so) that the Great Wall is long. The event consisting in my coming to have this true belief was significantly beyond, if not *entirely* beyond, my control. LCAL answers the question 'What is epistemic luck?' in a way that implies that my belief about the wall of China being long is lucky. Intuitively, I still count as *knowing* the true belief I form with untrammelled vision next to a giant wall I have no problem detecting as large.¹¹⁰ Problematically, then, LCAL offers no principled way for saying what epistemic luck is that would allow us to answer the second question, which asks us to explain the conditions under which epistemic luck is incompatible with knowledge. On LCAL, some lucky true beliefs would count as knowledge and other lucky true beliefs would not. This is of no help. At the end of the day, the fact that LCAL can't be defended as an account of luck that is non-circular or for that matter free from obvious counterexamples invites us to look elsewhere for an account of luck.

9.4 Motivating an account of luck that doesn't appeal to control

¹¹⁰ *Ibid.* Lackey (2007) gives a similar sort of worry, though hers is directed to the sufficiency condition of LCAL as it applies to events generally rather than specifically epistemic events consisting in forming true beliefs.

It might be thought that, by abandoning LCAL, we've at the same time abandoned the prospects of providing what Pritchard (2009) has called a *robust virtue epistemology*—a term he applies to those projects aimed at preserving that 'knowledge is in some way a product of cognitive ability.' (Pritchard 2009: 45) ¹¹¹

Now Pritchard has his own reasons for thinking a *robust virtue epistemology* fails, but they lie outside the scope of this chapter (though we'll be considering some of these reasons in the next chapter). Presently, though, I maintain that we haven't *yet* disqualified ourselves from the resources of a robust VE simply by dumping LCAL. My argument here is just that it remains open to us to show how the connection between truth and ability needed for knowledge might ensure that the belief is not lucky in the way relevant for undermining knowledge. After all, we're still engaged with the task of giving an account of luck and specifying the sense in which it is incompatible with knowledge. As such, we don't need to concede just yet Pritchard's (2009) suggestion that our idea that knowledge arises from abilities cannot be accommodated in a way that ensures the belief won't be relevantly lucky, nor the related point that the insight that knowledge excludes luck can't be accommodated by a theory of knowledge in a way that ensures that non-lucky beliefs are connected in the right way to our abilities. We are at this point innocent with respect to commitments either way on these points.

The Modal Account of Luck

That said, let's move from LCAL to the most notable alternative account of luck in the literature, which is Pritchard's (2005). Following Lackey, let's call this the Modal Account of Luck (MAL), an account Lackey takes to have two conditions.

MODAL ACCOUNT OF LUCK (MAL):

- (1) If an event is lucky, then it is an event that occurs in the actual world but which does not occur in a wide class of the nearest possible worlds where the relevant initial conditions for that event are the same as they are in the actual world (Pritchard, 2005, p. 128).
- (2) If an event is lucky, then it is an event that is significant to the agent concerned (or would be significant, were the agent to be availed of the relevant facts (Pritchard 2005: 132).¹¹²

¹¹¹ Indeed, we've rejected and in doing so prevented ourselves from appealing to an account of luck that seemed very friendly to our guiding idea that knowledge must consist in some connection between truth and ability. And *if* our theory of knowledge can accommodate the luck intuition only by adding on a separate condition that blocks the relevant sort of luck in a way that the mere connection we want a knower's true belief to have to her ability *can't*, then indeed—knowledge would no longer be *essentially* understood in terms of the true belief and ability connection. And thus, we'll have given up on providing a *robust virtue epistemology*.

¹¹² This is Lackey's (2007) description of MAL.

MAL, unlike LCAL, does not bring with it worries of circularity, and so this is already a heavy strike in its favour. Also, it clearly gets right straightforward cases of lucky events, such as lottery wins. If my lone ticket wins a reasonably large fair lottery, it stands that were we to hold fixed the relevant initial conditions (i.e. my having bought a ticket in that lottery), then in most nearby worlds, I'll not be celebrating but lamenting my losing ticket. If MAL is right, then luck must essentially be a modal intuition—an intuition we have about events (significant to us) that could have easily been otherwise.

Pritchard adds the second condition to MAL to distinguish lucky events from events that are chancy but not lucky—e.g. ones that could easily have not occurred but about which no one cares.

Now, applied to the sort of events we're interested in—true beliefs—Pritchard's MAL account becomes:

LUCKY TRUE BELIEFS (LTB) *S's* true belief is lucky *iff* there is a wide class of possible worlds in which *S* continues to believe the target proposition, and the relevant initial conditions for the formation of that belief are the same as in the actual world, and yet the belief is false. (Pritchard 2007*a*: 3)

LTB easily gets the right result in Gettier cases; for example, Smith continues believing that the man who will get the job has 10 coins in his pocket in a wide class of worlds where he had more, fewer or no coins in his pocket. Also, LTB gets right cases where we take it that *environmental luck* undermines knowledge; for example, Henry continues believing what he sees is a barn in a wide class of possible worlds in which he's looking at a barn façade¹¹³—as Goldman's case goes.

Moreover, notice that LTB does not count as lucky the sort of true beliefs that depend on luck in a way that is compatible with knowledge. For example, suppose by dumb luck I slip and fall in such a way that I notice, while on the ground, that there is a dollar underneath the refrigerator. My true belief that there is a dollar under the refrigerator depends on luck in the sense that most nearby worlds are ones in which I don't acquire the evidence for believing this that I have in the actual world. However this sort of luck, which Pritchard calls *evidential luck*, is benign in the sense that knowledge is not undermined by it. LTB is not concerned with true beliefs lucky in this way; this is because LTB holds fixed the relevant initial conditions that give rise to the belief, and these will require holding fixed that you formed the belief about the dollar on the basis of the evidence you had from your vantage point on the floor, where you clearly spotted it. LTB, as Pritchard intends

¹¹³ Kvanvig (2008) "Critical Notice of Pritchard's Epistemic Luck" contests that this is obviously true; he thinks it would be only if the intentions and practicers of the barn erectors are held fixed. I'll be discussing his idea here in a later section of this chapter.

it, is meant to capture what he calls *veritic luck*—the sort at issue when we say that it is a matter of luck that the belief formed was true. Importantly then, *LTB* is not an account of the conditions under which it's lucky that we have true beliefs. Instead, it's an account of the conditions under which it's lucky that the beliefs we have are true. Pritchard quotes Unger (1968) as saying that this sort of luck 'comes between the man and the fact' (Pritchard 2005: 133-8). It's *veritic* luck then that, according to Pritchard, we should think of as incompatible with knowledge, and the sort that *LTB* is meant to capture as it would apply to true beliefs. (Pritchard 2005: 146)

If this much is right, then we've made headway past the task of just giving an account of luck; we'll have answered the questions associated with the first *two* steps of the anti-luck project in that we've said what luck was via *MAL* and specified the sort incompatible with knowledge in terms of the sort picked out by *LTB*—the sort that would make it lucky that a belief we have is true.

All that would be left, if *LTB* is correct, is to show how a theory of knowledge could 'block' the kind of luck that *LTB* specifies. Because *LTB* counts as lucky those beliefs that—to put the idea simply—could (holding fixed the relevant initial conditions) easily have been false, then already we can see that the third step of the anti-luck project would have our theory of knowledge include a symmetrical condition that requires known beliefs *couldn't* (holding fixed the relevant initial conditions) easily have been false.

This is precisely the strategy Pritchard (2005) pursued in his book *Epistemic Luck*. On Pritchard's (2005) account, the condition needed by a theory of knowledge to ensure that known beliefs couldn't easily have been false is a *safety condition*. This is the requirement that:

Safety: S's belief is safe *iff* in most near-by possible worlds in which S continues to form her belief about the target proposition in the same way as the actual world, her belief continues to be true. (Pritchard 2005)

The obvious symmetry between *safety* and *LTB* reveals an important point about the relationship between the second and third components of the anti-luck project when the account of luck is modal. That is, *once we've got* a modal account of the sort of luck incompatible with knowledge, the final step will naturally be to say that knowledge insulate itself from these modal properties. Thus, if knowledge-incompatible luck occurs when true beliefs could have easily been false, then knowledge will require that they lack this property: that they *couldn't* easily have been false, and so the corresponding modal condition on knowledge must ensure (as *safety* aims to) *this*.

Safety not only stands in symmetry with LTB, but it also is preferable to *sensitivity* as a modal condition. Sensitivity (Nozick 1981) states:

Sensitivity: *S*'s belief that *p* is sensitive iff, were *p* false, *S* would not have believed *p*.

It's now well known that sensitivity, unlike safety, runs into problems preserving closure principles. And as Sosa (2007) has recently pointed out, sensitivity also runs into problems with the sceptic. He says:

Safety is preferable to sensitivity as a requirement for knowledge. For one thing, it is in line with our commonsense presupposition that we are not currently envatted, nor even just misled. If one were misled, one's belief that one was not misled would remain, on the same basis. That simple belief hence fails the test of sensitivity. But it passes the test of safety, which requires only that not easily would it be held yet false. (Sosa 2007: 98)

Pritchard's safety-based anti-luck approach has been the most widely discussed to date, and unsurprisingly, there have been a variety of objections to it. Now that we've seen the close connection between the three tasks of an anti-luck project, it should be an obvious point that an objection at any of the three stages will also threaten an anti-luck account at its other two stages. Let's refer then to Pritchard's own approach, as a collective response to the challenges posed by each task, the *safety-based account of knowledge*—an account that relies on MAL at the first stage and LTB at the second stage. In the next section, we'll consider some important recent objections to the *safety-based account of knowledge* and in doing so, get clearer about the prospects a VE account has for preserving the anti-luck insight conceived of as a fundamentally modal intuition.

9.5 A Conundrum for Pritchard's Safety: The Lottery Problem and the Garbage Chute Case

Recall here Pritchard's *safety principle*, which he uses to support the idea that knowledge, in order to exclude knowledge-undermining luck, must be *safe*:

Safety Principle (SP): *S*'s belief is safe iff in most near-by possible worlds in which *S* continues to form her belief about the target proposition in the same way as the actual world, her belief continues to be true. (Pritchard 2007a: 6)

While SP counts as safe mundane beliefs we take ourselves to know (e.g. such beliefs are not ones about which we could have easily been wrong), it also accommodates our intuitions about standard Gettier cases and *environmental luck* (fake barn) cases—e.g. by ruling that the true beliefs at issue in these cases won't be *safe*. This is all good; however, Pritchard (2007a) has recognized that SP runs into a trouble when we consider what it has to say about lottery propositions—i.e. propositions such as 'My ticket is a loser'—which we hold on

the basis of excellent statistical grounds. SP counts these propositions as *known* even though, intuitively, we don't know such propositions because we don't *know* we'll lose the lottery. This result occurs for SP because it counts lottery propositions as safe: our belief that our ticket is a loser remains true in most nearby worlds, which are worlds in which we've lost the lottery.

How can SP avoid counting lottery propositions as known? Pritchard suspects that one way to do this would be to strengthen the account of safety so that lottery propositions would fail to count as safe. He supposes that such a strengthened account might be:

(SP*) S's belief is safe *iff* in nearly all (if not all) near-by possible worlds in which *S* continues to form her belief about the target proposition in the same way as in the actual world the belief continues to be true. (Pritchard 2007*a*: 6)

Presumably, since some nearby worlds are worlds where the ticket wins and the belief that it loses is false (even though most nearby worlds aren't like this), it's false on SP* that you know your ticket will lose (as this belief wouldn't be safe). Beefing up safety in this way seems like a mistake, though, because it leaves safety too strong a requirement to suppose it is satisfied in cases where we intuitively *do* know. Pritchard notes that a worrying case here is Sosa's (2000) 'rubbish chute' case. (Pritchard 2007*a*: 6)

On my way down to the elevator I release a trash bag down the chute from my high rise condo. Presumably, I know my bag will soon be in the basement. But what if, having been released, it still (incredibly) were not to arrive there? That presumably would be because it had been snagged somehow on the way down (an incredibly rare occurrence) or some such happenstance (Sosa 2000: 13).¹¹⁴

Although presumably Sosa's rubbish chute case is one where you have knowledge, it wouldn't count as such on the strengthened SP* because the relevant belief doesn't qualify as safe—some worlds very close to the actual world are ones where the bag gets snagged and the target belief is false.

We can see that the prospects of blocking knowledge-undermining luck by appealing to a safety requirement is doubly threatened by the combinatory challenge of the Lottery Problem and the Garbage Chute Case. Basically, it seems like a theory of knowledge requires a stronger safety requirement than SP in order to prevent lottery propositions from counting as known, but once strong enough to imply that lottery propositions aren't safe, then the notion of safety seems to be too strong to preserve that we have knowledge in garbage chute cases by failing to count *those* beliefs as safe. A catch-22.

¹¹⁴ *Ibid.* p. 6. Pritchard's quotation.

Now if this conundrum shows that the safety requirement for knowledge isn't plausible, then this calls into doubt not just the prospects that a theory of knowledge with a safety condition would have for effectively accommodate the luck intuition, but *also* whether our corollary account of the sort of luck incompatible with knowledge was right in the first place. For *if it was*, then a safety condition—defined in perfect symmetry to the knowledge-incompatible luck it's meant to block—should not allow for a disconnect (as appears to be the case in the Lottery Problem and Garbage Chute Cases) between beliefs that are safe and beliefs that are known.

Pritchard's (2007a) way out of this dilemma has been to revise his original account of safety (SP) to make it stronger than it originally was, not *as* strong as SP*, and yet weak enough to preserve that garbage-chute beliefs (appropriately understood) will be safe. Here's the revised account:

SP** S's belief is safe iff in most near-by possible worlds in which S continues to form her belief about the target proposition in the same way as in the actual world, and in all very close nearby possible worlds in which S continues to form her belief about the target proposition in the same way as in the actual world, the belief continues to be true. (Pritchard 2007a: 7)

Is this strong enough to prevent lottery propositions from qualifying as safe? Pritchard thinks so for the reason that some *very close* nearby worlds¹¹⁵ will be ones in which S's belief is false; his reasoning is that:

...the possible world in which I win the lottery is a world just like this one, where all that need be different is that a few coloured balls fall in a slightly different configuration. (Pritchard 2007a: 7)

Let's now see why SP** is also supposed to handle the garbage chute case. On this score, Pritchard thinks that it's important to see how very minor details of the case affect our intuitions. His idea is that, first, unless there is something on which the bag might easily snag to, then it's not clear that 'there are very many (if any) near-by possible worlds in which his belief is false.' (Pritchard 2007a: 7) And if that's so, then SP** gets the right result by counting the belief as safe. However, if we include within the details of the case that there is something in the chute on which the garbage bag 'is *almost* snagging on each time' (*Ibid.* 7), then SP** fails to count the belief as safe, though this Pritchard thinks would be the right verdict since, with these details in place, Sosa's belief *really could* easily be false—and it would not be intuitive to say that he *knows* it will arrive safely at the bottom.

¹¹⁵ As Kvanvig (2008a) has pointed out, Pritchard fails to give us any principled way to sort 'very close' nearby worlds from merely 'nearby worlds,' and as a result, he worries that the distinction is one that could be applied in an *ad hoc* way. I'll discuss this deeper motivations behind this objection later, but for now it will suffice to point out both the worry as it applies here and note Pritchard's reply: Pritchard denies that we should be expecting some especially refined formulae for ordering worlds; he takes it that intuitive notions about what worlds are closer than others will suffice for possible-world semantics to be useful to the project of clarifying modal notions. I'm inclined to agree with Pritchard here, although there will be more to say later about a more general worry that Kvanvig's objection raises.

If Pritchard is right to think that amending SP in the form of SP** gets around these problems, then—in order to keep the needed symmetry in tact—we need to also amend the corresponding account (LTB) specifying the conditions under which a belief will be lucky in a way incompatible with knowing. Thus, LTB will now read LTB**:

LTB**: S's true belief is lucky iff in most near-by possible worlds in which S continues to form her belief about the target proposition in the same way as in the actual world, and in *some very close* nearby possible worlds in which S continues to form her belief about the target proposition in the same way as in the actual world, the belief is false.

9.6 Kvanvig's Closure Objection

Kvanvig (2008*a*) raises several challenges to the thought that a theory of knowledge can rely on a safety condition such as Pritchard's in order to accommodate the luck intuition. Among these objections is the criticism that an account that places a safety condition on knowledge will fail to preserve closure. Siding here with the majority, I take it to be a strike against a theory of knowledge if it indeed must deny closure.

Kvanvig's argument from closure is as follows:

Kvanvig's Closure Argument

1. Safety doesn't get the right result in standard (Ginet) fake barn cases unless the practices and intentions of the Wisconsinites (i.e. the barn erectors) are held fixed.
 2. But if safety requires holding such practices and intentions fixed, then cases can be constructed where closure fails.
-
3. Therefore, safety gets the right result in standard fake barn cases only if it gives up closure.

Kvanvig defends (1) by arguing that if we *don't* hold fixed the intentions and practices of the Wisconsinites that (in the original Ginet barn façade case) led them to salt the valley with barn facades whilst leaving one real barn, then:

...there will be little reason for thinking that there are fake barns in any nearby worlds...and thus little reason to think that in a fair number of nearby worlds one's perceptual belief that the object is a barn is false. (Kvanvig 2008*a*: 6)

But *why* would there be little reason to suppose this? Here his point is an interesting one. Kvanvig says:

After all, it is exceedingly unlikely that the Wisconsinites' approach to boredom or the desire to engage in frivolity is directed precisely this way, and it could easily be that none of the other remotely likely combinations of intentions and practices would have any empirical results involving fake barns. (Kvanvig 2008*a*: 6)

The idea then is that without holding fixed whatever the Wisconsinites' intentions and practices were that culminated in their erecting barn facades in the actual world, we can't rely on the thought that nearby worlds are ones that include the barn facades in the vicinity of the real barn, and *thus*, we lose our explanation for why the 'there's a barn' belief in the actual world is not safe—i.e. we can no longer explain why nearby worlds are ones where the hero fails to know it.

If Kvanvig's defense of premise (1) is correct, and the barn-erectors' intentions and practices must be held fixed across worlds, then according to premise (2) an implication is that closure must be given up. His defense of premise (2) involves a thought example—one where we stipulate that the intentions and practices of the Wisconsinites in the actual world are ones where 'chance' is built into their practices. Kvanvig has us suppose that:

At each barn site, they flip a fair coin to decide what to do. If the coin lands heads, they replace the barn with a red fake barn. If the coin lands tails, they paint the barn green and leave it in place. (Kvanvig 2008*a*: 8)

Now, given that the odds that what you point to when saying 'there's a barn' *really is* a barn are about 50/50, it's clear that your belief 'there's a barn' isn't safe, and so your belief 'there's a barn' isn't one you know. And further, if your belief was instead 'there's a green barn' you wouldn't know that either; after all, you don't even know that it's a barn. Problematically, though, Kvanvig says:

...the 'green barn in the field' belief is safe because in every nearby world in which you form this belief, your belief is true. But then closure fails for safety, since you know that the green barn belief entails the barn belief. (2008*a*: 8)

Given that the Wisconsinites left only the real barns green and that we're holding this as part of what's fixed in nearby worlds, the presence of the red barn facades doesn't threaten the safety of the green barn belief because nearby worlds are ones where you believe something is a green barn only if it is. And then the problem is that that belief is safe (and counts as knowledge) while we know that it entails the 'there's a barn' belief, which was itself not safe and not knowledge. And so closure fails for safety because the safety account implies that you can know some p and know that p entails q while failing to know q .

Kvanvig claims that Pritchard's attempt to dodge this problem¹¹⁶ would work only if the requirement were lifted that the intentions and practices of the Wisconsinites be held fixed, a requirement we may lift only on the pain of failing to get the right result in the original fake barn case.

¹¹⁶ Kvanvig cites Pritchard as supposing that in a wide-class of nearby worlds, the Wisconsinites will fail to have painted some of the barns red and left some facades green, and so the belief 'there's a green barn' isn't safe given that the belief fails to be true in a number of close nearby worlds.

What this case may well show is that on an account of knowledge with a safety condition, knowledge fails to be closed under known entailment. However, Kvanvig suspects that a related closure principle that trades on *competent deduction* rather than *known entailment* could be preserved on a safety account. On the view that knowledge must be closed under *competent deduction*, the idea is:

(1) *Closure for knowledge (competent deduction)*: If S knows p and S competently deduces q from p , then S knows q .

As opposed to:

(2) *Closure for knowledge (known entailment)*: If S knows p and S knows that p entails q then S knows q .

Because beliefs reached by competent deduction are ‘paradigmatically safe’ (2008: 9) there will be no threat from the problem case to the view that knowledge is closed under competent deduction; you count as knowing the ‘there’s a barn’ belief only in instances in which you competently deduce it from the ‘green barn belief’—where as closure under known entailment counts you as knowing the barn belief so long as you know the green barn belief.

The extent to which this result is unwelcome depends on the importance of preserving closure for knowledge over known entailment. Given that the modal account of luck appears to be the most plausible (after all, the problems for control accounts seemed intractable) and that the most promising way to capture modally the sort of luck that’s incompatible with knowledge is one for which the symmetrical requirement for blocking it will be a safety condition, we might reason that a safety requirement for knowledge is motivated more strongly than the closure principle across known entailment that seems to be at odds with it. Thus, Kvanvig’s objection might incline us to think ‘all the worse for closure across known entailment’ as opposed to ‘all the worse for a safety condition on knowledge.’ However, this is the most optimistic reading. Others will take incompatibility with closure over known entailment to be a devastating blow against a safety condition on knowledge, even if safety can preserve a competent deduction closure principle. It seems then that, with a lack of any clear conception of the seriousness Kvanvig’s closure threat carries, we have no positive reason to jump ship because of it.

9.7 Becker on Luck and Reliability

If the luck relevant for undermining knowledge is as the present safety-based proposal claims, then the fact that we are restricted to an account on which knowledge is true belief connected in some way to ability is in no obvious trouble. As I've suggested earlier on in the chapter, it remains open for us to show how true belief is connected to ability in cases of knowing in a way that ensures such true beliefs will be relevantly safe. *Now*, that said, it *would* be a problem for us if the sort of luck that undermines knowledge could only be satisfied by some condition that requires we stray from our original true-belief-through-ability template. And this is precisely what Kelly Becker (2008) has recently suggested, and so it's needed that his point be addressed.

According to Becker, there are two strands of *veritic* epistemic luck, and therefore, an anti-luck account must have two separate components, one for dealing with each. (Becker 2008: §1) Becker, following Pritchard, takes *veritic* epistemic luck to be the sort of luck usually thought to undermine knowledge. Knowledge can be undermined, he notices, in some cases in which beliefs are true though produced by unreliable processes. For example, I might believe correctly what the weather is like by reading tea leaves. Though my belief is true, this process is unreliable, and it's just a matter of luck that my belief wasn't false. Contrast this case now with fake barn cases, where my true belief *is* the result of a reliable belief-forming process. According to Becker, the knowledge-undermining luck in the former case consists in the fact that an unreliable process produced a true belief. We block that luck by requiring for knowledge that beliefs be reliably produced; this will straightforwardly prevent beliefs that owe to this sort of luck from counting as knowledge. A reliability condition doesn't help the latter sort of veritic luck, though, and so this sort of luck requires a modal condition. (Becker 2007: §1)

If Becker is right, then we've overlooked something important—that is, that *two* kinds of luck are incompatible with knowledge.¹¹⁷ This would require that we no longer take LTB** to capture the essential features of lucky true belief incompatible with knowledge; even more, it would require that the true-belief-through-ability account be modified to include an additional process-reliabilism condition to deal with both kinds of knowledge-undermining luck.

I'll argue now that Becker is mistaken on both these points. In doing so, I want to first show that he's wrong to suspect that there is a unique strand of veritic luck that we must employ a reliabilist condition to handle. On this point I want to draw attention to his reasoning, which begins with the correct observation that we can make a distinction between two sorts of luckily true beliefs that fail to count as knowledge: some of

¹¹⁷ Becker goes on to argue for a non-veritic variety of epistemic luck that he thinks is incompatible with knowledge, but the example he used to defend this was not compelling enough to sufficiently motivate it.

them, such as the true belief reached via the method of reading tea leaves, were produced by unreliable belief forming processes, and others, such as true barn beliefs in fake barn country, are produced by reliable belief forming process. While this is true, it is a mistake to draw the inference Becker does that this distinction motivates the thought that there is a strand of epistemic luck—to which a distinct condition on knowledge must answer—that has anything to do with reliability. For one thing, we can expect that an evidentialist will be inclined to provide a competing diagnosis of the tea-leaf and fake-barn cases and could run a parallel argument about luck as follows:

An evidentialist parallel diagnosis: Whereas the tea-leaf case is one in which the agent's lucky true belief was one for which her belief failed to fit the evidence, the barn façade case is one in which the agent's lucky true belief was one that did fit the evidence. Therefore, (in line with Becker's reasoning), there are two kinds of luck incompatible with knowledge: luck that undermines knowledge when the true belief fits the evidence, and luck that undermines knowledge when the true belief doesn't fit the evidence. Thus, we need two anti-luck conditions to treat both kinds: we treat the first kind with an evidentialist condition, one that requires a belief fit the evidence. And then, because the second kind of knowledge-undermining luck occurs when true beliefs do fit the evidence, we need a modal condition to accommodate this latter sort.

Is there any reason to prefer Becker's story to the evidentialist's? There would be only if perhaps you, like Becker, already hold a reliabilist theory of knowledge to be true. (Becker defends a reliabilist account of knowledge in his 2007 book). But we should certainly not expect that an account of the sort of luck incompatible with knowledge is one whose specification requires an antecedent commitment to reliabilism.

We might ask, then, whether Becker's suggestion could be generalized in a way that is threatening to us. That is, we should ask whether the fact that some of the instances where knowledge undermines luck are ones where we're justified (in a wide sense) in holding the true belief we do and others aren't suggests that there must be two 'kinds' of knowledge-undermining luck, each which requires a separate condition? This would be so *only if* the modal account of luck we are using (LTB**) fails to deal with both kinds. On closer inspection, though, it seems that we can capture this luck in a way that requires no special consideration to how the belief was produced. LTB** simply tells us that *however* the belief was produced, this is something we have to hold fixed when we go from the actual world to nearby worlds and assess whether the belief continues to be true. If the belief was reliably produced, well supported by evidence, etc. or not, these are simply things that LTB** indiscriminately tells us we need to count as part of the relevant initial conditions that gave rise to the belief. The sort of luck that undermines knowledge then doesn't attach to distinctions within the initial conditions held

fixed. Luck undermines knowledge just when the conditions that gave rise to the belief, being what they are, could have easily yielded us a falsehood. Thus, we can dismiss Becker's challenge.

9.8 Lackey's Necessity and Sufficiency Objections

A final set of objections I want to consider to the modal account of luck I've been defending here alongside the thought that a safety condition would be best suited to handling it is one advanced recently by Jennifer Lackey (2008). Lackey, who it should be noted argued with equal force against the LCAL account of luck we dismissed in the first two sections, claims that that the modal account of luck we've taken to be clearly preferable to LCAL fails to specify conditions that are either necessary or sufficient for an event's being lucky.

If Lackey is right here, then this would be an unwelcome result to say the least. Indeed, if Lackey is right that luck can't be accounted for modally, then with control accounts and modal accounts equally off the table, it's not clear how the project of accommodating the anti-luck insight is supposed to begin—as both the obvious starting points will have reached dead ends. For Lackey, there is no third alternative, just the following somewhat depressing result:

Proponents of the LCAL and [modal account of luck]...are looking in the wrong places for capturing luck since both views propose conditions that are neither necessary nor sufficient for an event's being lucky. Hence, we see in the current philosophical literature what luck is *not*—it is *not* a matter of the absence of either control or counterfactual robustness. We have, unfortunately, yet to see what luck *is*. (Lackey 2008: 17)

We'll see now whether the arguments she gives to our modal account of luck actually warrant that we surrender it in exchange for Lackey's own defeatism.

Lackey on the necessity condition

The necessity condition at which Lackey takes aim is one that applies to the modal account of luck that we considered at the outset:¹¹⁸

MODAL ACCOUNT OF LUCK (MAL):

¹¹⁸ Lackey's intention of rejecting that the modal account of luck [MAL] captures conditions necessary for an event's being lucky should be distinguished from the recent projects of Comesana (2005) and Kelp (2007) who have rejected the idea that safety (which itself arises out of MAL) is a necessary condition for knowledge. While I've not been able to discuss each and every challenge to the modal account of luck and the safety condition it requires, a brief note concerning Comesana and Kelp—and in particular, why I didn't consider their particular challenges here—is for the reason that I think that Pritchard (2007) in his paper "Safety-Based Epistemology: Whither Now?" gives a satisfactory reply to Kelp's Frankfurt-style objection, and that this objection was a stronger one than Comesana's.

- (1) If an event is lucky, then it is an event that occurs in the actual world but which does not occur in a wide class of the nearest possible worlds where the relevant initial conditions for that event are the same as they are in the actual world (Pritchard, 2005: 128).
- (2) If an event is lucky, then it is an event that is significant to the agent concerned (or would be significant, were the agent to be availed of the relevant facts (Pritchard 2005: 132).¹¹⁹

Now, to be precise: Lackey's rejection of the necessity condition¹²⁰ for MAL requires that she defend that an event can be lucky whilst, at the same time, *not being such* that it occurs in the actual world but not in a wide class of the nearest possible worlds where the relevant initial conditions for that event are the same as they are in the actual world. She tries to do this with her BURIED TREASURE case:

BURIED TREASURE (Lackey): Sophie, knowing that she had very little time left to live, wanted to bury a chest filled with all her earthly treasures on the island she inhabited. As she walked around trying to determine the best site for proper burial, her central criteria were, first that a suitable location must be on the northwest corner of the island—where she had spent many of her fondest moments in life—and, second, that it had to be a spot where rose bushes could flourish—since these were her favorite flowers. As it happens, there was only one particular patch of land on the northwest corner of the island where the soil was rich enough for roses to thrive. Sophie, being excellent at detecting such soil, immediately located this patch of land and buried her treasure, along with seeds for future roses to bloom, in the one and only spot that fulfilled her two criteria.

One month later, Vincent, a distant neighbor of Sophie's, was driving in the northwest corner of the island—which was also his most beloved place to visit—and was looking for a place to plant a rose bush in memory of his mother who had died ten years earlier—since these were her favorite flowers. Being excellent at detecting the proper soil for rose bushes to thrive, he immediately located the same patch of land that Sophie had found one month earlier. As he began digging a hole for the bush, he was astonished to discover a buried treasure in the ground. (Lackey 2008: 11)

Lackey claims that Victor's discovery of the buried treasure is both (i) a lucky event, and (ii) one that fails to count as such on Pritchard's modal account of luck. Therefore, she claims, Pritchard's account of luck does not specify conditions necessary for an event's being lucky.

Let's grant that Victor's discovery is a lucky event and take issue with Lackey's thought that MAL fails to count it as such. Her reasoning is that:

Victor's discovering a buried treasure when he did is an event that not only occurs in the actual world, it also occurs in a wide class of the nearest possible worlds where the relevant initial conditions for such an event are the same as in the actual world. (Lackey 2008: 11)

Now, given that the original account of safety (SP) is one Pritchard discarded for his preferred account (SP**) which was able to deal with the conundrum raised by lottery propositions and the garbage chute case, we

¹¹⁹ Lackey's description of MAL is in her 2007 "What Luck is Not" p. 9.

¹²⁰ As a caveat, I should mention that I'm considering this objection now, rather than earlier, because I wanted to first motivate and defend the anti-luck project that arises out of MAL so that we have the right sort of perspective for assessing what's at stake if this account must be given up.

should note that this required us to amend the original account of the conditions under which luck undermines knowledge from LTB to LTB**. Plausibly, the original account of lucky events will need to be amended also so that it stands in symmetry with the others.

MAL** An event is lucky iff in most near-by possible worlds where the relevant initial conditions for the event are the same as in the actual world, and in *some very close* nearby possible worlds where the relevant initial conditions for the event are the same as in the actual world, the event doesn't occur.

If we hold fixed Victor's intentions and knowledge of rose-bush-thriving soil, and hold fixed as well Sophie's intentions and soil-detecting abilities, Lackey is right to suppose that Victor's discovery occurs in a wide class of nearest possible worlds. However, it remains true that there are some *very close* nearby worlds in which all of these things are held fixed and yet the event doesn't occur. This much will be enough for MAL** to count the event as lucky. In defense of this suggestion, just consider that even if we hold fixed all of the intentions and knowledge of soil that Lackey wants us to hold fixed, there are worlds *very close* to the actual world where, despite all that has been held fixed, Sophie buries her treasure in a slightly different spot just a few feet away—perhaps she's had a momentary lapse of concentration. These worlds, as well as ones in which Victor digs a foot or so in the other direction, or is preoccupied or in a hurry, as ones that are both (i) very close to the actual world; and (ii) not ones where he finds the treasure. And so, *even if* Victor finds the treasure in a wide class of possible worlds with the relevant initial conditions fixed, the event is nonetheless lucky so long as we can hold these same conditions fixed in very close nearby worlds where he fails to discover the treasure.

Lackey on the sufficiency condition

Lackey supposes that *whimsical* events that are of significance to us constitute counterexamples to the thought that the two conditions of the MAL (the first of which we're now replacing with MAL**) are jointly sufficient for an event's being lucky. For example, Lackey imagines a case where she, on a whim, decides to go to Paris, and that this turns out to be significant for her. She says that MAL counts this as a lucky event, although according to Lackey, it's not:

For even if my choosing to go to Paris for the weekend is based on a whim, I am still consciously choosing to perform this action and am, therefore, responsible for whatever consequences—either positive or negative—result from it. (Lackey 2008: 14)

Her thinking here reflects that she thinks that two features of the case count against the whimsical trip to Paris being lucky: first, the fact that it is a product of conscious choice, and secondly, that she is responsible for whatever consequences result from it. We can respond to both of these suggestions: Abelard presumably made a conscious choice to introduce himself to Heloise, though this fact shouldn't lead us to think that he was not lucky to have met the love of his life—the direct result of this conscious choice. Similarly, a lottery winner's decision to take the financial risk associated with playing is one by which she is responsible for 'whatever consequences—positive or negative—that result from (her decision to gamble)' (2008: 14) in a contest where she willingly incurs some risk for an opportunity for reward. After all, the bill for the ticket was mailed to her and so was the winning check. But this would not prevent her win from being lucky. Thus, Lackey's reasons for denying that whimsical events are lucky don't hold up. And on top of that, it's worth noting that we actually talk about whimsical events that turn out to be significant to us as lucky. For example, if I am soul seeking and just happen to grab a book that changes my outlook and gives me hope, I'd refer to the event of my having come across that book as lucky, even though I'll admit it to be a result of my own choice and responsibility.

So ultimately, then, Lackey's challenges to the necessity and sufficiency conditions of our modal account of luck fail to sufficiently cast doubt upon it.

9.9 Concluding Remarks

What we have now are working answers to the questions that motivated Pritchard's three anti-luck tasks.¹²¹ These fruits are bittersweet. While they constitute an important step toward our goal of accounting for the anti-

¹²¹ We've uncovered several important insights about luck. First, we saw that although a lack-of-control-account of luck (LCAL) would appear to be the most promising for the purposes of supplementing a robust VE account of knowledge, LCAL accounts were ultimately indefensible. The most promising strategy for LCAL was to give an account of control in terms of ability; however, we saw that ability and luck come apart in a way that required an account of control to specify some other notion in addition to ability—and that notion was shown to be luck. Thus, it was concluded that LCAL accounts are *inevitably* circular—control is essentially a function of both luck and ability—which means that luck cannot be defined in terms of control, as the latter *must* be defined partly in terms of the former. This intractable problem with LCAL motivated that we adopt a different notion of luck from which to engage in what we said were Pritchard's three anti-luck tasks: providing an account of luck, specifying the sense in which luck is incompatible with knowledge and finally developing a condition suitable for blocking this sort of luck. Toward the end of engaging in these three tasks, the modal account looked comparatively more promising than the LCAL account. Taking Pritchard's (2005) view as our starting point, we were able to answer the three tasks by way of what seemed the three most attractive options: (i) the Modal Account of Lucky Events (MAL) as a specification of the conditions under which an event is lucky, (ii) Pritchard's Lucky True Belief (LTB) account as a plausible way to articulate the conditions under which luck is incompatible with knowledge; and (iii) Pritchard's (2005) Safety Principle (SP) as a condition capable of blocking the sort of luck LTB specified to be incompatible with knowledge. As Pritchard's more recent work has shown us, though, SP ran into problems with lottery propositions and garbage chute propositions, which together implied a conundrum for SP in that it seemed to imply no formulation could be strong enough to prevent lottery propositions from counting as known and at the same time weak enough to count garbage chute cases as knowledge. Pritchard's way out of the conundrum we saw was to amend SP so that it required that no *very close* nearby worlds could be ones in which a safe belief is false. We subsequently revised Pritchard's account of LTB so as to be in symmetry with his new account of safety. What this implied was that knowledge undermining luck had to be changed so that lucky true beliefs (in the sense relevant to undermining knowledge) will be ones that are false in at

luck insight about knowledge, this step forward comes at the expense of inviting a tension that is perhaps obvious by now to the reader. In having concluded that Pritchard's modal account provides the strongest, most defensible approach to accommodating the anti-luck insight, we appear to be faced with two restrictions on an account of knowledge pulling us in different directions. Before proceeding to the task of the next chapter, I think it's important that we examine this tension carefully as doing so will reveal more clearly what stands to hang in the balance.

I think we can grasp the tension best in the context of the various commitments that gave rise to it: The opening chapters, in having clarified the value insight and what is needed to account for it, motivated the thought that this guiding insight about the value of knowledge could be preserved only on an account where knowledge is *essentially* understood in terms of a connection between true belief and ability. The value insight thus motivated some variety of *robust virtue epistemology*, and the thought was that in order to provide an analysis of knowledge in line with this requirement, we'd need to *specify* the precise way that true beliefs are connected to abilities in cases of knowledge.

Because we've taken it from the start that a theory of knowledge must answer to *both* the value and luck insights, the thought was that we could specify the connection between true belief and ability essential to knowledge only if we could show that this connection rules out luck—(a project Pritchard claims is not promising for a *robust VE*, but which is still open to us). Now our intention of getting clear about luck, the sense in which it's incompatible with knowledge and the way to block it out left us with clear ideas on each of these points, and thus, a clear idea about what the connection between true belief and ability will have to rule out. And on this point, the fact that we've learned that the sort of luck incompatible with knowledge is itself

least *some very close* nearby worlds. Having defended this picture taken to be the most attractive, we showed how Kvanvig's closure-objection failed to have any devastating consequences, partly because closure for knowledge under competent deduction does not fail on safety, and partly because the closure principle for knowledge over known entailment doesn't carry any obvious plausibility that the thought that our modal account of luck and its related safety requirement lack. We saw then that what might have seemed initially to be a devastating objection appeared to end in a stalemate. Further, we saw that Becker's attempt to undercut our theory of the sort of luck incompatible with knowledge by imposing dual criteria was itself based on the mistaken thought that different conditions for blocking epistemic luck would be needed to account for the fact that some non-known luckily true beliefs are reliably produced and others not. Finally, having by this point shown that following Pritchard's recent account of safety and knowledge-incompatible luck are not only much more plausible as an anti-luck approach than was LCAL but also defensible against numerous challenges, we considered Lackey's attempt to undercut our approach at its core, by arguing that the modal account of luck fails to specify conditions that are either necessary or sufficient for an event's being lucky. We saw that her BURIED TREASURE case seemed to threaten the necessity condition employed by MAL; we got around this problem though by bringing MAL into symmetry with our account of knowledge-incompatible luck and safety—that is, by requiring of lucky events that they not occur in *some very close worlds* where the relevant initial conditions leading to the event in the actual world are held fixed. Further, we saw that Lackey's challenge to the sufficiency condition was not successful in its appeal to whimsical events as ones she thinks are by mistake counted as lucky by MAL. We showed that her arguments for this point didn't hold up, and that she lacked any good case for supposing that such events—ones that could easily not have transpired yet did and had significant consequences for us—aren't lucky; thus, we showed that it was not a mistake that MAL counts them as lucky.

one we must represent modally suggests that our *robust VE* approach could defend an essential connection between true belief and ability that blocks the relevant kind of luck only if a modal safety condition can be somehow *built into* our specification of the conditions under which a true belief is connected to ability in a way sufficient for knowledge. If this cannot be shown, then a consequence would be that we must reject a *robust VE* approach in favor of a *modest VE* approach on which we accommodate the ability and luck insights by employing separate conditions for each: on such an account, we'd be surrendering the thought that an ability condition can do 'all the work.' We'd have to say instead that knowledge requires that a true belief be through ability and that, additionally, the true belief must be safe (whilst conceding that the satisfaction of the former does not imply the satisfaction of the latter). While some might opt for such a view, I contend that it should be resisted.

Firstly, if we abandon *robust VE* for *modest VE*, then we will have implied that the luck insight is more important than the value insight. After all, our sustained attention to the value insight showed that we might promisingly preserve it only on an account where the true belief and ability connection captures all that is essential to knowledge. If we give up on this, *we give up on what we saw to be the only type of account suitable to accounting for the value of knowledge*. Conversely, though, if we latched on stubbornly to the *robust VE* template in an effort to accommodate the value insight, while at the same time unable to show how the safety of beliefs is an implication of satisfying this connection, then we'll have *given up on what we saw to be the only type of account suitable to accounting for the insight that knowledge excludes luck*.

Highlighting this tension will set the scene for the next chapter; our goal is simply to preserve the insights about luck developed in this chapter in our articulation of the connection between truth and ability essential to knowledge. What hangs in the balance are our master intuitions, at least of which will be such that we'll fail to accommodate it if the stated goal is not achieved.

10

Knowledge in the Balance of Luck and Ability: Toward a Robust Virtue Epistemology

10.1 Introduction

It was argued that our initial insight about the value of knowledge motivated a theory of knowledge according to which knowledge is thought to be true belief that stands connected in the right way to intellectual virtue or ability.

What we saw in the previous chapter, though, was that the *luck insight* about knowledge turned out to be a modal insight and that this motivated a *modal condition* on a theory of knowledge. Having seen that the sense in which knowledge is incompatible with luck (i.e. expressed by LTB**) is one that a theory of knowledge best accommodates by requiring that known beliefs be *safe* (i.e. in the sense specified by SP**), we recognized that a theory of knowledge is subject to a *safety constraint*.

Safety constraint: Knowledge requires of true beliefs that they be appropriately *safe*.

The present view is thus one on which a true belief counts as knowledge only if it is (i) connected in the right way to intellectual virtue or ability; and (ii) appropriately safe. These features of our view bring it into close alignment with Pritchard's (2009) present view. I want to now explain some points of similarity and difference between Pritchard's variety of what he calls *anti-luck virtue epistemology* and the view I am developing here. A grasp of this will be useful for understanding why, despite certain shared commitments, the project I develop in this chapter will take on a notably different form than Pritchard's.

Anti-Luck Virtue Epistemology: What it is not

According to Pritchard (2009):

The general structure of the account of knowledge offered by an anti-luck virtue epistemology can now be described as follows: knowledge is safe belief that arises out of the reliable cognitive traits that make up one's cognitive character. (Pritchard 2009: 52)

By and large, this general description of *anti-luck virtue epistemology* (hereafter ALVE) suggests that our present project falls happily within such a description of ALVE. This is not surprising. After all, Pritchard thinks that a theory of knowledge must:

...incorporate two conditions on knowledge, an anti-luck condition and an ability condition.... (Pritchard 2009: 51)

To be precise, then, we share with Pritchard's view a commitment to thinking that a defensible account of knowledge must answer to *both* the anti-luck *and* ability insights. Consequently, it would be right to locate our present view somewhere under the heading of *anti-luck virtue-epistemology*, defined as follows:

Anti-luck virtue epistemology (ALVE): A theory of knowledge that places *conditions* on knowledge that arise from the intuitions that knowledge (i) arises from ability and (ii) excludes luck. An ALVE analysis of knowledge must count a true belief as known only if both conditions (i) and (ii) are satisfied.

Now, importantly, not all analyses of knowledge under the heading of ALVE are going to look alike. There is room for sharp disagreement about whether satisfying either one of the two conditions would imply that the other will be satisfied, and on this matter, quite different pictures can be developed. To be sure, a commitment to ALVE is only a commitment to offering an analysis of knowledge that upholds *that* conditions (i) and (ii) be satisfied; such a commitment is neutral with respect to the matter of whether either condition entails the other, or whether we should instead view them as independent conditions. Given this argument space, it's possible to disagree on the crucial matter of how to analyze knowledge whilst agreeing on the separate point of *whether* an analysis of knowledge must include both an ability and an anti-luck condition.

What this means is that different strands of ALVE will match up with the different possible ways to defend the matter of whether either condition entails the other.

It's clear to see how Pritchard's view lines up with one of these strands. He says:

For just as there is no formulation of the ability condition that can obviate the need for an anti-luck condition, so there is no formulation of the anti-luck condition that can obviate the need for an ability condition. (Pritchard 2008 ALVE: 51)

His view, then, is that neither the ability nor the anti-luck condition entails the other. It is worth noting here that Sosa (2007) holds the same position at least with respect to whether one sort of condition would entail the other. Sosa's ability condition, which he calls *aptness* stands independent to his anti-luck condition, which is a basis-relative safety condition. Neither entails the other. (Sosa 2007: 29) According to Sosa:

A performance is apt if, and only if, it is correct attributable to a competence exercised by the performer, in conditions appropriate for its exercise. By contrast, a performance is safe if, and only if, for some basis that it has, it would not easily have been incorrect if based on that basis. So defined, safety and aptness can come apart. (Sosa 2007: 82)

Let's call strands of ALVE on which neither condition is supposed to imply the other *Dual-condition ALVE*, to reflect a commitment to holding that these conditions come apart in both directions and are thus independent requirements. A generic formulation of *Dual-condition ALVE* will be:

Dual-condition ALVE: *S* knows that *p* if, and only if, *S*'s true belief that *p* is (i) appropriately connected to *S*'s ability; and (ii) safe.

On dual-condition ALVE, neither condition is redundant in the analysis, as neither implies the other. Dual-condition ALVE can be distinguished from the strands of ALVE that maintain that at least one condition implies the other; on these strands, at least one condition will be redundant within an analysis of knowledge, and so these strands will offer in that respect different analyses of knowledge from dual-condition ALVE. For one who defends that it is the ability condition that implies the anti-luck condition, but not vice versa, then the resulting analysis will be one we'll call *Ability Condition ALVE*:

Ability Condition ALVE: *S* knows that *p* if, and only if, *S*'s true belief that *p* is appropriately connected to *S*'s intellectual virtue or ability.*

*When appropriately connected, the belief will be safe.

According to Ability-Condition ALVE, you satisfy the ability condition only if you satisfy the anti-luck condition. An important point about Ability-Condition ALVE is that it does not imply that all safe true beliefs are known, even though it does imply that all true beliefs appropriately connected to ability are known. Ability-Condition ALVE, then, places an ability condition and an anti-luck condition on knowledge while denying that

safety is sufficient for knowing. In this sense, we can see that although the ability and anti-luck condition are afforded equal weight insofar as an analysis of knowledge requires they both be satisfied, Ability-Condition ALVE gives a more fundamental place within the analysis to the ability condition.

If one takes it that the anti-luck condition implies the ability condition, but not vice versa, then the resulting analysis will be one we may call Anti-Luck ALVE:

Anti-Luck ALVE: S knows that p if and only if S 's true belief that p is safe (e.g. satisfies SP**).

According to Anti-Luck ALVE, you satisfy the anti-luck condition only if you satisfy the ability condition. An important point about Anti-Luck ALVE is that it does not rule that all beliefs appropriately connected to ability are known, even though it does imply that all safe true beliefs are known. Anti-Luck ALVE, then, places an ability condition and an anti-luck condition on knowledge while denying that having a true belief appropriately connected to ability is sufficient for knowing. In this sense, we can see that although the ability and anti-luck condition are afforded equal weight insofar as an analysis of knowledge requires they both be satisfied, Anti-Luck ALVE gives a more fundamental place within the analysis to the anti-luck condition.

So which type of ALVE is correct? Dual-condition ALVE, Ability-Condition ALVE, or Anti-Luck ALVE? Each is saying something different about knowledge.

10.2 Choosing the best version of ALVE

10.2.1 Against Anti-Luck ALVE

Having then shown three distinct ways that a commitment to an ability and anti-luck condition on knowledge could be defended (and thus, three distinct kinds of ALVE accounts) I want to look more closely now at what was the third of them, *Anti-Luck ALVE*, and suggest why I think it is problematic.

Anti-Luck ALVE: S knows that p if, and only if, S 's true belief that p is safe (i.e. satisfies S**).

Now, as we said, Anti-Luck ALVE (hereafter, AL-ALVE) counts as a *bona fide* form of *anti-luck virtue epistemology* because both an ability and an anti-luck condition would be preserved by the analysis—that is, on

AL-ALVE, you count as knowing only if your belief is both safe and connected appropriately to ability. That much seems good.

However, an obvious objection to AL-ALVE surfaces once the question is raised as to why we should ever think that the ability condition is *really* implied by the anti-luck condition. What story could one tell to explain this? It's not clear. And without a compelling story here, we shouldn't just take the AL-ALVE defender at her word when she tells us that all safe true beliefs have the property of being connected the right way to our abilities.

Suppose, though, that the AL-ALVE defender could produce for us a compelling story (again, I can't imagine what it would be), sufficient for convincing us that the anti-luck condition implies the ability condition. Our first objection would be met, but serious problems will remain. One concerns the fact that AL-ALVE rejects the sufficiency of the ability condition for knowledge. Consequently, AL-ALVE allows that some true beliefs are appropriately connected to ability but not known. If this is allowed, AL-ALVE cannot address the Meno Requirement; we've said that the distinctive value of knowledge the Meno Requirement holds a theory of knowledge to account for is one that we can preserve only on an account of knowledge on which an essential connection between true belief and ability is what explains this distinctive value. Having rejected that this connection is sufficient for knowledge, AL-ALVE allows that true beliefs that fail to count as knowledge could nonetheless be just as valuable as knowledge. This is an unwelcome implication to say the least.

Even if, somehow, AL-ALVE could meet our first objection by coming up with a plausible story for how a belief's being safe implies that it is appropriately connected to abilities, and somehow managed to address the second objection concerning the Meno Requirement (how this would go, I'm not sure), there is third objection to AL-ALVE that still remains.

We might posit this objection by reflecting a bit on the anti-luck project. The third step—the step consisting in giving a *condition* on an analysis that blocks the sort of luck that was said in the second step to be incompatible with knowledge—was straightforwardly one of placing a *necessary* condition on knowledge. After all, the thought that luck can undermine knowledge motivated a condition on knowledge that would be satisfied only by true beliefs that weren't lucky in this way. And so, we think that in order for a true belief to be known, it is necessary (not sufficient) that it not be (veritically) epistemically lucky, and therefore, it is necessary (not sufficient) that it be *safe*. AL-ALVE therefore places an anti-luck and ability condition on knowledge in a way that requires safety to be not only necessary, but *sufficient* for knowing. Given that the anti-luck condition only motivated a necessary condition, the burden is on AL-ALVE to explain how it is sufficient, and without such

an explanation, AL-ALVE couldn't be maintained (and this is so *even if* AL-ALVE was to magically come up with answers to our first two objections where there seemed to be none). After considering these three objections, the problems they pose for AL-ALVE, and the fact that AL-ALVE has no promising way to avoid the consequences of the first two and no obvious way to defend against the third—the only fair conclusion will be to think that the viability of the *anti-luck virtue epistemology* could be maintained only if we can defend one of the two remaining versions: *dual-condition ALVE* and *single-condition (Ability) ALVE*.

10.2.2 Against Dual-Condition ALVE

Dual-condition ALVE: *S* knows that *p* if, and only if, *S*'s true belief that *p* is (i) appropriately connected to *S*'s ability; and (ii) safe.

Already, *dual-condition ALVE* (hereafter, DC-ALVE) is a major improvement over AL-ALVE. First off, DC-ALVE does not claim, as AL-ALVE did, that a safety condition implies the ability condition. Thus, DC-ALVE is not pained (as AL-ALVE was) to have to come up with an explanation for this. Secondly, DC-ALVE does not commit itself to the claim that safety is *sufficient* for knowledge, and thus, does not incur what was another problem of AL-ALVE.

DC-ALVE also fares, by comparison, better than AL-ALVE in its prospects for preserving the Meno Requirement, although this is a somewhat complicated matter and deserves I think some sustained attention. First, it should be noted that AL-ALVE's prospects on this score were indeed very bleak, moreso than was perhaps first indicated. Consider that AL-ALVE allowed that the connection between true belief and ability required for knowledge could be present in the absence of knowledge—and so insofar as we should think it is that connection that would account for the distinctive value of knowledge, we're in trouble when it is pointed out to us that beliefs AL-ALVE counts as not known would nonetheless count as just as valuable as if they were. AL-ALVE, then, was stuck in the unfortunate position of having to locate the source of the distinctive value of knowledge in some property knowledge always has, and for AL-ALVE, that was safety. But because we saw that the anti-luck intuition motivates safety only as a *necessary* but not a *sufficient* condition on knowledge, AL-ALVE has left unmotivated that the idea that safety really is sufficient for knowledge—something it would have to demonstrate (by appealing no less to something other than *just* the intuition that knowledge excludes

luck) *before* arguing further that the distinctive value of knowledge the Meno Requirement demands a theory of knowledge account for can be accounted for by safety alone.

It doesn't take a whole lot, we can now see, to have better prospects for addressing the Meno Requirement than AL-ALVE—its prospects were especially bad. DC-ALVE unsurprisingly fares better. For one thing, it is not burdened to locate the distinctive value of knowledge the Meno Requirement requires a theory account for *exclusively* in the modal properties of the belief, as AL-ALVE was. DC-ALVE, instead, will imply that the connection between true belief and ability that stands when true beliefs are safe can be at least part of the explanation for why knowledge is distinctively valuable (indeed, it *must* be part of such an explanation). And so, what we've argued for earlier—that we'll need to account for the Meno Requirement on some view on which knowledge requires some connection between true belief and ability—aligns better with DC-ALVE because at least DC-ALVE must specify this connection as *part* of what accounts for the distinctive value of knowledge.

That said, the overall prospects for DC-ALVE are nonetheless not promising for satisfying the Meno Requirement, despite the advantages it has over AL-ALVE. For one thing, DC-ALVE—like AL-ALVE—allows for the possibility that true beliefs appropriately connected to ability are not known (a consequence of leaving the ability condition *necessary* but not *sufficient*). Thus, DC-ALVE faces a problem very similar to its less-promising competitor: the problem of having to admit that, if the distinctive value of knowledge is located in the connection between truth and ability, then some true beliefs that are not known will be just as valuable as they would be if known. Whereas AL-ALVE could sidestep this worry (i.e. the worry that it commits itself to counting un-known true beliefs as ones that would become no more valuable if known) only if it could explain how the distinctive value of knowledge is located exhaustively in its modal properties (a very bad outcome), DC-ALVE can sidestep the same worry only by explaining how the distinctive value of knowledge is accounted for just when *both* conditions are met, though not when either separately is met. DC-ALVE thus has to scrap the final value explanation that we showed could be defended on a view where a connection between true belief and ability captures the essential features of knowledge.

Pritchard (2009) supposes this isn't necessarily a bad thing. For one thing, recall that Pritchard rejects that knowledge is finally valuable, and so for him it's not a problem for a theory of knowledge that it fails to account for the *final value* knowledge has. He says:

Robust virtue epistemology is...not a viable theory of knowledge, and this means that the response to the value problem [i.e. the *final value* response] that it advertises is not viable either. (Pritchard 2009: 47)

So even though Pritchard maintains that cognitive achievements are finally valuable—for reasons similar to the ones I gave in defence of this claim in Chapter 7—he rejects the *final value* solution because he rejects the *robust virtue epistemology* projects on which he takes it that that solution could be defended. He rejects *those* projects—ones that suppose that an ability condition captures all that is essential to knowledge—on two central grounds, which he characterizes in what he calls the ‘Jenny case’ and the ‘Barney case.’ Thus, whether Pritchard’s DC-ALVE is worse off for implying that the *final value* solution must be taken off the table is something we may assess in part by considering whether the ‘Jenny case’ and the ‘Barney case’ actually succeed in undermining *robust virtue epistemology* by showing knowledge and cognitive achievements to come apart.

Fortunately, half of our work is already done. Pritchard’s ‘Jenny’ case is a variation on Jennifer Lackey’s testimony example, and Chapter 8 was (entirely) devoted to arguing why this sort of case fails to show that knowledge and cognitive achievements come apart. The Lackey case was one that, recall, threatened the idea that cognitive achievements are *necessary* for knowledge. Although we saw that her example failed to demonstrate this, Pritchard’s ‘Barney’ case poses a different challenge—one to the idea that cognitive achievements are sufficient for knowledge. Though the ‘Barney’ case has not received near the attention that the ‘Jenny’-style cases have, it would—if successful—no less show a way that knowledge and cognitive achievements come apart, and so would carry the same implications threatened by the Jenny case.

The Barney case is structurally simple, much like a barn façade case¹²². It is one where Pritchard says:

...the agent [Barney] fails to have knowledge because his belief is subject to knowledge-undermining epistemic luck but where, nonetheless (and unlike Gettier-style cases), the agent does exhibit a cognitive achievement (i.e., the agent’s true belief is primarily creditable to his cognitive abilities). (Pritchard 2009: 47)

Now, a point of clarification: Pritchard does *not* think that the sort of epistemic luck that undermines knowledge in Gettier cases—which Pritchard calls *intervening* luck—is compatible with cognitive achievements. This seems right; after all, we don’t credit Smith for the truth of his belief about who gets the job. *Environmental luck*, such as the luck that undermines knowledge in barn façade cases, is thought by Pritchard to be *compatible* with cognitive achievements, a view that is shared by Sosa (2007) *per* his discussion of the kaleidoscope perceiver. (Sosa 2007: 100-101)¹²³

¹²² Pritchard’s sheep in the field case here is relevantly like a barn façade case, and so I’ll be discussing it in terms of the latter as the relevant details of such cases are already familiar.

¹²³ Sosa’s kaleidoscope perceiver has, according to Sosa, an apt belief (correct due to a manifestation of competence) despite the fact that the belief could have easily been false. His view is that animal knowledge, which only requires apt belief, can persist in such circumstances, whilst reflective knowledge, which is apt belief aptly noted (or apt belief in one’s first-order apt belief) is undermined.

This suggestion has admittedly a lot going for it. After all, if we consider the case discussed in the previous chapter of the archer who shoots skillfully and successfully at a target which is itself (unbeknownst to the archer) the only one not surrounded by a forcefield that would have deflected his arrow, it seems right to call what he's done an achievement, even though the environmental luck present in the situation undermines the safety of his success. And further, to Pritchard's credit, it seems right to credit Barney and the archer for their lucky successes even though we're *not* inclined to credit Smith for his lucky success. These considerations indeed suggest that environmental luck *is* compatible with cognitive achievements whereas intervening, Gettier-style luck is not.

Despite all of this, though, I think this assessment is mistaken, and that this can be shown in a such way that will allow us to diagnose and then explain away Pritchard's insight that we're comparatively more inclined to bestow credit to the agent for her skillful successes in environmental luck cases than we are in cases where successes are subject to intervening luck.

Cognitive achievement, cognitive success environmental luck: a closer look

My argument here turns on a distinction about true beliefs that was discussed in the previous chapter not in the context of achievements but in the context of discussing the sense in which knowledge requires that true beliefs can't be lucky. Recall here the case where I by dumb luck slipped and fell face-first onto the kitchen floor, where I found myself staring underneath the refrigerator and spotting quite clearly a dollar bill that, had I not by chance slipped, I'd never have believed was there. My true belief depended on luck (indeed, I'd not have had that true belief if not for the dumb luck that I'd fallen right there) but this luck was compatible with knowledge, and in a way that showed it to be irrelevant to the matter of whether one knows.

This invited a distinction: luck undermines knowledge *not* when it's a matter of luck that you have a true belief (as was the case with the dollar under the refrigerator) but when it's a matter of luck *that the belief you have is true*. This distinction, I shall argue, is relevant not just in the context of epistemic luck, but *also in the context of cognitive achievement*.

To see why this is, let's simply place the claim that knowledge requires cognitive achievement side by side the claim that knowledge excludes luck. A natural way to recast these ideas permits us to do so by asking: what is required of a *true belief* in order that it be known? Intuitively, we'll say: (1) It must not be a matter of luck that the belief one has is true; (2) It must be of credit to the knower that the belief she has is true.

Now, we saw whether it's a matter of luck *that one has a true belief* (e.g. the 'finding a dollar' case) was irrelevant to whether one knows. The relevant point was what (1) says: it must not be a matter of luck *that the belief one has is true*. Plausibly, we should say in the context of cognitive achievements that it must be to a knower's credit not simply *that she has a true belief* (a matter the luck discussion showed to be irrelevant to whether one knows) but that it be to her credit *that the belief she had was true*.

This subtle distinction surfaces in Sosa's (2007) discussion of testimonial knowledge and credit. He says:

The error resides... in supposing that what explains there being my true belief is what explains why my belief is true. (Sosa 2007: 96)

Sosa (2007) hits upon this idea also in an earlier discussion about apt belief.¹²⁴

What must be attributable to the competence is not just the belief's existence but its correctness. (Sosa 2007: 33)

If we are right to make use of this distinction, then the cognitive achievements thought to be of final value, and which consist in an agent's cognitive success through her cognitive ability are ones that require that we think of the success condition—the true belief—in the refined way that we think about true belief in the sense in which it's relevant to knowledge whether it's lucky. If what is relevant to knowledge is whether it's a matter of luck that the belief you formed was true, we're saying now equally that relevant to knowledge is whether it's of credit to your abilities that the belief you formed was true.

Bearing in mind this distinction, we are in a position to show how the Barney case fails to demonstrate that knowledge and cognitive achievements come apart, and in a way that explains away Pritchard's inclination to think that cognitive achievements are compatible with environmental luck but not intervening luck.

To this end, let's revisit the case of Archie: here Pritchard's assessment is that the environmental luck constituted by the forcefields renders Archie's skillful shot unsafe (e.g. he could have so easily shot at one of the targets that would have deflected his arrow) though nonetheless an achievement. Now, what made his shot unsafe was that it was a matter of luck *that the shot he had was a success*. While it's fine and well to give Archie credit for *successfully shooting the target*, is it to his credit *that the shot he fired was successful*? An intuitive way to assess this is to ask whether it is to Archie's credit that the shot he fired hit the good target rather than any of the other targets that would deflected his arrow? Clearly not. That was a matter of luck.

¹²⁴ This discussion is found in Sosa's (2007) second lecture, "A Virtue Epistemology."

And the same goes for Barney: it is to his credit that he has a true belief that there's a barn, but it is not to his credit that the belief that he has is true—after all, it's not to his credit that he formed his belief about the real barn rather than any of the barn facades. That was a matter of luck.

Now, let me say why Pritchard's thought that environmental luck is compatible with cognitive achievements seemed intuitive. This is because even though it was not to Archie's or Barney's credit that their arching and doxastic efforts were successful, respectively, we *are* inclined to credit Archie for successfully shooting the target and Barney for his true belief that there was a barn. This is more than we can say for Gettier's heroes: we do not say that it is to Smith's credit that the belief he had was true, *but neither do we say that he deserves credit for having a true belief!* After all, unlike Barney, Smith's abilities were entirely disconnected from his success.

We can thus explain away our inclination to think that cases of true beliefs subject to intervening luck are not cognitive achievements while cases of true beliefs subject to environmental luck are. Simply, only in cases of environmental luck will we be inclined to credit *her having a true belief* to the agent. However, as we saw, the kind of cognitive achievement *knowledge is concerned with* is an achievement whose success (i.e. true belief) will be specified differently so as to capture the sense in which the achievement is relevant to knowledge. Relevant to knowledge is whether it's of credit to you that the belief you formed was true, not whether it's of credit to you that you formed a true belief. The latter may well be a cognitive achievement, but it would be the sort of cognitive achievement that comes apart with knowledge, whereas the sort I've specified is *actually* relevant to knowledge wouldn't.

The prospects of DC-ALVE

As it turns out, it *is* a problem for DC-ALVE that the final value solution is off limits to it; given that the arguments against the necessity and sufficiency of the cognitive achievement thesis did not show knowledge and cognitive achievements to come apart, the final value solution is not only *not* viable, but decidedly the most attractive option—it matched the demands of the Meno Requirement and is uniquely suited to a theory of knowledge as true belief connected to ability.¹²⁵

¹²⁵ Now, *even* though the *final value* solution is off the table for DC-ALVE, there's a further problem for DC-ALVE when we consider how it could account for the value of knowledge given the resources it has. Pritchard suggests, *via* an appeal to Craig's (1990), that his DC-ALVE view accommodates the Primary and Secondary Value Problems given that—as Craig argues—“the central importance of the concept of knowledge resides in the practical need to pick-out reliable informants—informants that one can rely on.” (Craig 1990) According to Pritchard, the distinction between informants who are reliable and ones we can *rely* on motivates the thought that it is with both anti-luck and

Now given that the view denies that either condition is sufficient in itself for knowledge, and (according to Pritchard), the ability condition is not *strictly* necessary (e.g. given his claim in the ‘Jenny’ case), the dual-condition view would lack the resources for accommodating what Chapter 1 clarified to be an *unrestricted* insight about the value of knowledge—an insight that calls for a diagnosis that should explain the value of knowledge in all instances. Though Pritchard denies that the insight is unrestricted (at least in the way he characterizes the three value problems), we can set that aside for the moment and simply make the distinction argued for in Chapter 4 between knowledge and mere true belief as *types* and *tokens*. Knowledge, as a *type* of epistemic standing, was said to have whatever value it has over mere true belief in virtue of essential properties—thus, an explanation here must be one that does not appeal to value knowledge has in virtue of meeting some conditions that aren’t essential to all tokens of the type. And so the DC-ALVE view seems to lack an explanation for the value of knowledge as a *type* of epistemic standing over true belief. This was a consequence of defining knowledge in terms of two conditions that themselves are thought of as disconnected from each other.¹²⁶

10.2.3 Ability-Condition ALVE

Despite these worries for DC-ALVE, it would be the only viable version of ALVE unless one of the single-condition versions could be defended. Anti-luck ALVE (AL-ALVE) was untenable for several reasons—one of those reasons concerned the obvious point that a safety condition seems incapable of implying the ability condition. From this, we may gather that Ability-Condition ALVE would be defensible only if the connection between ability and truth could be defended as one that ensures the safety of the corresponding belief.

An interesting point to note is that Ability-Condition ALVE falls within two classification types: it is an instance of *anti-luck virtue epistemology* because the ability intuition and the anti-luck condition are *implied by what the analysis requires* for knowledge. Additionally, it is an instance of *robust virtue epistemology* for the reason that it

ability intuitions that the concept of knowledge would be grounded in as a way to pick out informants we can rely on. If he’s right about this explanation of the “geneology of the concept of knowledge” then this counts in favor of *some version* of ALVE, but not in favour of DC-ALVE over either of the single-condition versions. Even if it did count specifically in favour of DC-ALVE, there would still be no reason to think that DC-ALVE could account for the *epistemic* value we take knowledge to have over mere true belief, and this (I argued) is what was relevant to the Meno Requirement. Says Pritchard: “For it seems that knowledge...marks-out a distinctive epistemic standing which is of particular instrumental value to us. We would thus expect knowledge to be of more instrumental value than that which falls short of knowledge, even though knowledge is not finally valuable.”

¹²⁶ Plausibly, we should expect a theory of knowledge—whether reductive or non-reductive—to sufficiently illuminate the *nature* of knowledge. Here, we’ll see that *single-condition* accounts have a distinct advantage over *dual-condition* accounts. The idea here concerns what I think is plausibly needed in order for analysis to sufficiently illuminate the nature of that which the analysis concerns: the requirement is that the analysis not only specify correct necessary and sufficient conditions but also be one that we can support with a description of what the state of affairs would be like when the conditions hold. The DC-ALVE view gives us no reason to think that the ability condition has anything to do with the anti-luck condition, and so invites only a bifurcated illumination of what is being analyzed.

takes the connection between true belief and ability to be ‘doing all the needed work.’ Part of the ‘needed work’ is blocking the sort of luck that undermines knowledge, and if the connection between true belief and ability can be shown to do this, then *robust virtue epistemology* would be defensible as a variety of anti-luck virtue epistemology.

To make matters simple, I’ll be referring to Ability-Condition ALVE under the more simple and equally apt description of *robust virtue epistemology*:

Robust VE (RVE): S knows that p if and only if S ’s true belief that p is appropriately connected to S ’s intellectual abilities.

In what follows, my aim will be to show that RVE can be successfully defended as a theory of knowledge. I’ll show that it avoids the problems that faced the other two varieties of ALVE and also has a lot of attractive features the others lack. Of course, in order to develop a defense of RVE, I’ll need to give an account of the connection between true belief and ability essential to knowledge on which safety is ensured by the satisfaction of this connection. If safety is not ensured by this connection, then it looks like Pritchard’s dual-condition ALVE would be the only remaining option for defending ALVE, and this, as we saw, does not bode well for the prospects of accounting for the distinctive value of knowledge.

10.3 Motivating a robust virtue epistemology

A note is in order to clarify the relationship between (i) the connection between true belief and ability essential to knowledge; (ii) cognitive achievements. Recall from our discussion in Chapter 9 that cognitive achievements consist in *cognitive success through cognitive ability*. We can appeal to the *final value* solution to accommodate the Meno Requirement only if knowledge and cognitive achievements do not come apart. And so therefore, it’s going to follow that the connection between true belief and ability essential to knowledge must be one on which the true belief is produced, in some sense, *through* ability. This gives us a starting point for formulating RVE:

RVE(1): S knows that p if, and only if, S ’s true belief that p arises *through* S ’s cognitive ability.

Now, because the final value associated with cognitive achievements is one that we can appeal to in responding to the Meno Requirement only if cognitive achievements and knowledge *don't come apart* (e.g. otherwise, beliefs not-known could be just as valuable as they would be if known), we can see that RVE(1) implies RVE(2):

RVE(2): *S* knows that *p* if, and only if, *S*'s true belief that *p* constitutes a cognitive achievement.

Further, because the connection between true belief and ability is supposed to ensure the belief's safety, (RVE2) implies (RVE3):

RVE3: If *S*'s true belief that *p* constitutes a cognitive achievement, then *p* is safe.

It's easy to make the mistake of thinking that RVE3 commits us to something different (and far less plausible) than it actually does. On this I want to take care to be thorough, and so let's consider more closely the nature of cognitive achievement as this will help to make more precise what RVE3 requires us to defend.

Borrowing from Greco *ala* Feinberg's account of moral credit, we may recall an important point about achievements generally.

Achievements: *S*'s success constitutes an achievement for *S* if, and only if, *S*'s success is creditable to *S*'s abilities.

Extended to specifically *cognitive achievements*, the achievement thesis tells us the following:

Cognitive achievements: *S*'s cognitive success constitutes a cognitive achievement for *S* if, and only if, *S*'s cognitive success is creditable to *S*'s cognitive abilities.

Let's consider now the cognitive success at issue when one believes truly and begin by thinking about the cognitive success one enjoys when he is the victim of what Pritchard calls *intervening* luck, as in standard Gettier cases. In such cases, for example, Smith's knowledge is undermined by luck in a way we saw would lead us to deny that there is any sense in which the cognitive success that is *his believing truly that the man who will get the job*

has ten coins in his pocket is creditable to his abilities. After all, that he was cognitively successful (i.e. had a true belief) here had nothing much to do with his abilities.

We saw in the previous section that things are different in cases where *environmental luck* undermines knowledge, as it did in Pritchard's 'Barney' case. Unlike the case of Gettier's Smith, the Barney case is one where it would be wrong to say that Barney's success had nothing to do with his abilities. This, I argued, might explain why Pritchard thinks that cognitive achievements and knowledge come apart: when true beliefs fail to count as knowledge because environmental luck renders them unsafe, we're inclined to credit the true belief to the agent's abilities, something we're clearly *not* prepared to do (in any sense) in standard Gettier cases where it's *intervening luck* at issue.

In arguing, contra Pritchard, that this much does not imply that knowledge and cognitive achievements relevantly come apart, I stressed the importance of distinguishing between *two distinct ways* to specify the cognitive success of believing truly and noted that a cognitive achievement could consist in *either* success being creditable to one's intellectual abilities.

The argument was that: in specifying the sense in which knowledge excludes luck, we had to distinguish between two kinds of cognitive successes one might have when having a true belief, and we said that only one of these success must not depend on luck. Specifically, it must not be a matter of luck that the belief (given the initial conditions) was true; one's *having a true belief* is also a cognitive success, although it didn't matter (for the prospects of whether you know) if *that* cognitive success was a matter of luck. After all, as the refrigerator case showed, I could have easily not *had that true belief*—but that fact wasn't relevant to whether I knew it.

This showed that 'believing truly' picks out either one of two distinct cognitive successes, and correspondingly, there are *two kinds of cognitive achievements* one might have when one believes truly. These two possible cognitive achievements, in the Barney case, would be:

Cognitive achievement 1 (CA1): When we credit to Barney's abilities that he has a true belief.

Cognitive achievement 2 (CA2): When we credit to Barney's abilities that the belief Barney has is true.

Now, the cognitive achievement captured by CA1, along with RVE3 imply RVE4:

RVE4: If *S*'s having a true belief that *p* is creditable to *S*'s cognitive abilities, then *p* is safe.

RVE4 is clearly false: we credit to Barney's abilities his *having the true belief p* (i.e. 'there's a barn', although his 'there's a barn' belief was not safe.) This shows that CA1 cognitive achievements are compatible with knowledge-undermining environmental luck, and so knowledge and CA1 cognitive achievements do come apart.

But the fact that RVE4 is false shouldn't concern us. After all, if it doesn't matter for knowledge whether one's *having some true belief p* owes to luck, why think it would matter for knowledge whether one's *having some true belief p* owes to ability? Indeed, those who contend that cognitive achievements are essential to knowledge would have no reason to think that achievements in the sense captured by CA1 should explain this.

Instead, we should think that the cognitive success relevantly at issue in the cognitive achievements we're taking to be an essential part of knowledge would be *the same cognitive success* that knowledge requires not be lucky: one that constitutes a cognitive achievement only if it owes to an agent's abilities *that the belief she has is true*. Here again, we may stress Sosa's (2007) point that:

What must be attributable to the competence is not just the belief's existence but its correctness. (Sosa 2007: 33)

Indeed, when the *robust virtue epistemologist* contends that cognitive achievements are an essential part of knowledge, this claim should be thought to be threatened by arguments that show how knowledge and achievements in the sense expressed by CA2 come apart, *not* by arguments that show how knowledge and achievements in the sense expressed by CA1 come apart.

With this much now suitably clear, I'll not belabour the point further, but instead show how this will change what RVE3 commits us to.

From

RVE3: If S 's true belief that p constitutes a cognitive achievement, then p is safe.

The implication will not be that we'll have to defend the obviously false:

RVE4: If S 's having a true belief that p is creditable to S 's cognitive abilities, then p is safe.

But instead, we're committed to defending:

RVE4*: If it is creditable to S 's cognitive abilities that S 's belief that p is true, then p is safe.

RVE4*, unlike RVE4, is not undermined by the Barney case. Although we credit to Barney's perceptual abilities his true belief as much as we credit to Archie's abilities his successful shot, recall that it wasn't to the credit of Archie's abilities that the shot he fired hit the unguarded target rather than any of the other eleven that would have deflected his arrow. Nor was it to the credit of Barney's abilities that it was a real barn, rather than any of the surrounding barn facades, that he believed was a barn. Thus, it's not creditable to Barney's abilities that the belief he formed was true.

This much shows only that *environmental luck* cases, while counterexamples to RVE4, do not pose a problem for what we're *actually* committed to defending—RVE4*. In order to *positively defend* RVE4*, though, I'll need to show that *when* it is creditable to S 's cognitive abilities that the belief she has is true, then the belief will be appropriately safe.

10.4 A positive defence for RVE4*

RVE4*: If it is creditable to S 's cognitive abilities that S 's belief that p is true, then p is safe.

A starting point: our recent discussion of the Barney case was one where we didn't just *say* that it wasn't to the credit of Barney's abilities that the belief he had was true; we also gave a reason to support this thought. Because the reason we gave was meant to explain why it wasn't to the credit of Barney's abilities that his belief was true, then naturally this is something we may recast as a *necessary condition* for it being creditable to one's abilities that the belief she has is true. With that in place, then the next step would be to try to show that satisfying this necessary condition will ensure that the target belief is appropriately safe. If we can do that, then *Voilà!* We'll have defended RVE4*, and from there, it will be easy to see that the rest of what we want follows.

*Initial attempts to defend RVE4**

Looking back at Barney, we could explain why it wasn't to the credit of his abilities that the belief was true by pointing out that his abilities had nothing to do with why his belief was about a barn (and true) rather than any of the nearby barn facades (and false).

Drawing from this observation, a first pass at the *necessary condition* we're looking for (and hoping to subsequently defend as a safety-ensurer) would be:

Attempt 1: It is creditable to S 's cognitive abilities that S 's belief that p is true only if S 's cognitive abilities have *something* to do with why S 's belief was true rather than false.

Attempt 1 needs more refining. One way to figure out what that 'something' is would be to consider more closely what's going on when, as in the case of Barney and Archie, our heroes' abilities seemed to have *nothing* to do with why their respective efforts resulted in success rather than failure. Here we might point out that one thing going on in these cases was that it seemed to be a matter of luck that Barney's and Archie's efforts were successes. Perhaps then *that's* why we weren't inclined to credit the successes of these efforts to their abilities. This thought invites a second attempt at a necessary condition:

Attempt 2: It is creditable to S 's cognitive abilities that S 's belief that p is true only if it is not a matter of luck that S 's belief that p is true.

Attempt 2 specifies a necessary condition for crediting to S 's cognitive abilities that her belief is true that ensures (though trivially) the safety of the belief. Even though that's what we said we wanted, it should be obvious that Attempt 2 runs into a problem. For one thing, it seems to leave the safety condition doing all the work—but more importantly, Attempt 2 rests on a dubious though easily taken-for-granted assumption about the relationship between luck and ability. To put the assumption behind Attempt 2 generally, the idea is that: if it is by luck that one's belief is true, then it's not due to her abilities that her belief is true; but if it *was* due to her abilities that her belief is true, then it would not be by luck that her belief is true. These assumptions, which must be in place in order for Attempt 2 to be plausible, express the thought that a true belief's owing to ability is incompatible with its owing to luck. Call this assumption *dependence incompatibilism*.

Dependence incompatibilism: If S believes p truly, then it is due to S 's abilities that the belief S has is true if and only if it is not due to luck that S 's belief that p is true.

Greco (2006), for one, is on board with dependence incompatibilism. This is implied straightforwardly by his view that a true believe will be *because of abilities* only when it's not *because of luck* and vice versa. This makes sense only if such dependences are taken to be mutually incompatible.

I'll argue now that *dependence incompatibilism* is false. In seeing why it's false, we'll be in a good position to improve significantly on Attempts 1 and 2. Toward this end, consider the following case, which I think will sharpen how it is we think of ability as something at odds with luck in accounting for our true beliefs.

[CAR-PARK CASE]: I park my car in a safe parking lot, lock it and walk 30 meters or so to a classroom, which has no windows offering a direct view of my car. As soon as I trek the 30 meters and step into the classroom, I form the belief that my car is still in the parking lot where I parked it, and indeed I am correct about this. Most any theory worth its salt will count this case as uncontroversially one in which I have knowledge. Suppose though that I stay in the classroom all afternoon. Do I still know my car is outside, 30 meters from the building? Plausibly, if I knew it was there when I walked in, I didn't lose that knowledge in the space of a couple hours. Nonetheless, my belief three hours later that my car is in the lot (which it is), even if I know it, owes to luck a bit more than my belief did as soon as I walked into the room. And owing even more to luck is the truth of my belief, at 11:00 at night, that my car is still there. And if I stayed in the room for three weeks, the truth of the belief would depend on luck significantly enough to say that I no longer know it is there. In three weeks time, the world will have had a much better chance to be different from what I'd thought. In three hours time, it would have had much less chance, and in three minutes, hardly any chance at all.

The CAR-PARK case invites us to think of luck and ability as standing in a contrast that is gradient rather than rigid. It would thus seem to be mistaken to assess whether a true belief depends on ability by asking whether it depends on luck or not (as Attempt 2 did). What's important isn't whether—but *how much!* Just as dependence on luck is a matter of degree, it is plausible to think that a given true belief an agent has will depend on ability by degree as well. After all, of the beliefs we could be rightly said to know, and which we think do not relevantly depend on luck, our abilities play a larger role in our having some than others. The fact that dependence of our true beliefs on luck and ability is somewhat gradient rationalizes that a theory be concerned not with incorporating an incompatibilist principle, but rather a *comparative* one. That is, we need a principle that contrasts the extent to which the success of one's cognitive efforts depends on luck with the extent to which it depends on ability.

This motivates then a third attempt at giving a necessary condition we can use to defend RVE4*. Let's call what Attempt 3 offers the *comparative condition*.

Attempt 3 (Comparative Condition) (CC): It is creditable to S 's cognitive abilities that S 's belief that p is true only if S 's having a belief that is true is a matter that depends on S 's cognitive abilities *more so* than it does on luck.

Our third attempt (CC)—a necessary condition on the ability requirement that itself has modal import—has special promise in that it appears to offer us some way to defend:

RVE4*: If it is creditable to S 's cognitive abilities that S 's belief that p is true, then p is safe.

But first, CC must be unpacked so that we have an idea how to go about *comparing* dependence: what do we mean when we say it would depend on ability *more so* than luck that the belief one has is true?

When does the success of one's cognitive efforts depend on ability more so than luck?

Here it helps (as it so often does) to bring in the Archie case. We credit Archie's abilities for why his shot was successful rather than unsuccessful in worlds where all the targets are normal—here that the shot he has is successful owes to his ability more than to luck; but we don't credit his abilities for why his shot succeeds rather than fails in worlds where he could have just as easily shot at a forcefield target as a normal target—here that the shot he had is successful owes to luck more than to his ability.

We may think of worlds where the success of Archie's shot owes primarily to ability as being like the former sort, and we may think of worlds where the success of Archie's shot owes primarily to luck as being like the latter (forcefield) sort. Therefore, if Archie were to shoot and hit a target right now in the actual world, his success would owe to his ability *more so* than to luck only if worlds like the *former* are closer to the actual world than worlds like the *latter*.

Worlds like the former are ones where, holding fixed Archie's abilities, he couldn't easily have intended to hit a target and failed. Worlds like the latter are almost all ones where, holding fixed Archie's abilities, he shoots and fails (e.g. he hits one of forcefields).

This leads us to say that: That the shot Archie fires in the actual world is successful depends on Archie's ability *more so* than luck just in case worlds where (holding fixed Archie's abilities) he couldn't easily have intended to succeed and failed, are closer to the actual world than worlds where, holding fixed Archie's abilities, he intends to succeed and fails.

Naturally, then, we can extend what we've said about Archie's successful shot to Barney's successful belief.

That the belief Barney formed was true depends on Barney's ability moreso than luck just in case worlds in which Barney's abilities employed in the actual world couldn't have easily led him to a false belief about whether or not p are closer to the actual world than worlds where, holding these abilities fixed, he forms a false belief about whether or not p .

In order to defend RVE4*, though, I'll have to show that S 's belief that p is *appropriately safe* in worlds where it depends on S 's ability moreso than luck that S 's belief that p is true. And so I'll now defend the following argument, where the crucial premise is (3), and the conclusion is RVE4*.

*Argument for RVE4**

- (1) If it is creditable to S 's cognitive abilities that S 's belief that p is true, then it depends on S 's ability moreso than luck that S 's belief that p is true.
 - (2) It depends on S 's ability moreso than luck that S 's belief that p is true just in case worlds in which S 's abilities employed in the actual world couldn't easily have led S to a false belief about whether or not p are closer to the actual world than worlds where, holding these abilities fixed, S forms a false belief about whether or not p .
 - (3) If worlds in which S 's abilities employed in the actual world couldn't easily have led S to a false belief about whether or not p are closer to the actual world than worlds where, holding these abilities fixed, S forms a false belief about whether or not p , then S 's belief that p is safe.
-
- (4) Therefore, if it is creditable to S 's cognitive abilities that S 's belief that p is true, then S 's belief that p is appropriately safe. (RVE4*)

Now, recall what we took from the previous chapter to be the best formulation of the safety requirement for knowledge—Pritchard's SP**:

SP** S 's belief is safe iff in most near-by possible worlds in which S continues to form her belief about the target proposition in the same way as in the actual world, and in all very close nearby possible worlds in which S continues to form her belief about the target proposition in the same way as in the actual world, the belief continues to be true. (2007*a*: 8)

SP**, recall, succeeded where Pritchard's original formulation of safety failed because SP** seemed both (i) strong enough a formulation to not count lottery propositions as safe, yet (ii) not so strong that it fails to count (the relevant kind of) garbage chute cases as safe. This was what made SP** seem so attractive in comparison to other formulations. Plausibly, I'll be able to defend that my own modal condition ensures a belief is appropriately safe only if I can show that: it *doesn't* depend on an agent's ability moreso than luck that her *my ticket is a loser* belief is true; yet it *does* depend on an agent's ability moreso than luck that her belief *my laundry is in the basement* is true in Sosa's garbage chute case. I'll need to defend this much in order to defend (3) in my argument.

10.4.1 Lottery propositions

Note that premise (2) here draws an equivalence between my requirement that the success of one's cognitive effort depend on ability moreso than luck and what I offered as a modal representation of this idea. To keep things nice and short, let's just refer to this as the *comparative view* (CV)

Comparative view (CV): It depends on S 's ability moreso than luck that S 's belief that p is true just in case worlds in which S 's abilities employed in the actual world couldn't easily have led S to a false belief about whether or not p are closer to the actual world than worlds where, holding these abilities fixed, S forms a false belief about whether or not p .

In order that I get the right result in the lottery case, I must—by appealing to the *comparative view*—show that it *doesn't* depend on an agent's ability moreso than luck that her *my ticket is a loser* belief is true. Therefore, by reference to (CV), I'll need to show that worlds in which S 's abilities employed in the actual world couldn't easily have led S to a false belief about whether or not p are *not* closer to the actual world than worlds where, holding these abilities fixed, S forms a false belief about whether or not p .

Now, the belief *my ticket is a loser* is one for which there is excellent probabilistic evidence. But nonetheless, as Pritchard (2007*a*, 2007*b*) has shown, this proposition is false in worlds very close to the actual world (worlds where a lottery ball bounces just slightly differently)—and so worlds very close to the actual world are ones where the abilities employed in the actual world (e.g. probabilistic reasoning) could easily have led you to a false belief about whether you won the lottery. After all, your ability to reason probabilistically would lead you to be equally convinced your ticket is a loser in the nearby worlds where you win as it does in

the actual world where the ticket loses. So *holding fixed this ability*, you could easily have been wrong about whether you won.

Worlds in which, holding fixed your ability to reason probabilistically, you *couldn't* easily have had a false belief about whether your ticket is a loser would be worlds where, holding fixed this ability, your ticket isn't (as it is in the actual world) a mere small bounce away from winning (otherwise, you *could* have easily been wrong). But these worlds are far away—they would be worlds where your ability to reason probabilistically about the odds of your ticket losing isn't based (as in the actual world) on normal evidence. For example, they might be worlds where your ability to reason probabilistically leads you to think your ticket is a loser because you hear that a corrupt conspiracy in New York rigged the ball machine in favor of the ticket they've agreed to share—from which you infer that your ticket is not a winner. Now this is a world where your ability to reason probabilistically couldn't have easily led you to a false belief—but indeed, this is a far away world, and certainly farther away from the actual world than worlds where you reason probabilistically that you've lost the lottery but win and are wrong. These worlds are only a ball-bounce (rather than a grand conspiracy rig) away from the actual world. Thus, we can see that the *comparative view* gets the right result about lottery propositions. Worlds in which *S*'s abilities employed in the actual world couldn't easily have led *S* to a false belief about whether or not *p* are *not* closer to the actual world than worlds where, holding these abilities fixed, *S* forms a false belief about whether or not *p*. And therefore, by reference to CV, it *doesn't* depend on an agent's ability moreso than luck that her *my ticket is a loser* belief is true.

10.4.2 The Garbage Chute Case

As the previous chapter showed, the problems posed by lottery propositions and the garbage chute case can work together as a double punch: have a view strong enough to rule out lottery propositions (as we ought to), and your view will also (problematically) rule out garbage chute cases—cases where intuitively we have knowledge, even though the evidential support for such propositions is not as strong as the evidential support we have for lottery propositions.

Let's look again at Sosa's Garbage Chute Case, which Pritchard showed SP** to handle in the previous chapter.

On my way down to the elevator I release a trash bag down the chute from my high rise condo. Presumably, I know my bag will soon be in the basement. But what if, having been released, it still (incredibly) were not to arrive there? That

presumably would be because it had been snagged somehow on the way down (an incredibly rare occurrence) or some such happenstance (Sosa 2000: 13).¹²⁷

What our comparative view needs to show is that the success of Sosa's belief *my bag is (or will soon be) in the basement* depends on Sosa's ability more so than luck. Thus, by reference to (CV), I'll need to show that worlds where Sosa's abilities employed in the actual world couldn't easily have led him (Sosa) to a false belief about whether or not his bag is in the basement are *closer to the actual world* than worlds where, holding these abilities fixed, Sosa forms a false belief (e.g. the bag gets stuck).

I think that Pritchard is quite right to say that how we fill out the details of the case actually matters a great deal. If we suppose that there is nothing on which bags ever come close to snagging (i.e. no nails, snags or hitches, just smooth metal aided by gravity), then worlds very close to the actual world will be ones where, given Sosa's ability to examine the garbage chute and assess accordingly whether the bag will end up in the basement, his belief couldn't easily have been false. It would take quite an anomaly to get a bag stuck in a smooth chute with nothing to catch on, and so the nearest world in which Sosa reasons that the bag will end up in the basement on the basis of his chute-observing abilities and is wrong will be a world where something quite peculiar happens—some aberrant nail emerges undetected into the chute zone from a moving wall, or a quirky magnetic field prevents the bag from going down. These are farther away worlds—and the nearer worlds will be those, like the actual world, where one's chute-assessing abilities couldn't easily have led to a false belief.

10.5 Putting it all together

The comparative view, we've seen, not only gets the intuitively right result in everyday cases, as well as in cases of intervening and environmental luck, but it *also* gets the right result in the lottery and garbage chute cases—something that was a virtue to SP**. Thus, the requirement that it depend on an agent's abilities more so than luck that the belief she has is true is one that captures a *very similar* idea as the requirement that a belief be safe in the sense that SP** specified. Because the same cases are given the same verdicts by our comparative view and SP**, it is plausible to suppose that satisfying our modal requirement implies that a belief will be appropriately safe. Indeed, unless there could be shown clear disconnect between the modal requirement I've offered and SP** that justifies thinking the latter is more intuitive, there's no reason to think otherwise. After all, if there

¹²⁷ *Ibid.* p. 6. Pritchard's quotation.

were such a disconnect, it would most likely have cropped up amidst its diagnoses of the lottery and garbage chute cases—where I showed the verdicts were the right ones.

Finally, the idea I've defended here preserves the thought that both the MAL and LCAL views have got *something* right about luck. The requirement that the success of an agent's cognitive efforts depend on ability moreso than luck speaks to the idea that—when the success owes to luck moreso than ability—it is out of the agent's control that his cognitive efforts were successful. This much is suggested even though luck itself is understood modally and not in terms of such control. Consequently, the view here accommodates the thought that the two most widely defended accounts of luck both hit at something captured by the condition that the success of one's cognitive efforts depend on ability moreso than luck—a *prima facie* good outcome for any view.

This concludes then a defense for Premise (3) of the larger argument (for RVE4*), upon which we'll now draw some further conclusions. The argument we'd set out to defend was as follows:

*Argument for RVE4**

- (1) If it is creditable to *S*'s cognitive abilities that *S*'s belief that *p* is true, then it depends on *S*'s ability moreso than luck that *S*'s belief that *p* is true.
 - (2) It depends on *S*'s ability moreso than luck that *S*'s belief that *p* is true just in case worlds in which *S*'s abilities employed in the actual world couldn't easily have led *S* to a false belief about whether or not *p* are closer to the actual world than worlds where, holding these abilities fixed, *S* forms a false belief about whether or not *p*.
 - (3) If worlds in which *S*'s abilities employed in the actual world couldn't easily have led *S* to a false belief about whether or not *p* are closer to the actual world than worlds where, holding these abilities fixed, *S* forms a false belief about whether or not *p*, then *S*'s belief that *p* is safe.
-
- (4) Therefore, if it is creditable to *S*'s cognitive abilities that *S*'s belief that *p* is true, then *S*'s belief that *p* is appropriately safe. (RVE4*)

With the conclusion (4) established, we can now go a few steps further and show how the larger picture unfolds. From (4) we can derive:

- (5) If *S*'s true belief that *p* constitutes a cognitive achievement, then *p* is appropriately safe (RVE3)

And if (5) is true, then both the ability condition and the anti-luck conditions will be satisfied by a theory of knowledge according to which:

(6) S knows that p if and only if S 's true belief that p constitutes a cognitive achievement.

And given that cognitive achievements themselves consist in cognitive successes that arise through cognitive abilities, we reach the *robust VE* conclusion:

(7) Robust VE: S knows that p if and only if S 's true belief that p arises *through* S 's cognitive ability.

10.6 Concluding Remarks

Put simply, the *robust VE* view I've defended claims that:

Robust VE: S knows that p if and only if S 's true belief that p arises *through* S 's cognitive ability.

Some of the key supporting claims I defended were:

- (i) S 's true belief that p arises through S 's cognitive ability iff it is creditable to S 's cognitive abilities that S 's belief that p is true.
- (ii) It is creditable to S 's cognitive abilities that S 's belief that p is true iff it depends on S 's ability moreso than luck that S 's belief that p is true
- (iii) It depends on S 's ability moreso than luck that S 's belief that p is true just in case worlds in which S 's abilities employed in the actual world couldn't easily have led S to a false belief about whether or not p are closer to the actual world than worlds where, holding these abilities fixed, S forms a false belief about whether or not p .
- (iv) If worlds in which S 's abilities employed in the actual world couldn't easily have led S to a false belief about whether or not p are closer to the actual world than worlds where, holding these abilities fixed, S forms a false belief about whether or not p , then S 's belief that p is appropriately safe.

(v) A belief will be appropriately safe if it satisfies Pritchard's safety: SP** S 's belief that p is safe iff in most near-by possible worlds in which S continues to form her belief about the target proposition in the same way as in the actual world, and in all very close nearby possible worlds in which S continues to form her belief about the target proposition in the same way as in the actual world, the belief continues to be true. (*Ibid.* 8)

Supporting claims (i) through (v) ground my RVE account in a way that shows it to be a *bona fide* variety of *anti-luck virtue epistemology*. The conditions in the analysis of knowledge will be satisfied *only if* the belief arises from ability and is appropriately safe—and this had the advantage of not requiring that we add a separate safety condition. It was argued that a separate safety condition could be added to the ability condition only on the pain of giving up rights to the *final value* response to the Meno Requirement. And so, having taken the route we did, we have a theory of knowledge that accommodates the value, ability and luck insights together in one analysis.

I've argued then that the most plausible variety of *anti-luck virtue epistemology* will be one on which the ability condition specified in the analysis would be satisfied *only if* the resulting belief is appropriately safe. A dual-condition account was not only problematic for the purpose of accommodating the Meno Requirement but also not necessary for accommodating both insights. And so an analysis of knowledge best preserves the ability and anti-luck insights when the former condition can be shown to imply the latter. The *robust VE* account of knowledge that resulted is now one that we can defend to be uniquely capable of accommodating our insights about knowledge that concern its distinctive value, the sense in which it arises from our abilities, and the way that it must exclude luck.

BIBLIOGRAPHY

- Adler, Jonathan (2006). 'Epistemological Problems of Testimony', in *Stanford Encyclopedia of Philosophy*, Sec. 4, 'Knowledge Transmission.'
- Ahlstrom, Kristofer (2007). 'An Argument Against Swamping' (*manuscript*)
- Alston, William (2005). *Beyond Justification': Dimensions of Epistemic Evaluation*, Cornell University Press
- Audi, Robert (1997). 'The Place of Testimony in the Fabric of Justification and Knowledge', *American Philosophical Quarterly* 34: 405–422.
- Axtell, Guy (2000). *Knowledge, Belief and Character*, Rowan & Littlefield.
- _____ (2008). 'Expanding Epistemology: A Responsibilist Approach,' *Philosophical Papers* 37, 1, 2008, 51-88.
- _____ (2007). 'Two for the Show' in *Synthese* 2007 issue on *Epistemic Luck*.
- Axtell, Guy & Carter, J. Adam (2008). 'Just the Right Thickness: A Defense of Second-Wave Virtue Epistemology' in *Philosophical Papers* Vol. 37 (3) 2008: pp. 413-434.
- Baehr, J (2009). 'Is There a Value Problem?' *The Value of Knowledge*, (eds.) A. Haddock, A. Millar, & D.H. Pritchard, Oxford University Press, Oxford.
- Battaly, Heather (2001). 'Thin Concepts to the Rescue: Thinning the Concepts of Epistemic Justification and Intellectual Virtue,' in Abrol Fairweather and Linda Zagzebski, eds. *Virtue Epistemology* (Oxford University Press, 2001).
- Becker, Kelly (2008). 'Epistemic Luck and the Generality Problem,' *Philosophical Studies* 139 (3).
- _____ (2007). *Epistemology Modalized*, Routledge, 2007.
- Beebe, Helen (2004). 'Causing and Nothingness', in L.A. Paul, E.J. Hall and J. Collins, eds., *Causation and Counterfactuals* (Cambridge, Mass.: MIT Press), 291-308
- Blackburn, Simon (2001). 'Reason, Virtue, and Knowledge' in Abrol Fairweather and Linda Zagzebski, eds. *Virtue Epistemology* (Oxford University Press, 2001).
- Brännmark, John (2007). 'Three Kinds of Organic Unity', (*manuscript*).
- Brentano Franz (1902). *The Origin of Our Knowledge of Right and Wrong*, London and Henley: Routledge & Kegan Paul.
- Brogaard, Berit. (2009). 'The Trivial Argument For Epistemic Value Pluralism' in *Epistemic Value* ed., D. Pritchard, Alan Millar, and Adrian Haddock, *Epistemic Value*, Oxford University Press.
- _____ (2005). 'I Know; Therefore I Understand,' (*manuscript*)

- _____ (2006). 'Can Virtue Reliabilism Explain the Value of Knowledge?' *Canadian Journal of Philosophy* 36 (2006) 335-354.
- Brogaard, Berit, and Salerno Joe (2008). 'Knowability Paradox' in Pritchard & Hendricks (2008) *New Waves in Epistemology*, Palgrave Macmillan.
- Brown, Campbell (2006). 'Two Kinds of Holism about Value', *The Philosophical Quarterly* 57: 228; 456–463.
- Carlson, Erik (2000). 'Organic Unities, Weak Lexicality, and the Additivity of Intrinsic Value' in *Value and Choice*, ed. Wlodek Rabinowicz. Lund: Lund Philosophy Reports.
- Carter, J. Adam (2007). 'The Meno Requirement: A Critical Survey' (*manuscript*).
- _____ (2007). 'Challenging the Meno Requirement' (*manuscript*) presented at *Value of Knowledge* Conference, Vrije University, Amsterdam. (August 2007)
- Chisholm, Roderick (1957). *Perceiving: A Philosophical Study*, Cornell University Press.
- _____ (1966). *The Theory of Knowledge*, Prentice Hall, rep. 1977.
- _____ (1986). *Brentano and Intrinsic Value*, Cambridge and New York: Cambridge University Press.
- Chrisman, Matthew, and Carter, J. Adam (2008). 'Is Epistemic Expressivism Incompatible with Inquiry?', (*draft*).
- Code, Lorraine (1987). *Epistemic Responsibility*, Brown University Publishing, Providence.
- Cohen, Stewart (1998). 'Contextualist Solutions to Epistemological Problems: Scepticism, Gettier, and the Lottery.' *Australasian Journal of Philosophy*, 76: 289-306.
- _____ (1999). 'Contextualism, Skepticism, and Reasons', in Tomberlin 1999.)
- Comesana, Juan (2005). 'Unsafe Knowledge', *Synthese* 146 (2005), 393-402.
- Conee, Earl & Feldman, Richard (2004). *Evidentialism*, Oxford: Oxford University Press.
- Craig, Edward (1990). *Knowledge and the State of Nature*. Oxford: Oxford University Press.
- David, Marian (2001). 'Truth as the Epistemic Goal,' in Steup, M. (ed) *Knowledge, Truth and Duty. Essays on Epistemic Justification, Responsibility and Virtue*. (Oxford, Oxford University Press), 151-69.
- Davies, Martin (2004). 'Epistemic Entitlement, Warrant Transmission and Easy Knowledge,' *Aristotelian Society Supplementary Volume*, 78 (2004), 213-45.
- Davidson, Donald (1967). 'Causal Relations', *Journal of Philosophy*, 64, reprinted in Davidson, 2001a
- DeRose, Keith (1999). 'Contextualism: An Explanation and Defense', in Greco and Sosa 1999.
- _____ (2002.) "Assertion, Knowledge, and Context," *Philosophical Review* 111: 167-203 Dougherty, Trent. (2007).

- Driver, Julia (2000). 'Moral and Epistemic Virtue', (essay) in Axtell's *Knowledge Belief and Character*, Rowan & Littlefield 2000.
- _____ (2001). *Uneasy Virtue*, Cambridge: Cambridge University Press.
- Elgin, Catherine (2007). 'Understanding and the Facts,' *Philosophical Studies*, Volume 132, Number 1 / January, 2007.
- Feinberg, Joel (1970). *Doing and Deserving: Essays in the Theory of Responsibility*. Princeton: Princeton University Press, 1970.
- Feldman, Richard (2007). 'Roderick Chisholm', Stanford Encyclopedia of Philosophy, 'Epistemology I: Epistemic Terms, Principles.'
- Gettier, Edmund (1963). *Is Justified True Belief Knowledge?*, *Analysis* V. 23.
- Goldman, Alvin (1967). 'A Causal Theory of Knowing,' *Journal of Philosophy* 64: 357-372 (1967).
- _____ (1999). *Knowledge in a Social World*, Oxford: Oxford University Press (1999).
- _____ (2004). 'The Unity of the Epistemic Virtues' *Pathways to Knowledge*, Vol. 1, No. 9., pp. 51-72.
- _____ (1992). *Liaisons: Philosophy Meets the Cognitive and Social Sciences*, MIT Press (1992).
- _____ (1976). 'Discrimination and Perceptual Knowledge,' *The Journal of Philosophy* 73: 771-791 (1976). Reprinted in *Liaisons* (1992).
- _____ (1979). 'What is Justified Belief?' in George Pappas, ed., *Justification and Knowledge*, D. Reidel (1979). Reprinted in *Liaisons* (1992).
- _____ (1992). 'Epistemic Folkways and Scientific Epistemology', in Goldman, *Liaisons: Philosophy Meets the Cognitive and Social Sciences*, MIT Press/Bradford Books (1992)
- Goldman, Alvin & Olsson, Erik (2009). 'Reliabilism and the Value of Knowledge', in *The Value of Knowledge*, eds. A. Haddock, A. Millar & D.H. Pritchard. Oxford: Oxford University Press.
- Goodman, Nelson (1951). *The Structure of Appearance*. Harvard: Harvard University Press.
- Greco, John (1995). 'Modern Ontology and the Problems of Epistemology,' *American Philosophical Quarterly* 32, 3, pp. 241-251.
- _____ (1999). 'Agent Reliabilism,' in *Philosophical Perspectives, 13, Epistemology*, James Tomberlin, ed. (Atascadero, CA: Ridgeview Press, 1999).
- _____ (2008). 'What's Wrong with Contextualism?', *Philosophical Quarterly*, 58, 232 (2008): 416-436.
- _____ (2009a). *Achieving Knowledge*, Cambridge: Cambridge University Press.
- _____ (2009b). 'The Value Problem' *The Value of Knowledge*, eds. A. Haddock, A. Millar & D.H.Pritchard, Oxford: Oxford University Press.

- Grimm, Stephen (2008). 'Epistemic Goals and Epistemic Values.' *Philosophy and Phenomenological Research* 77:3
- _____(2009). 'Epistemic Normativity.' In *Epistemic Value*. Eds. Adrian Haddock, Alan Millar, and Duncan Pritchard. Oxford University Press.
- Hawthorne, John (2004). *Knowledge and Lotteries*. Oxford: Oxford University Press.
- Hiller, Avram & Neta, Ram (2007). 'Safety and Epistemic Luck' in *Synthese* edition on Epistemic Luck, 2007.
- Hintikka, Jaako (1962). *Knowledge and Belief*. New York: Cornell University Press.
- Hookway, Christopher (2000). 'Regulating Inquiry', in Axtell (ed.) (2000), 149-160.
- _____(2006). 'Epistemology and Inquiry: The Primacy of Practice,' in S. Heatherington (ed.) *Epistemology Futures*, 95-110.
- Hudson, Hud (2001). *A Materialist Metaphysics of the Human Person*. New York: Cornell University Press.
- Hursthouse, Rosalind (1999). *On Virtue Ethics*, Oxford: Oxford University Press.
- Jones, Ward (1997). 'Why Do We Value Knowledge?' *American Philosophical Quarterly* 34:4: 423-439.
- Kagan, Shelly (1989). *The Limits of Morality*, Oxford: Oxford University Press.
- Kelp, Christoph (2008). 'Knowledge and Safety.' *Journal of Philosophical Research* 33.
- Korman, Daniel (2004). 'Unrestricted Composition and the Argument from Vagueness' (*manuscript*)
- Kornblith, Hilary (2002). *Knowledge and its Place in Nature*. Oxford: Oxford University Press.
- Korsgaard, Christine (1983). 'Two Distinctions in Goodness' *The Philosophical Review*, Volume 92, Number 2: 169-195.
- Kvanvig, Jonathan (2003). *The Value of Knowledge and the Pursuit of Understanding*. Cambridge: Cambridge University Press.
- _____(2005). 'Truth and the Epistemic Goal' *Contemporary Debates in Epistemology*, eds. Matthias Steup and Ernest Sosa.
- _____(2008a). 'Critical Notice of Pritchard's Epistemic Luck' *Philosophy and Phenomenological Research* 77 (July 2008), pp. 272-81.
- _____(2008b). 'Pointless Truths.' *Midwest Studies in Philosophy* Volume XXXII (2008), pp. 199-212.
- _____(2009a). 'Responses to Critics,' *Epistemic Value*, edited by Pritchard, Haddock, and Millar, (Oxford: Oxford University Press, forthcoming 2009).
- _____(2009b). 'The Swamping Problem Redux: Pith and Gist' *Social Epistemology*, forthcoming, 2009. Formerly entitled: "Further Thoughts on the Swamping Problem."

- Lackey, Jennifer (2006). 'Why We Don't Deserve Credit for Everything We Know,' *Synthese* 158 (2007): 345-61.
- _____ (2003). 'A Minimal Expression of Non-Reductionism in the Epistemology of Testimony,' *Notus* 37: 706-23.
- _____ (2008). 'What Luck is Not,' *Australasian Journal of Philosophy*, Volume 86, Issue 2 June 2008 , 255 – 267
- _____ (forthcoming). 'Knowledge and Credit' (typescript).
- Leonard, Henry. S. & Goodman, Nelson (1940). 'The Calculus of Individuals and its Uses', *The Journal of Symbolic Logic* 5: 45-50.
- Lesniewski, Stanislaw (1983). 'On the Foundations of Mathematics' *Topoi*, 2, 7–52.
- Levin, Michael (2004). 'Virtue Epistemology: No New Cures,' *Philosophy and Phenomenological Research* Volume 69 Issue 2: 397- 410.
- Lewis, David (1970). 'Causation,' *The Journal of Philosophy*, Vol. 70, No. 17, Seventieth Annual Meeting of the APA Eastern Division (Oct. 11, 1973), pp. 556-7
- _____ (1986). *On the Plurality of Worlds*. New York: Blackwell Publishing.
- Lynch, Michael (2008). 'Truth, Value and Epistemic Expressivism,' *Philosophy and Phenomenological Research*.
- McDowell, John (1994). 'Knowledge by Hearsay.' in B. K. Matilal and A. Chakrabarti, eds, *Knowing from Words* (Kluwer, Dordrecht, 1993; Synthese Library vol. 230), pp. 195-224
- Meylan, Anne (2007). 'The Value Problem: Against a Reliabilist Solution,' *Latin Meeting in Analytic Philosophy Genova 20-22 September 2007 CEUR-WS Proceedings ISSN 1613-0073. pp. 85-9*
- Miscevic, Nenad (2007). 'Virtue-Based Epistemology and the Centrality of Truth,' *Acta Analytica* Volume 22, Number 3 / September, 2007
- Montmarquet, James (1991). *Epistemic Virtue and Doxastic Responsibility*, Rowman and Littlefield.
- Moore, G.E. (1903). *Principia Ethica*, Cambridge: Cambridge University Press.
- _____ (1912). *Ethics*, Oxford: Oxford University Press.
- _____ (1922). *Philosophical Studies*, London: Routledge and Kegan Paul.
- _____ (1939). 'Proof of an External World' *Proceedings of the British Academy* 25 (1939) 273-300. Reprinted in *Philosophical Papers* and in *G. E. Moore: Selected Writings* 147-70.
- Olsson, Erik (2007). 'Reliabilism, Stability and the Value of Knowledge?'
- Olson, Jonas (forthcoming). 'A Question about Supervenience and Value-Making Properties'
- Plato. (*Meno*. tr. Benjamin Jowett. Place: Publishing Co.

- Pritchard, Duncan (2005). *Epistemic Luck*, Oxford University Press: Oxford.
- _____ (2007a). 'Anti-Luck Epistemology.' *Synthese* 158: 277-97.
- _____ (2007b) 'Safety-Based Epistemology: Whither Now?' *Journal of Philosophical Research* 33.
- _____ (2007c). 'Recent Work on Epistemic Value', *American Philosophical Quarterly* 44: 85-110.
- _____ (2008a). 'Sensitivity, Safety, and Anti-Luck Epistemology.' In J. Greco (ed.), *The Oxford Handbook of Scepticism*. Oxford: Oxford University Press.
- _____ (2008b). 'The Value of Knowledge.' (manuscript).
- _____ (2008c). 'Knowledge and Value' (manuscript)
- _____ (2008d). 'Virtue Epistemology and Epistemic Luck, Revisited', *Metaphilosophy* 39 (2008), 66-88
- _____ (2009). 'Anti-Luck Virtue Epistemology' *forthcoming* in A. Haddock, A. Millar & D. H. Pritchard (eds.), *Epistemic Value*. Oxford: Oxford University Press.
- _____ (forthcoming a). 'What is the Swamping Problem?' *Reasons for Belief*, (eds.) A. Reisner & A. Steglich-Petersen, (Springer, forthcoming).
- _____ (forthcoming b). 'Evidentialism, Internalism, Disjunctivism', *Evidentialism and its Discontents*, (ed.) T. Dougherty, (Oxford UP, forthcoming).
- Pritchard, Duncan H. and Hendricks, Vincent (2007). *New Waves in Epistemology*, Palgrave and MacMillan.
- Pryor, James. (2004). 'What's Wrong With Moore's Argument?' *Philosophical Issues*; Vol. 14.
- Quine, W.V.O. (1969). 'Epistemology Naturalized', reprinted in E. Sosa & J. Kim. *Epistemology: An Anthology*. Malden, MA: Blackwell Publishing. pp. 292–300
- Rabinowicz, W. and Ronnow-Rasmussen, T. (2000). 'A Distinction in Value: Intrinsic and for Its Own Sake', *Proceedings of the Aristotelian Society* 100: 33-49.
- Reichenbach, Hans (1951). *The Rise of Scientific Philosophy*, University of California Press.
- Riggs, Wayne (2004). 'Insight, Open-mindedness and Understanding.' (manuscript).
- _____ (2009). 'Luck, Knowledge, and Control.' In A. Haddock, A. Millar & D. H. Pritchard (eds.), *Epistemic Value*. Oxford: Oxford University Press.
- _____ (2007). 'Why Epistemologists Are So Down On Their Luck,' *Synthese* 158(3), 2007.
- _____ (2008). 'The Value Turn in Epistemology' in *New Waves in Epistemology*, Vincent Hendricks & Duncan Pritchard, eds. Palgrave Macmillan, 2008.

- _____ (2009). 'Getting the Meno Requirement Right,' in *The Value of Knowledge*, eds. A. Haddock, A. Millar & D.H. Pritchard, Oxford University Press, *forthcoming*.
- Ronnow-Rasmussen, Toni (*forthcoming*). 'Subjectivism and Objectivism', (*typescript*) p. 248.
- Sider, Ted (2001). *Four-Dimensionalism*. Oxford: Clarendon Press.
- _____ (2007). 'Parthood', *Philosophical Review* 116: 51-91.
- Simons, Peter (1987). *Parts: A Study in Ontology*. Oxford: Clarendon Press.
- Sosa, Ernest (1991). *Knowledge in Perspective: Selected Essays in Epistemology* (Cambridge University Press).
- _____ (1993). 'Epistemology, Realism, and Truth,' in *Philosophical Perspectives* 7 (Ridgeview Press), pp. 1-16.
- _____ (2000). 'Reliabilism and Intellectual Virtue' in *Knowledge, Belief and Character*, ed. G. Axtell, Rowman and Littlefield Publishers.
- _____ (2001). 'For the Love of Truth', In Abrol Fairweather and Linda Zagzebski (eds.), *Virtue Epistemology* (OUP, 2001).
- _____ (2002). 'The Place of Truth in Epistemology,' in *Intellectual Virtue: Perspectives from Ethics and Epistemology*, edited by Michael DePaul and Linda Zagzebski (Oxford University Press).
- _____ (2007). *A Virtue Epistemology: Apt Belief and Reflective Knowledge*. Oxford: Oxford University Press.
- _____ (2007). 'Epistemic Normativity' in *A Virtue Epistemology: Apt Belief and Reflective Knowledge*, Vol. 1. Oxford: Oxford University Press.
- Swinburne, Richard (1999). *Providence and the Problem of Evil*. Oxford: Oxford University Press.
- Unger, Peter (1968). 'An Analysis of Factual Knowledge', *Journal of Philosophy* 65: 157-70.
- Van Cleve, James (1986). 'Mereological Essentialism, Mereological Conjunctivism, and Identity Through Time', *Midwest Studies in Philosophy* 11: 141-156.
- Van Inwagen, Peter (1990). *Material Beings*. New York: Cornell University Press.
- Van Woudenberg, Rene (2006). 'Which Value for What Knowledge?' (*manuscript*)
- Varzi, A.C. (2000). 'Mereological Commitments', *Dialectica* 54: 283-305.
- Williams, B. (2002). *Truth and Truthfulness: An Essay in Genealogy*. Oxford: Oxford University Press.
- _____ (1985). *Ethics and the Limits of Philosophy*. Boston: Harvard University Press.
- Williamson, Timothy (2000). *Knowledge and its Limits*. Oxford: Oxford University Press.

Wittgenstein, Ludwig (1953). 'On Certainty' ed. G.E.M. Anscombe and G.H. von Wright, Translated by Denis Paul and G.E.M. Anscombe, Basil, Oxford: Blackwell Publishing, 1969-1975.

Wright, Crispin (2007). 'The Perils of Dogmatism', To appear in *Themes from G. E. Moore: New Essays in Epistemology and Ethics*, eds. Susana Nuccetelli and Gary Seay, Oxford.

Zagzebski, Linda (1996). *Virtues of the Mind: An Inquiry into the Nature of Virtue and the Ethical Foundations of Knowledge*.

_____(2002). 'The Search for the Source of Epistemic Good', in *Moral and Epistemic Virtues*, ed. M. Brady and D.H. Pritchard, New York: Blackwell Publishing.

