

VIRTUE EPISTEMOLOGY:
ITS PROPER FORM AND ITS APPLICATIONS

by
Sarah Wright

A Dissertation Submitted to the Faculty of the
DEPARTMENT OF PHILOSOPHY
In Partial Fulfillment of the Requirements
For the Degree of
DOCTOR OF PHILOSOPHY
In the Graduate College
THE UNIVERSITY OF ARIZONA

2005

THE UNIVERSITY OF ARIZONA
GRADUATE COLLEGE

As members of the Dissertation Committee, we certify that we have read the dissertation prepared by Sarah Wright entitled *Virtue Epistemology: Its Proper Form and Its Applications* and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy

_____ Date: 7/19/05
Keith Lehrer

_____ Date: 7/19/05
Terence Horgan

_____ Date: 7/19/05
David Schmitz

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copies of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

_____ Date: 7/19/05
Dissertation Director: Keith Lehrer

STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at The University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgment of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the head of the major department or the Dean of the Graduate College when in his or her judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

Sarah Wright

ACKNOWLEDGEMENTS

I wish to thank Keith Lehrer and Terry Horgan for helping me throughout the process of creating this dissertation. Both have been amazingly generous with their time and have always been willing to listen to my ideas (however off the wall) and offer their comments, suggestions, and encouragement. Both of them have helped me immensely in the production of this dissertation and in my development as a philosopher. I also wish to thank Dave Schmitz for being willing to join my committee and for being quite generous with his time and comments. Even though we were not working directly on the topics on chapter 6, many of my ideas in this chapter grew out of my experience of writing an article with him. The experience of writing with him was also pivotal in helping me to develop as a writer of philosophy.

I would also like to thank the University of Arizona Philosophy Department community for providing a wonderful environment in which to work. In particular working with John Pollock has given me the background I needed to even being writing this dissertation. Julia Annas' wonderful seminar on the Stoics first introduced me to central ideas in virtue ethics. Cara Nine and Mark Timmons both gave me very helpful suggestions on earlier drafts. I would like to acknowledge a deep gratitude to Justin Fisher for his patience and support in the processes of writing this dissertation, as well as his insightful suggestions and commentary.

Finally I am indebted for the comments and suggestions given to me by members of the dissertation reading group at the University of Arizona, and members of the audiences at colloquia at the University of Arizona, the University of Georgia, the University of North Carolina, and the Western Canadian Philosophical Association.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS	7
ABSTRACT	8
CHAPTER 1: INTRODUCTION	9
What is Virtue Epistemology?	9
Virtue Epistemology is a Historical Theory	13
Virtue Theory should be Internalist	19
An Overview of My Arguments for Internalist Virtue Epistemology	22
CHAPTER 2: ON THE EXISTENCE OF CHARACTER TRAITS	25
The Situationist Objection to Virtue Theories	25
Objections to the Methodology of These Studies	30
Objections to the Claim That These Studies Show There is No Virtue	34
Epistemic Virtues?	35
Empirical Evidence Favoring a Virtue Approach to Epistemology	38
CHAPTER 3: THE PROPER STRUCTURE OF THE EPISTEMIC VIRTUES	41
The New Evil Demon Problem	43
Goldman’s Use of the Virtues in his Two-Stage Reliabilism	44
Sosa’s Contextualist Reliabilism and the New Evil Demon Problem	47
Zagzebski’s Virtues and the Generality of the New Evil Demon Problem	50
The Historical Argument for Internalism	54
The Argument from the Sources of Value	58
The Proper Structure of the Epistemic Virtues	62
Knowledge and Gettier: Undermining a Motivation for Externalism	69
CHAPTER 4: CONTEXT SENSITIVITY	77
Worries about Skepticism	77
Other Versions of Context-Sensitivity	81
Comparing Inferential Contextualism to Attributor Contextualism	87
Comparing Virtue Epistemology to Inferential Contextualism	94
CHAPTER 5: VIRTUE, REDUCTION, AND TESTIMONY	99
The Humean Position	102
Coady’s First Objection	103
Coady’s Second Objection	106
Coady’s Third Objection	114
Using Virtue Epistemology to Justify Reliance on Testimony	119
Conclusion	123

TABLE OF CONTENTS -- *Continued*

CHAPTER 6: REASONING ABOUT PROBABILITY	125
Puzzling Cases	126
<u>The Sleeping Beauty Problem</u>	126
<u>Doomsday Problem</u>	130
The Partitioning Problem	132
Bayesian Response to the Partitioning Problem	136
Belief-Based Theories and the Partitioning Problem	138
Virtue Theories and the Partitioning Problem	141
WORKS CITED	145

LIST OF ILLUSTRATIONS

FIGURE 1 Alternate Partition for the Doomsday Problem	133
FIGURE 2 Alternate Partitions for the Sleeping Beauty Problem	135

ABSTRACT

Epistemology is an evaluative enterprise. But what should we take to be the primary unit of evaluation? Traditionally, individual beliefs have served as the primary units of epistemic evaluation. I argue that epistemology should instead use a different unit of evaluation--the agent and her character traits. Such a theory is a virtue epistemology.

What makes a character trait a virtue? There are two competing answers to this question. The externalist holds that it depends on the relationship between the character trait and the world. The internalist holds that it depends on the ways the character trait motivates us to respond to our perceptions of the world. I argue, contrary to recent developments in virtue epistemology, that we should accept an internalist conception of virtue.

How universal are the standards for virtue and vice? Rather than holding that standards are universal and do not depend on context, the contextualist holds that the standards for virtue and vice vary depending on the particulars of the context. I argue for a contextualist version of virtue epistemology, and show why context-sensitive virtue theory is superior to other potential versions of contextualism.

Finally, I apply the developed notion of a contextually-sensitive, internalist virtue epistemology to two intriguing areas in epistemology. I argue that my view is better able to account for certain otherwise puzzling phenomena, including questions about the epistemic relevance of the testimony of others and about how we could have the capacities with reasoning about probability that we routinely exhibit.

CHAPTER 1: INTRODUCTION

What is Virtue Epistemology?

Epistemology is a normative enterprise. It sets up norms for good epistemic behavior. We use these norms to evaluate the epistemic behavior of others, and we sometimes hope to use them to regulate our own epistemic behavior. We use epistemic norms to apportion praise and blame for our epistemic practices. Given that epistemology is normative, we need to give an account of its normativity.

Although epistemology is normative, this fact is often not fully explored and explained in developing different epistemological theories¹. These theories do not altogether lack a normative component; rather the details of the normative component are often overlooked or underdeveloped. But the normative theory endorsed by an epistemology can radically affect its structure and the goals it endorses. Traditionally epistemology has concerned itself primarily with giving an account of knowledge and of justified belief. These goals reflect a belief-based approach to normativity that parallels an act-based approach to normativity in ethics. Act-based theories of ethics include both deontological and consequentialist theories. Although deontological and consequentialist theories will lead us to very different conclusions, they both start with the same question, “What makes an act a right act?” Starting with this question gives a certain structure to the theories aimed at answering it. Those theories will take as basic the projects of defining right action, and will begin by

¹ Notable exceptions include Chisholm (1957) and (1977) who takes it as a central project to provide an ethics of belief, and Alston (1993) and Goldman (1999), who both argue that the concept of justification has been taken as deontological in character and that this deontological structure is not appropriate for application to justification of beliefs.

focusing on finding the characteristics of actions that make them right. If these theories include any reference to virtues, these virtues will be defined in terms of the more basic concept of right action. Thus on an act-based ethical theory the direction of analysis will lead from acts to virtues.

A similar direction of analysis is found in belief-based epistemologies. These theories begin with the question, “What makes a belief justified or an instance of knowledge?” This question might be summarized to parallel the question asked by the act-based ethical theorist as, “What makes a belief a good belief?” By focusing on justified belief and knowledge as the basic concepts, these theories will begin by searching for the relevant characteristics of beliefs that make them either justified or an instance of knowledge. If these theories develop a concept of epistemic virtue, that concept will be derivative from the concepts of justified belief and knowledge. Thus starting with a belief-based approach will set the agenda for and give structure to the inquiry for an epistemological theory.

Although the direction of analysis starting with acts and beliefs has been dominant in both modern ethics and modern epistemology, there is another alternative. This is the direction of analysis that starts not with the characteristics of individual acts and beliefs, but rather with characteristics of persons. This is the direction of analysis given by virtue theories. Virtue theories in ethics will begin by asking not “What makes an act a right act?” but “What makes a person a morally good person?” The answer to this question will begin by looking at character traits and by determining which of these make a person good person—that is, which character traits are virtues. Only then will we move on to define right actions, and

right actions will be defined in terms of the moral virtues thus the direction of analysis will lead from virtues to acts. Although the virtue approach has not generally been taken in modern ethics, it was the generally accepted direction of analysis in the moral theorizing of the ancient Greeks². Virtue approaches have been making resurgence in recent ethics, as people explore this alternative direction of analysis as a way to solve or get around problems in act-based theories of ethics³.

The direction of analysis endorsed by virtue ethics can also be pursued in epistemology. Virtue epistemology will not begin in the traditional way by analyzing justified belief and knowledge. Rather than asking “What makes a belief a good belief?”, virtue epistemology begins with the question, “What makes a believer a good believer?” Starting with this question will lead us to begin our epistemic enquiry with an analysis of the character traits of different believers. Those character traits that a good believer will have will then be the epistemic virtues. These epistemic virtues will then be used to give definitions of the belief-based concepts of justified belief and knowledge. The direction of analysis on virtue epistemology will thus lead us from the epistemic virtues to justified beliefs and knowledge.

One further difference between act or belief-based theories and virtue theories is that some act-based theories take as their aim looking for a non-normative criterion by which the right actions (or good beliefs) can be identified. This project in epistemology is called by

² See Annas (1993).

³ G.E.M. Anscombe (1958), Iris Murdoch (1967), and Philippa Foot (1978) are early examples of this resurgence of virtue theory, and Hursthouse (1999) is a more recent example.

Chisholm the “the problem of ‘the criterion’” and he describes it as a search for a right-making characteristic of beliefs that fits the following desiderata:

A ‘right making’ characteristic is one that can be described and identified in ethically neutral language.”⁴

Taking this project as a goal of epistemology puts Chisholm squarely in the belief-based approach to epistemology. However looking for a solution to the problem of the criterion is not a central aim of a virtue theory. For virtue theory does not take a definition of right acts or good beliefs as conceptually prior to the idea of virtues, either epistemic or moral. Indeed even at the level of analyzing the virtues, virtue theory does not look for an analysis that is devoid of normative language. For the virtues will be captured in terms that are both normatively loaded while also being descriptive. Such terms are said to be “thick moral concepts” in which the normative and descriptive are intertwined in such a way that we cannot eliminate either component⁵. Courage, generosity, and temperance are moral virtues, while intellectual courage, open-mindedness, and intellectual carefulness are epistemic virtues. None of these can be defined in a way that is entirely free of a normative element. This is because these character traits are not the sort that can be specified entirely by rules, rather they require that we develop a skill at both identifying them and acting according to them⁶. Thus while virtue epistemology does take as one of its goals the giving of a definition of justified belief and knowledge, these definitions of the right-making properties of beliefs will not be entirely free from normative content⁷. We will define justification and knowledge

⁴ Chisholm (1957) p 31.

⁵ This terminology of “thick” and “thin” moral concepts is due to Bernard Williams (1985).

⁶ For more on this topic see the discussion of codifiability in Chapter 6

⁷ Compare this to Keith Lehrer’s discussion of the “virtuous loop of reason” (2000) p. 142.

in terms of the epistemic virtues, and the epistemic virtues will be defined in terms of their role in the overall good life for human beings.

The presence of normative terminology does not mean that the concepts of epistemic and moral virtues are mysterious or unanalyzable. We are perfectly capable of engaging into a discussion of which character traits ought to be considered virtues, and discussion that refines our concepts of those virtues. The methodology to be used here is that of reflective equilibrium⁸. We can use this method to refine our thick concepts, including those of the epistemic and moral virtues. This process is not unlike the process that we use to refine other epistemic concepts, such as that of a reliable process, a foundational belief, or coherence. Thus virtue theory does not aim at giving a non-normative criterion of the right-making characteristics of beliefs, yet it can still give us a robust and useful account of those characteristics.

Virtue Epistemology is a Historical Theory

Ernest Sosa first suggested the application of virtue theory to epistemology in his “The Raft and the Pyramid” (1980). In this article, Sosa examines the debate between coherentism and foundationalism, and finds both views wanting. He suggests that the primary problem with coherentism is that it cannot explain the connection between our beliefs and our sensory experiences. There is a similar problem for the foundationalist, but he faces a dilemma. Either the relationship between sensory experience and foundational beliefs is primitive, in which case it is hard to see how it is justifying, or it is derived from other principles, in which

case foundational beliefs are not foundational. Sosa proposes a virtue epistemology as a ground from which the justification for the basing of beliefs on sensory experiences derives. He points to our “stable dispositions for truth acquisition” as the objects of “primary justification”⁹. These dispositions are justified by their reliability at reaching the truth, and hence Sosa identifies his view as reliabilism.

Sosa’s move to considerations of reliability is similar to that proposed by Alvin Goldman in “What is Justified Belief?” (1979). Goldman proposes an account of justified belief which depends upon the reliability of the belief forming processes from which they result¹⁰. This theory of justification is what he calls a “Historical or Genetic” theory, as opposed to a “Current-Time Slice” theory¹¹. On a Current-Time Slice theory the justification of a belief must be a function of only the properties of a believer and his other beliefs at the time of the target belief. A “Historical or Genetic” theory allows historical factors to play a role, such as requiring that the past and its causal connections to current beliefs be correct in order for a belief to be justified. The move to a Historical theory can be motivated by considerations of the distinction between having justification and being justified¹². Having justification consists in having beliefs (or other psychological states) that have the right evidential relationship to the belief in question. Being justified requires more than simply having justification because it requires the correct relationship between the evidence (either beliefs or other psychological states) and the target belief. To be justified, the target belief needs to

⁸ As presented by John Rawls in (1971).

⁹ Sosa (1980) p 189.

¹⁰ When these beliefs are inferred from other beliefs, then the conditional reliability of the reasoning process used to make the inferences also matters. See Goldman (1979) p 13-14.

¹¹ Goldman (1979) p 14. This terminology is borrowed from Robert Nozick (1974).

be based upon or caused by the evidence, rather than there simply being evidence that could bear upon it. We can see how having justification and being justified can come apart by considering a case from Chisholm which ignores this distinction:

If a man knows, say, that the hall has been painted, then, according to our definition, he accepts the hypothesis that it has been painted, he has adequate evidence for this hypothesis, and, finally, the hall *has* been painted. On the other hand, if he accepts the hypothesis, has adequate evidence for it, but does not *know* that it is true, then the hall has not been painted.¹³

In this description of the case, Chisholm seems to be ignoring the possibility that the man might have adequate evidence for the hypothesis that the hall has been painted, and yet be basing his belief that it has been painted on other, inadequate evidence. Perhaps he reaches his belief on the basis of wishful thinking, after hoping for weeks that they would get done with the messy painting job. Or perhaps he bases his belief on the testimonial evidence of a co-worker that he knows is a prankster and apt to tell lies. In either case it seems that the man could accept the hypothesis, *have* adequate evidence for it, and the hypothesis could also be true, and yet because the man does not base his belief on that evidence, he still does not *know* that the hall has been painted. Thus it seems that in order to give an adequate account of justification and knowledge, we will need to pay attention to the distinction between having justification and being justified, and that the latter will be required for an agent to have knowledge.

¹² See Horgan and Henderson (2000).

¹³ Chisholm (1957) p 16. Emphasis in original.

The importance of the distinction between having justification and being justified leads us to adopt a historical theory of justification. For in order to capture what it is for a belief to be justified we need to be able to look into the history of that belief and see how it was caused and how it is sustained¹⁴. Thus we need to have an historical account of justification.

Reliabilism is not the only possible historical theory; virtue epistemology is another.

Although reliabilism and virtue epistemology are sometimes conflated because they are both historical theories¹⁵, there can be important differences between them. The virtue theory that I will develop in the following chapters will both reject the requirement of reliability for a character trait to be a virtue, and require more than mere reliability in order for a character trait to be a virtue. Indeed the account of epistemic virtues that I will develop will require only that the belief be motivated in the correct way, and that that motivation be developed in the agent in an appropriate way. This move within historical theories might be seen as a response to problems that the historical reliabilist faces. One example of this is the Norman case given by Lawrence Bonjour. Norman is a clairvoyant, which is reliable with respect to certain circumstances and subject matters. Norman however has no evidence that he is reliably clairvoyant, or even that it is possible for people to be clairvoyant in the way he is. Norman comes to believe that the president is in New York, and this belief results from his

¹⁴ Keith Lehrer's Mr. Raco example shows us that we must look not only to how a belief was formed but also to how it is sustained. Mr. Raco comes to his belief that members of a particular racial group are susceptible to a particular disease on the basis of his racism. He later comes to study medicine and comes to have scientific evidence that in fact members of this race are susceptible to that particular disease. Even though his belief was caused by his racism, it might some to be sustained by appropriate scientific evidence. However since his belief would then be overdetermined, we need something more than a counterfactual analysis of causation to determine what is the true basis of his belief. See (2000) 196ff. For more on this topic see Chapter 3 on the role of motivation in virtues and belief.

¹⁵ See Sosa (1980).

clairvoyant ability¹⁶. We can now ask if Norman is justified in his belief. Even though his clairvoyant capacity is by hypothesis reliable, it seems counterintuitive to claim that Norman is justified. The virtue epistemologist can explain why Norman does not seem to be justified. For Norman, even though he is exercising a reliable faculty, doesn't seem to be epistemically virtuous. In the first place we might wonder about his motivation for accepting the results of his clairvoyance. It seems that in accepting the outputs of this process Norman is not being sufficiently intellectually careful. Intellectual carefulness would require that Norman find out more about his clairvoyance before trusting its outputs, or that Norman look for other confirming evidence about the president's whereabouts before coming to the belief that he is in New York. We might also worry about the way in which Norman has developed his clairvoyant capacity. The virtue theorist will require that in order to be a virtue, a character trait must develop through observing and following the actions of others who already possess that virtue. This developmental picture is also missing from Norman's story. There are different possible approaches that Norman could take in order to become justified in his beliefs. He might study the capacity of clairvoyance or look for evidence for the president's whereabouts that is based on his other epistemic virtues. (His choice might depend on whether he is more interested in clairvoyance, or in the president's whereabouts¹⁷.) But without further inquiry his belief is not yet justified.

By looking at the Norman case we can see how virtue epistemology, while offering a historical theory of justification, goes beyond the type of historical requirement offered by

¹⁶ Lawrence Bonjour (1985) p 41.

¹⁷ See Chapter 4 on the context sensitivity of our epistemic virtues based on our epistemic roles.

the reliabilist. The reliabilist looks only at the reliability of the process that produces the belief in question. The virtue theories look at the characteristic motivations that make up an agent's character traits, and look at how those character traits have been developed. We might say that virtue epistemology is an even more historical theory than reliabilism in that it requires an appropriate history in to the development of the capacities, rather than simply looking at their current reliability.

Virtue theory focuses on the development of the character traits that make up a virtue because the concepts of the virtues and the types of character traits that make them up are central to the direction of analysis the virtue theorist offers. If virtue theories were to define virtues in terms of the actions or beliefs that they produce, this would be to return to the traditional direction of analysis. Instead virtue theories start with virtues and the character traits that make them up. These virtues are not reducible to actions or beliefs because they are not codifiable in any general rule about action or belief¹⁸. Because they cannot be learned by learning a set of rules, virtues must be learned in the same way that a skill is learned. Perhaps there are some rules of thumb that one follows early in one's training, but developing a robust skill requires that one can transcend those rules of thumb and tell when it is appropriate to apply them, and when it is appropriate to break them. Thus the fact that virtues are not reducible to acts or beliefs leads us to model their development on the development of skills.

¹⁸ For more on non-codifiability of virtues see chapter 6

Once we have seen that virtues must develop like skills, we can then look at the way the development of the different virtues interacts. Just as we might need one skill to fully develop another, we might need one virtue to fully develop other virtues. It seems plausible that in order to develop the epistemic virtues one will need the moral virtue of honesty, otherwise one might lie to oneself. In order to develop the virtue of generosity one might need to develop the epistemic virtue of intellectual courage, so that one can think of one's own ways of discovering and ameliorating need. Looking at the virtues and the way that they develop as skills, shows us that the virtues will have a reciprocal relationship with each other. This thesis, which follows from the fact that the virtues will have develop like skills, is called the reciprocity of the virtues.

These features of virtue theory, that virtues are developed like skills, and that virtues are reciprocal, can be used to help solve existing problems in epistemology. (We will see the solution that they offer in the last two chapters.) Thus virtue theory will be shown to have an advantage over belief-based theories, in that these useful tools are a direct outgrowth of taking the virtue theoretic direction of analysis.

Virtue Theory should be Internalist

Although the virtue approach to epistemology was first proposed by Ernest Sosa, and his subsequent development of his own virtue epistemology is a version of reliabilism, which is a version of externalism, I will argue that the proper structure for the epistemic virtues is internalist. My argument for this position is contained in Chapter 3, but here I would like to clarify the type of internalism I am arguing in favor of. Traditionally debates about

internalism and externalism in epistemology have focused on justification. Internalism about justification has been defined in many ways, but it has generally held that whether or not a belief is justified depends only on the existence of and interrelations between mental states that are accessible to the agent. These mental states are normally beliefs, since these are the type of mental state that is both accessible to us, and properly structured to serve as justifiers for other beliefs since both have propositional content. Thus internalism about justification has generally held that the only justifiers of our beliefs are other beliefs.

This approach to internalism cannot be easily generalized to apply to an account on which justification depends on the epistemic virtues and their relation to the belief. For epistemic virtues are character traits of the believer, which are not beliefs. Thus if the possession of epistemic virtues can be a justifier we need a broader notion of internalism if we are to preserve the possibility of their being an internalist virtue epistemology. Fortunately John Pollock and Joseph Cruz have developed a taxonomy of types of theories that will be of use here. They make a distinction between doxastic theories and non-doxastic theories.

Doxastic theories are those that hold the following “doxastic assumption”:

If one holds precisely the same beliefs in two possible circumstances, then no matter how those circumstances differ with respect to things other than what one believes, there will be no difference in what beliefs are justified under those circumstances¹⁹.

To hold this assumption is essentially to hold the traditional type of internalism discussed above. But by calling this the doxastic assumption, and the resultant view doxastic internalism, we can call attention to the claim that only beliefs can be justifiers, and this

opens up the possibility of denying this claim to arrive at other forms of internalism. In particular we should consider including in the class of justifiers other cognitive states of the agent, or even other non-cognitive states of that agent. Since we are concerned with epistemic virtue, and since virtues are not discreet states, we should further include those characteristics of a person that are not captured by states at a time, including character traits, and capacities. Thus a non-doxastic internalism about justification will allow justifiers to be any internal state, capacity or character trait of the person for whom the belief is justified.

Since in developing a virtue epistemology, epistemic virtues will be our primary notion from which those of justification and knowledge will be derived we should also give a definition of non-doxastic internalism with respect to epistemic virtues. An epistemic virtue is internalist in this sense if it depends only on the internal states, capacities and character traits of the person who possesses the virtue. Since the developmental history of the virtue is important it should be emphasized that this is not limited to current internal states, but includes all those in the agent's history.

While non-doxastic internalism is a weaker position than doxastic internalism, because it expands the class of potential justifiers, it is still a position significantly different from externalism. For externalism allows justifiers to be not only internal states, capacities, and character traits of an agent, but also the relationships these internal features have to the external world. Thus the externalist may require not only that the agent's internal states be organized correctly, but also that they reliably reflect the truth, or that they track the external

¹⁹ Pollock and Cruz (1999) p22.

truth in particular ways. Thus the debate between non-doxastic internalism and externalism is a substantive one, and one that I will explore further in Chapter 3.

An Overview of My Arguments for Internalist Virtue Epistemology

Before turning to my arguments for internalism in virtue epistemology I will first be concerned to defend the very possibility of doing virtue theory either in epistemology or in ethics. In **Chapter 2** I will address recent arguments to the effect that people do not have stable character traits of the sort that are required for them to have virtues. If the conclusions of these arguments were true it would undermine the very possibility of doing virtue theory. These arguments depend on recent results from psychology, and in discussing them I will discuss both problems with the experiments themselves, and problems with the interpretation that they have been given by philosophers. By uncovering these problems I will show that virtue theories are still a viable option as our theory of normativity in epistemology.

In **Chapter 3**, I will move on to present my arguments for the non-doxastic internalist version of virtue epistemology. I will present a general argument that any version of virtue epistemology that includes externalist requirements will fall prey to counterintuitive results. I will also give an argument for internalism based on preserving a parallel between ethical virtues and moral virtues. Historically moral virtues have been interpreted as also being non-doxastic internalist, because they depend only on the internal states, capacities, and character traits of individuals. Introducing non-doxastic internalist epistemic virtues will allow us to develop epistemic virtues that have the same structure as moral virtues, thus gaining

theoretical simplicity. I will also argue that externalist theories get the source of values wrong—externalism gives virtues only derivative value, while internalism can give them basic value. I will then move on to give an analysis of the more traditional epistemological concepts of justified belief and knowledge. In developing these analyses I will also consider an argument that Linda Zagzebski makes about the general structure of the Gettier problem and how to avoid it. I argue that properly understood this argument undercuts possible remaining motivation for externalism in virtue epistemology.

In **Chapter 4** I explain why the accounts of justified belief and knowledge given in Chapter 3 do not lead us to skepticism about either knowledge or justified belief. I do this by introducing the concept of context sensitivity. Virtue epistemology is sensitive to context, but in a way very different from the mainstream of current contextualist debate in epistemology. I explain why the context sensitivity of a virtue approach, based as it is on the epistemic roles we take on, is superior to both attributor and inferential versions of contextualism.

In Chapters 5 and 6 I move on to applications of internalist virtue epistemology in order to show how it can be a useful tool for solving existing problems in epistemology. In **Chapter 5**, I apply my virtue approach to the problem of how our beliefs based on testimony can be justified. This is a difficult problem because it seems that we are justified in accepting the testimony of those that we know little about. By unifying epistemic and moral virtues under the same analysis, and showing that the epistemic and moral virtues are reciprocal, we are able to discover another source of information about others in their moral behavior. This

connection between the moral and epistemic virtues can then help explain why we are justified in holding beliefs on the basis of the testimony of others.

Finally in **Chapter 6**, I apply my virtue epistemology to explaining how we reason about probabilities. Through looking at some strange cases I will show how important partitioning is in our reasoning about probabilities. I will then set out to explain how it is that we can perform this partitioning so easily in normal cases. Explanations based on Bayesian updating and belief-based theories will be considered, and then I will show why the explanation offered by virtue epistemology, based on the development of a virtue paralleling the deployment of a skill, is preferable to either of these approaches.

CHAPTER 2: ON THE EXISTENCE OF CHARACTER TRAITS

Before arguing for my own version of virtue epistemology, which I will do in the next chapter, I would like to address some recent worries that have been raised against the very possibility of virtue theories. These are concerns that stem from recent psychological studies purporting to show that people do not have the sort of robust character traits that would be needed to be the basis of virtues. Although these objections have been aimed at claims about ethical character traits and ethical virtues, the proponent of epistemological virtues must answer them as well. If there are no ethical character traits to stand as the basis of virtues, then there may well not be epistemological character traits either. In order to develop a virtue epistemology, this possibility must first be addressed. In addition I am particularly concerned to defend the ethical virtues from this assault because I offer an account of the epistemological virtues that puts them on a par with the ethical virtues. To develop this parallel between the ethical and epistemological virtues I also need to defend the ethical virtues from recent attacks.

The Situationist Objection to Virtue Theories

There has been a recent argument that virtue ethics is impossible, because people lack character traits. Since virtue theories require that there be character traits, in order for some of them to be virtuous and others vicious, this would be a blow against the virtue framework. What evidence is there for the audacious claim? Most of the evidence comes from a line of thought from social psychology—situationism. Situationism holds that

people's reactions in a given situation are determined more by the details of the situation than by any underlying character trait. These claims are supported by empirical behavioral studies aimed at comparing the influence of aspects of the situation and character traits. Gilbert Harman (1999, 2000) and John Doris (1998) have both taken the results from social psychology and tried to apply them to the foundations of virtue theories. Harman thinks of character traits as "relatively long-term stable disposition[s] to act in distinctive ways"²⁰.

Turning to the empirical studies in this area he concludes that:

It seems that ordinary attributions of character traits to people are often misguided, and it may even be the case that there is no such thing as character, no ordinary character traits of the sort people think there are, none of the usual moral virtues and vices... Since it is possible to explain our ordinary belief in character traits as deriving from certain illusions, we must conclude that there is no empirical basis for the existence of character traits²¹

These conclusions, plus considerations about how the concept of virtue and character traits might be misused, lead Harman to conclude that we ought to abandon virtue approaches in ethics.²²

Doris takes the empirical work in this area to have a less radical import, but he still thinks that it shows virtue ethics as it is currently conceived to be misguided. Doris' conclusion from the situationist literature is that there are no broad based character traits which would lead to similar behavior across a wide range of different situations. However since it has

²⁰ Harman (1999), p 317

²¹ Harman (1999), p 316.

discovered regularities in behavior in an individual in very similar situations across time, situationism does not entirely get rid of character traits. Rather, according to Doris:

Situationism allows that a suitable fine-grained inventory of local traits may provide an account of personality that is both empirically adequate and theoretically useful.²³

Although this is a moderate conclusion Doris thinks that it still leads to radical changes in virtue theory, because it eliminates the broad based traits on which virtue ethicists have focused. This means that while virtue ethics might be saved, its approach to the character traits that are the basis of moral evaluation would be quite different.

What is the empirical evidence on the basis of which Harman and Doris come to their conclusions? Much of it is taken from Ross and Nisbet (1991) a seminal work in situationism, and a collection of many empirical studies about the role of situation in behavior. Of the cases considered I will focus on two, which are taken as central examples of empirical support of situationism. By looking carefully at the details of the experiments, and by comparing the concepts used in the experiment with those used in virtue theory, we will see that these experiments are far less damning to virtue theories than originally thought.

The first experiment was carried out on Princeton Theological Seminary students. They were put in a situation where they might help someone in need, a person sitting a doorway coughing and groaning. The experimenters looked for correlation between subjects's

²² Harman (1999), p 327-9 and (2000), p 224

²³ Doris (1998), p 508.

helping behavior and three other variables. The first, intended to capture a character trait, was the basis of their religious belief. They were categorized by their own classification of themselves as being either interested in religion as a means (e.g. to salvation) or as an end (e.g. as a quest for meaning in life). The second was a situational variable. Some of the subjects were asked to read the parable of the Good Samaritan while the others read a text about vocational alternatives to ministry. The third was also a situational variable. The students were asked to give a talk, and sent to the next building to do so. Some were told that they were already late, some that they were on time, and some that they had time to spare. Experimental results showed only an insignificant correlation between helping and either the character trait or the situational variable of the text read. The only variable that was correlated with helping the person in distress was the amount of time the subjects thought they had. Those with more time were more likely to help than those who thought that they were already late. The experimenters conclusion was that the subject's response was determined not by character but by situation²⁴.

The second experiment is really a collection of experiments intended to identify the character trait of honesty across a number of different contexts. Subjects in these experiments were children and they were put in situations to tempt them to be dishonest. These situations included the opportunity to cheat on tests, on homework, in athletic competition and in party games, as well as to steal money, or to lie about their own behavior or the behavior of others. All these behaviors were taken to be instances of the general trait of dishonesty, and the experimenters were looking for correlations between these behavior

²⁴ Darley and Batson (1973)

types. The correlation between behaviors of the same type (eg. the probability that a child who cheated on a test would cheat again) was quite high. However the correlation between behaviors of different types (eg. the probability that a child who cheated in the test would take the money) was low.²⁵ Situationists conclude from this experimental evidence that if there are stable dispositions, they are quite specific, not nearly as general as the character traits that we assume in our everyday discourse.

These experiments, and others like them, are taken by situationists to show that it is the situation, rather than any character trait, that drives behavior in these circumstances. Indeed, the lack of correlation of behavior across different situations can be taken as evidence that behavior is not controlled by anything like a character trait. This conclusion will have negative consequences for virtue theories in so far as they appeal to character traits to explain the virtues, and to give an account of what the good person is like. Before jumping to a negative conclusion I think that we should be more careful in our interpretation of these experiments. There are two types of reasons to doubt the standard interpretation of these experimental findings. The first has to do with the methodology of the experiments and the way that they connect with our concept of character traits. The second has to do with our understanding of what a virtue really is. Once we understand what the nature of the virtues offered by a (classical) virtue theory, we will see that the experimental results just considered do little to tell against this conception.

²⁵ Hartshorne and May (1928).

Objections to the Methodology of These Studies

In these studies, the experimenters code the subject's behavior as, for example, being an instance of helping (the person in the doorway) or not. In order to do this the experimenters operationalize the character trait in question; they give a definition of what types of behaviors will be classified as acting according to the trait, and which count as acting against the trait. The way that a character trait is operationalized will affect the way that the behavior is understood, and the way that the resultant data emerges. Gopal Sreenivasan, in defending the virtue theorist, has given a set of what he calls "generic requirements a behavioral measure must satisfy in order to properly operationalize a character trait." They are:

- (i) Each behavioral measure must specify a response that represents a central or *paradigm case* of what that trait requires;
- (ii) the concrete situation each specified must *not* have any features that *defeat the reason* on account of which that trait requires the response in question; and
- (iii) the *subject and the observer must agree* on these characterizations of the specified responses and situations.²⁶

Why are these generic requirements an important constraint on a behavioral measure? To see this let us turn to parts of the experiments above that violate these three requirements, and show how they undermine the conclusion that the subjects lack character traits.

First let us consider the requirement that the response in question be a *paradigm case* of an application of that trait. In the Hartshorne and May experiments, subjects were given the

opportunity to lie about the behavior of others, and whether those others had cheated. If the other subject had cheated, then telling the truth would be to “rat them out” and lying would be to protect them. It is quite plausible to claim that these are not paradigm cases of honesty. Many people may disagree about whether honesty requires telling on others, or whether this type of lie is acceptable. Thus it is an open question whether lying in this case should be counted as a lack of required honesty. We might think that a person is honest, even if she does not tell the truth in situation.

Second let us move on to the issue of *defeating reasons*²⁷. In order to demonstrate that someone has acted contrary to the requirement of a character trait or a virtue, it is not sufficient to consider a case where there is a defeating reason not to act in accordance with that virtue. For an example we can go back to the classic case in the historical debate about justice, which begins in Plato’s Republic. Although we may consider it a requirement of justice to return to people those things they have left in our keeping, we do not think that justice requires returning a weapon to a friend who has gone crazy. This is because the requirement to return the weapon is defeated by a larger concern about the well being of others. Similarly in the cases given in the experiments. For example, the subjects in the Good Samaritan experiment had two competing claims upon them—the requirement to help those in need, and the requirement to give a speech one has promised to give. Even if we do not think that this defeating reason is sufficient to trump the requirement to assist those in need, we should allow that the subjects may have thought (at least at the time) that

²⁶ Sreenivasan (2002), pp. 61-2.

it was a defeating reason. Thus the subjects in these experiments may well have had the character traits in question, but acted differently because of competing claims²⁸.

Finally let us consider the requirement that the subject and experimenter agree about their characterization of the responses. This is a crucial requirement, because we are looking for cases where people act in a way that they think is discordant with the requirements of a particular virtue. It does not show that someone is dishonest if they act in a way that they perceive as being the honest thing to do in the situation. Sreenivasan gives the example of a subject who is honest, but who also holds that “finders keepers”²⁹. In such a situation, the children who are subjects in a set-up intended to get them to steal, may think that it is entirely in keeping with the requirements of honesty, and not stealing, to take stray change. If the subject does not think that this is dishonest, then we have not found a situation when details of the context have made him act out of character.

Thus we can see how the details of the experiments considered do not meet the three requirements on operationalizing a character trait. As a result we cannot conclude from these experiments that the subject lacked stable character traits. For their behavior that is out of keeping with those character traits can be explained by defeating (or at least conflicting claims) or by their own alternate interpretation of the situation.

²⁷ For more on defeating reasons and the way that defeaters can work in epistemology see Pollock (1990) and (2001).

²⁸ In the debriefing for this experiment, many of those who did not help the person in need reported feeling a conflict between their duties to the victim and their duties to give to talk.

²⁹ Sreenivasan (2002), p59.

I have only raised this structural objection to two of the numerous experiments that are considered by situationists. Rather than working through an incredibly long list of experiments one by one, it is more fruitful to consider the general characteristics of these experiments that make them unlikely to meet the three requirements on operationalizing a character trait. When constructing these experiments, experimenters want to choose situations in which each of the available options is taken by some subjects. If they considered cases in which subjects were invited to steal someone's wallet at gunpoint and those responses were compared with subjects' willingness to lie in a court of law, they would get near perfect correlation of the trait of honesty in both cases. In order to make a lack of correlation possible, the experimenters need to choose tricky cases, ones that pull a subject in both directions. But these cases, by their very nature are going to be ones in which there are competing (possibly defeating) concerns. Similarly they will not be paradigm cases of lacking a virtue³⁰. Thus it will be difficult to get cases that both provide the variation in response needed to prove situationism, and which are central and unproblematic enough to clearly count against character traits. This is not to say that it is impossible to get such cases; if that were true then the claim that we have character traits would be unfalsifiable. Rather it will be a difficult matter to choose appropriate cases to use in experiments aimed at proving the situationist position.

³⁰ It is important to notice that situationists are only partially concerned with cases of ethical character traits. They are concerned with character traits of all types. In the case of non-moral character traits (e.g. being outgoing or shy) it may be easier to get variation among subjects even in paradigm cases.

Objections to the Claim That These Studies Show There is No Virtue

The situationist literature is aimed at proving claims about the lack of influence of character traits on behavior. As a claim in social psychology, it is not directly concerned with virtues. However both Harman and Doris take the situationist experiment to tell not just against a role for character traits, but also against a role for virtues. This is because both think that character traits, as understood by the situationist are relevantly similar to the virtues. But are they? Both Julia Annas (ms) and Rachana Kamtekar (forthcoming) argue that the ancient conception of virtue is quite different from the conception of character traits used in this empirical research. Recall that Harman thinks of character traits as “relatively long-term stable disposition[s] to act in distinctive ways”³¹. This means that character traits have a direct connection with action. This assumption is played out in the interpretation offered of the empirical evidence. Behavior is taken as evidence of dispositions and hence of character. If one has a particular character we expect to see it give rise to the relevant behavior. Annas thinks that situationists expect this tight connection because they start with a “notion of character as an unreflective habit of stereotypical response”³².

But this notion of a character trait does not fit well with the conception of the virtues endorsed by virtue ethicists. Annas summarizes this conception (which is similar across a number of ancient theorists) as holding that:

[V]irtue is not a habit in the sense in which habits can be mindless; it is not a source of actions within the agent which bypasses the agent’s practical

³¹ Harman (1999), p 317

³² Annas (ms), p 9.

reasoning... This is because a virtue, unlike a mere habit, is a disposition to act which is exercised in and through an agent's practical reasoning.³³

Thus virtue has a much more complex and subtle relation to behavior than the type of character traits situationists think of. Since the dispositions of virtue are inputs to practical reasoning, they can be sensitive to details of the situation, and especially to competing concerns. As with Sreenivasan's earlier concern, we may explain variation in behavior by the presence of defeating reasons. Even someone with a robust virtue does not always act in the ways that are stereotypically associated with that virtue, because in particular situations other concerns, perhaps motivated by other virtues, may be more important. In this way the behaviors of the virtuous person are appropriately situation-sensitive, but this sensitivity to situation does not mean that the agent lacks a virtue. Thus sensitivity to details of a situation is compatible with, even required for, the possession of virtue. The situationist experimental evidence is thus not incompatible with a virtue ethics.

Epistemic Virtues?

Although experiments like those above tell against ethical character traits, we might also be concerned with the possibility that a parallel argument could be made for epistemological character traits. Cases like those above aim to demonstrate that we do not have the underlying dispositions to act well that we think we do. But do we have underlying dispositions to think well? Although not a central part of the situationist literature in social psychology, there are some experiments that might be interpreted to tell against the existence of epistemological character traits. Consider the study in which subjects were asked to judge

³³ Annas (ms) p 10.

line lengths when visually presented next to each other³⁴. This was a fairly simple task in which the control subjects could easily identify which line was equal in length to the test line 99% of the time. However the test subjects were placed in a room with other apparent subjects who were really conspirators. These conspirators incorrectly identified the lines of equal length to the test line, sometimes in extremely obvious cases. After hearing these misidentifications, many subjects agreed with the misidentification— some even incorrectly identified the equal length line 50% of the time. Here it seems that in a specific situation, subjects are acting out of keeping with their epistemic character traits. They are perfectly capable of correctly identifying equal length line, and yet they do so incorrectly. We might then say that their misidentifications are not driven by an underlying epistemological character trait, but only by the details of the situation.

However I think that there are two different ways to respond to cases like these without moving to a situationist position. The first possible response is to doubt that the subject's reports of their own beliefs correctly reflect their beliefs. After all, in social situations our belief reports are frequently tempered by what we think is polite or socially acceptable. For example, if you are discussing a movie after having exited the theater, and all of the people you are talking with express negative opinions, you might be tempted to express a negative opinion too, even if your real opinion of the movie is positive. It seems obvious to me that this effect is more pronounced in situations where there are numerous others, where there is unanimity of opinion, or where the others' expressions of their opinions is strongly held. One might respond to the social situation by agreeing with others, and even by making

³⁴ Asch, S. E. (1956).

claims about the movie that she thought were false, simply in order not to “rock the boat”. Yet this behavior is perfectly consistent with really believing that the movie was a good one. One possible explanation is that the same type of phenomenon is appearing in the experimental set up. Notice that many elements that might lead people to polite misreporting are present in these experiments. There are a number of others in the room, and all of them misidentify the shorter line as longer. The others also express their opinions confidently. Thus we might object that while the subject reports that they believe a falsehood, they might not believe it at all, secretly believing the truth.

A second type of response to the experimental findings might be to claim that they are perfectly compatible with the existence of epistemological character traits, even if the subjects really do change their minds and agree with the conspirators. Rather than simply being swayed by the situations they find themselves in, the subjects may be responding appropriately as epistemic agents to evidence that the situation is providing them. Presumably epistemically virtuous agents ought to possess a certain amount of epistemic humility. Even though they trust their senses, and do so with good reason, they are also sensitive to evidence that their senses may not be working properly. The reports of the confederates, who are in perceptual circumstances that are just as good as the subject’s own, is some evidence that the subject’s perceptual mechanism is not working correctly. If enough others give evidence that contradicts one’s perceptual beliefs it might be reasonable to doubt or even recant on one’s original belief. This is particularly so when the epistemic agent has no reason to believe that the others are attempting to influence her with their reports. Thus even a change in belief is not inconsistent with the subject having constant

epistemic character traits that respond differently to different evidence available in different situations.

Empirical Evidence Favoring a Virtue Approach to Epistemology

Having considered some empirical evidence that has been taken to tell against virtue approaches, and having shown while these experiments do not undermine virtue theory, I would like to move on to consider some empirical evidence that shows the advantages of a virtue approach in epistemology. Ever since the modern era epistemologists have considered reason and emotion to be very different things, and have advocated epistemological theories favoring the use of reason alone. Emotion is thought to be only a possible interference with the process of reason, not in any way providing assistance to reason. However recent studies of the neurological processing of information have shown that emotion is a necessary component of many reasoning processes. Here I will focus on evidence that the calculation of probabilities from experience depends on feeling emotional responses.

Angela Sirigu et al. (2004) have run an experiment in which they studied the decision making processes of patients with lesions in the orbitofrontal cortex and compared it to that of normal control subjects. The characteristic disabilities of patient with orbitofrontal lesions are an inability to experience emotion—essentially the messages from the brain to the body which result in the physical responses of emotion are interrupted. These patients are also characterized as not having lost any abstract processing capacity. Therefore they are perfect

subjects in which to consider the role of emotion in processing information. Both sets of subjects participate in an instance of decision making under risk. Subjects were asked to draw cards from two different decks. The cards had numerical values representing either the prize they received for drawing that card, or the penalty for drawing that card. The subjects drew from these decks and attempted to figure out which deck gave them better chances of more and larger reward and few and smaller penalties.

Subjects in these experiments were also connected to a device that measures skin conductivity response. When normal subjects are nervous, their skin conductivity increases. This is not true for the orbitofrontal patients who show little change in this measure. Earlier experiments have shown that normal subjects in situations like these exhibit skin conductivity responses before they draw from the “bad” deck. Surprisingly, they exhibit these responses before they can report which deck is the “bad” deck. In Sirigu’s experiment subjects were asked to draw cards, and to report their level of regret after having drawn each card. This was then compared to their skin conductivity responses. The normal subjects’ skin conductivity response (after drawing a card) was closely correlated with their reported regret about their choice. It seemed that patients used this experienced regret to isolate the “good” deck, and fairly quickly began drawing only from it. The orbitofrontal patients also had high correlation between experienced regret and skin conductivity, but only because they had very little of either. This lack of emotion response hurt their ability to learn which deck was the good deck. It took these subjects much longer to isolate the “good” deck and begin drawing exclusively from it. This result suggests that the experience of emotion, in this case

regret, is an important component of reasoning. Without it, subjects were much worse at figuring out the probabilities that the deck they were drawing from would give them a prize rather than a penalty.

This result goes contrary to the traditional assumption that emotions can only serve to interfere with reasoning, not to assist it. The results of this and similar studies show that an epistemology that captures how humans reason, and should reason, ought to include a role for emotions within the epistemological enterprise. Virtue epistemology has an advantage in this respect because it already includes a role for emotion. For a having a virtue gives us a characteristic motivation to act in particular ways. In the case of moral virtues, this might be the feeling of pride, or love of those we are protecting, which leads us to act courageously. In the case of epistemic virtues this might be a feeling of the weight of the responsibility on our shoulders in a particular instance which leads us to be epistemically careful. Because virtues have a motivational component built in, they automatically make a role for out affective responses to play within epistemology. This is not to say that non-virtue theories cannot make room for emotions within their theories. But to do so would require some radical changes from the traditional versions of non-virtue epistemologies.

Chapter 3: The Proper Structure of the Epistemic Virtues

In this chapter I intend to lay out what I think is the proper structure for the virtues needed in a full account of virtue epistemology. I will begin this development by considering a thought experiment, the new evil demon problem, which will help us to see the importance of internalism generally. I will then consider the uses of epistemic virtue put forth by two prominent externalist virtue epistemological theories, those proposed by Alvin Goldman and Ernest Sosa, and the purposes to which they put the idea of an epistemic virtue. In each case the idea of an epistemic virtue is introduced to help the theory amend its counterintuitive results in cases like the new evil demon problem. However in both cases the use of the epistemic virtues does not fully solve the intended problem, because it does not allow the theory to agree with our intuitions. It merely gives us a way to explain the intuitions away. Thus both theories gain little from the use of epistemic virtues to help with counterintuitive cases.

After looking at purely externalist versions of virtue epistemology, I will then move on to discuss the mixed account of the virtues put forward by Linda Zagzebski, in *Virtues of the Mind*. Her epistemic virtues contain both internalist and externalist elements, and as a result of the way that her theory runs afoul of the new evil demon problem, we will be able to see the general problem for any account of the epistemic virtues that contains an external element.

Having explained why the new evil demon problem will affect all virtue theories with an externalist element, I will move on to give two more arguments in favor of taking an internalist approach to the epistemic virtues: a historical argument, and an argument from the sources of value. Both of these arguments apply equally well to epistemic and moral virtues. This is a good result because in the end I want to follow Zagzebski at least in her overall goal of giving a unified treatment to epistemic and moral virtues.

My historical argument is based on the ancient Greek conception of a virtue. I will argue that the internalist version of the virtues fits much better with the ancient conception of virtue. The final argument is based on the source of value in an account of the virtues. Here I will argue that an externalist virtue theory gives up on the idea that the virtues are independently valuable, and makes that value of the virtues a merely instrumental value. This I think undermines the appeal of a virtue theory, and reduces virtue theory to a type of rule consequentialism. I propose to take the difference between these types of theories seriously, and to offer an internalist account of virtue that respects this difference.

Having given these arguments for an internalist conception of the epistemic virtues, I will return to more the traditional epistemic topics of justification and knowledge. I argue that a proper understanding of the relationship between epistemic virtue, justification, and knowledge serves to undermine one possible motivation that we might have for externalism in our accounts of epistemic virtue and justification. A proper understanding of this relation allows us to give a purely internalist definition of virtue in general, epistemic virtue, and

justification, while at the same time preserving the important externalist elements needed for knowledge.

The New Evil Demon Problem

Consider two agents, Prudence and Faith, who are, unfortunately, living in a world where evil demons systematically deceive all humans³⁵. Although their experiences of their world are coherent and consistent just like ours (and may in fact be just like ours), those experiences do not correctly represent the world in which they are living. Prudence is trying hard to be a responsible epistemic agent, and forms her beliefs on the basis of careful observation, diligent research and open-minded enquiry. She is careful in considering her evidence, never jumps to conclusions, searches for confirming evidence where appropriate, and so on. She seems to be the model of a good epistemic agent, but yet because of the deceiving demons, none of her beliefs are true, and none of the processes that she uses are reliable. Faith on the other hand is irresponsible in her belief formation, jumping to conclusions, following wishful thinking, and close-mindedly refusing to consider objections. Faith's beliefs are also all false because of the influence of the deceiving demons. It seems intuitive that Prudence is virtuous epistemic agent and that her beliefs are justified, while Faith is not a virtuous epistemic agent, and her beliefs are not justified.

The New Evil demon problem is a problem for reliabilism because both Prudence and Faith have the same level of reliability. They are both wrong in all their beliefs and so are both

perfectly unreliable. On a reliabilist theory then, both Prudence and Faith are not justified. Furthermore, since for the reliabilist justification depends on reliability alone, Prudence is not any more justified than Faith. They are both equally unjustified in their beliefs. But intuitively we feel that Prudence is a better epistemic agent than Faith, and hence is more justified than Faith in her (admittedly false) beliefs. So reliabilism runs afoul of our intuitions in the new evil demon scenario.

Goldman's Use of the Virtues in his Two-Stage Reliabilism

Goldman's updated version of reliabilism, as presented in "Epistemic Folkways and Scientific Epistemology" (1996) incorporates the idea of epistemic virtues in order both to respect the fact that we use intellectual virtues in evaluating others, and to show how reliabilism can accommodate our intuitions about troublesome cases like the new evil demon problem.

Let us look at the use of the epistemic virtues offered by Goldman. Before presenting his position he is careful to make a distinction between the descriptive and normative jobs of epistemology³⁶. The descriptive element that he focuses on is not our actual cognition, but rather our actual practice of evaluating our epistemic situations. Science can be relevant to this pursuit in the study of our current linguistic practices and usage of our epistemic evaluation terms. Our descriptive study then can give us a good idea of the folk notions of

³⁵ This scenario was first presented by Lehrer and Cohen (1985) and was given the label "the new evil demon problem" for reliabilism by Sosa (1991). The names that I use here are an addition by Henderson and Horgan (forthcoming). Lehrer applies the new evil demon problem to the issue of epistemic virtue in his (2001).

³⁶ Goldman (1996) 423-424.

justification and knowledge, as well as of good epistemic practice in general. This descriptive element will then serve as a foundation for normative epistemology. Goldman argues that our normative epistemology cannot stray too far from our folk epistemology; otherwise we risk simply changing the subject and not leaving ourselves with a unified phenomenon to study. However normative epistemology need not simply reiterate the folk conception; it can embody substantive changes, just not too many of them.

With this distinction in hand, Goldman begins to describe the role of the epistemic virtues in our folk epistemology. It is here that the two-stage structure becomes apparent. In ordinary cases, we reach our judgments about justification by comparing the instance at hand with a list of what we consider to be epistemic virtues. If the process used in this instance falls clearly under one of the virtues, or if the methods used in generating the belief are sufficiently similar to some method on our list of epistemic virtues, then we judge that the belief is justified. This is the second stage of Goldman's two-stage reliabilism. Thus our epistemic judgments depend heavily on our concepts of the epistemic virtues. Where do we get these concepts? This is first of the two stages of two-stage reliabilism. We come to believe that certain processes are virtuous because we believe that they are reliable³⁷. The process of coming to judge processes as virtues is not as straightforward as a mere reliability measure though. Lacking information about reliability of our methods, we stumble along, refining our notion of reliable, and hence virtuous, processes. Society helps in this process. Rather than each of us figuring out a list of virtues on his or her own, individuals can count on the community to share with them concepts of virtuous processes that have been refined

over time and by large numbers of people. In this way we come to share most of our concepts of virtuous processes with each other, and such concepts become useful for communication.

By presenting us with this two stage process, Goldman seems to be able to reconcile his own notion of reliabilism with our folk notion of epistemic virtues, as well as with our folk intuitions about cases. Probably the biggest benefit of such a picture is its ability to explain cases that have long been a problem for the reliabilist. One example of this is the way that two-stage reliabilism can explain our intuitions in the new evil demon case. When we consider Prudence's belief forming processes, we find that are very similar to, or are perhaps included in, the list of processes that we think of as virtuous. Thus we classify Prudence's behavior as virtuous, and hence are tempted to call her resultant beliefs justified. When we consider Faith's belief forming processes, we have the opposite reaction, noticing the similarities between her processes and those that we classify as vices. Thus we classify Faith's beliefs as unjustified. Of course, in both cases the processes that the agents use in their epistemically unfortunate world are not reliable. This is where the two-stage structure of our folk epistemology plays a role. Although we are to consider the reliability of a process when deciding whether to add it to our list of virtues, we don't consider the reliability of a process when we are applying our list of virtues. At the second stage of evaluation, that of actual application, it is only being on the list of virtues that matters, not reliability.

³⁷ Goldman (1996) 425.

Thus Goldman can handle the new evil demon case by explaining the source of our intuitions that agents in a demon world can be differentially justified. However it is important to notice the limitations of this approach. Goldman can explain why we have this intuition, but not why this intuition is correct. That is, he cannot tell us why the irresponsible agent is *really* less justified than the responsible agent, but only *why it seems to us* that she is. Yes, by looking at the descriptive, folk-epistemology we can see why people would call Prudence virtuous and her beliefs justified. But the normative epistemology of reliabilism holds that only beliefs generated by a process that is actually reliable are justified, and so the normative theory must insist that Prudence is not really justified. Thus Goldman is stuck saying that the intuition that Prudence is justified is importantly wrong. And thus Goldman cannot use his appeal to virtue to solve his original problem, which was to accommodate our intuitions in the new evil demon scenario into his theory.

Sosa's Contextualist Reliabilism and the New Evil Demon Problem

Sosa embraces a version of virtue epistemology that, like Goldman's focuses on reliability. Taking reliability as the measure of virtues means that Sosa faces the same problems that reliabilism has faced. These are cases where reliabilism seems to give us unintuitive conclusions about justification, and include the new evil demon case. Just as Goldman introduces his two stage reliabilism to address these cases, Sosa also offers his own version of reliabilism that is aimed at getting around these hard cases. He suggests that reliabilism should be more sensitive to the contexts in which an agent might be justified³⁸. Subtleties of this context can affect the reliability of a process and as a result can effect whether or not an

agent is justified. Sosa's suggestion is to relativize epistemic virtue to a particular environment. This is a partial solution to the generality problem for reliabilism. Rather than asking about the reliability of a process in general, we can ask about the reliability of a process in the relevant environment. Sosa defines intellectual virtue as follows:

A subject *S*'s intellectual virtue *V* relative to an "environment" *E* may be defined as *S*'s disposition to believe correctly propositions in the field *F* relative to which *S* stands in conditions *C*, in "environment" *E*³⁹.

This definition of intellectual virtue allows that we can assign intellectual virtue to an agent either relative to her actual environment, or, more importantly, relative to another environment that she could be in. Thus we can talk about an agent's conditional intellectual virtue, so long as she is conditionally reliable. Even if she is not currently in environment *E*', it might be the case that if she were in environment *E*' she would be reliable. In this case we can say that the agent is intellectually virtuous with respect to environment *E*'.

Sosa's conditional reliabilist position gives us the mechanisms needed for his solution to the new evil demon problem. We should note that while Prudence is not reliable with respect to her actual, demon controlled environment, she is reliable with respect to the environment she assumes that she is in. Absent any evidence that she is in an abnormal environment, since her experiences are coherent, Prudence can assume that she is in a normal

³⁸ Sosa (1991) 135-142.

³⁹ Sosa (1991) p140.

environment, and relative to that environment her processes are reliable. Thus when we compare Prudence to an epistemically virtuous agent in a more cooperative world, we find that they are both responding in the same way. Both, having no evidence to the contrary assume that they are in a normal environment, and both believe that their processes are reliable in that environment. Given this parallel, we might hold that, if the responsible agent in the good world is justified, then Prudence must be justified as well.

But of course all this shows us is that the two agents believe that they have the same reliability. What reliability they actually have is another matter. Sosa seems to allow us to focus only on the conditional reliability based on the environment that the agent thinks she is in. Although it might be acceptable for Prudence to assume she is in a normal environment, it is not safe for us to assume that—we know, by stipulation that she is in a demon environment. Thus we need to consider Prudence's reliability in an environment like her actual one, an environment where she is persistently misled. Sosa has given us an explanation of why Prudence might think she is justified, but he has not given us an explanation of why we think she is justified. She is not intellectually virtuous with respect to her actual environment, and we know this.

Thus it seems that relativizing intellectual virtue to an environment or context is not an adequate solution to the new evil demon problem. We want to say that Prudence is intellectually virtuous, even in her problematic environment. It is not just that she *would* be virtuous if she were in a better circumstance; she *is* virtuous in a bad circumstance.

Even though Prudence's virtue does not lead to success, as it would normally, she still has the virtue. So Sosa's proposal does not explain our intuitions that Prudence is both virtuous and justified. His use of the intellectual virtues does not help him to answer problems within his own theory, and thus does not offer us much hope as an approach to virtue epistemology in general.

Zagzebski's Virtues and the Generality of the New Evil Demon Problem

Zagzebski is concerned to give a proper structure of the virtues, both epistemic and moral. One of her goals is to give an account of the virtues that applies equally well to the epistemic and the moral virtues. By giving a universal account of the virtues, Zagzebski is able to claim that all virtues should be treated the same way. To this end Zagzebski offer the following general definition of a virtue:

A **virtue** then can be defined as **a deep and enduring acquired excellence of a person, involving a characteristic motivation to produce a certain desired end and reliable success in bringing about that end**⁴⁰.

Zagzebski breaks this definition down into what she calls the motivational component of virtue and the success component of virtue. The motivational component of virtue is the requirement that the agent have the right kinds of motivations, not just in some particular instances but regularly. In order to have a virtue an agent must possess a regular disposition

to have the right kinds of motivations. Zagzebski thinks of the virtues as identifiable by their characteristic motivations. Thus the virtue of courage is identified with the motivation to protect others and oneself. The virtue of charity is identified with the motivation to help others. This motivational element is an entirely internalist element, concerning the internal states of the believer. This stands in contrast to the to the element of reliable success required by the externalist theories of virtue we have just considered.

Zagzebski does not content herself with the motivational component though. She also requires that a character trait be reliable in bringing about success, and hence adds an element of externalism as well. Zagzebski tries to motivate the inclusion of a success component by arguing that the term “virtue” is a success term⁴¹. She suggests that were we to learn that a characteristic previously thought to be a virtue, was not successful in reaching the desired ends characteristic of the motivation, we would stop calling it a virtue. Thus having success in reaching the end of the motivation is an essential element in having a virtue. Zagzebski moderates this view by requiring only *reliable* success in reaching the ends of our motivation. She thinks that while we would change our opinion about a putative virtue’s being a virtue if we were to discover that it was not reliably successful in reaching its ends, isolated individual failures to reach success would not change our mind about a character trait being a virtue. A character trait might still be a virtue even if it fails to succeed in certain isolated circumstances, so long as it is reliably successful.

⁴⁰ Zagzebski (1996) p 134. Emphasis in original.

⁴¹ Zagzebski (1996) p. 135-136.

By requiring that a virtue be reliably successful, Zagzebski has built a strongly externalist element into her account of virtue. This externalist component comes from the elements of reliabilism that Zagzebski adopts. The motivational component of a virtue is entirely internal to the agent. An agent can be motivated to act in a particular way even if circumstances stop her from so acting. The motivational component of virtue then depends only on elements within the agent. The success component of virtue depends on elements outside of the agent. In order to guarantee success, even only reliable success, the world must cooperate with the agent who is acting. If the agent is confused or misled about the way the world is, the agent may act from the right motivations, and still not have success. If this failure to obtain success is systematic, and hence she is not reliably successful, then the agent will not count as having a virtue.

We can see why this element of externalism poses a problem in the case of Zagzebski's virtues by returning to the new evil demon scenario. As with the reliabilist theories discussed earlier, Zagzebski's structure of the virtues clashes with our intuition that Prudence is a virtuous epistemic agent. To make contact with Zagzebski's account let us clarify the new evil demon case by adding that Prudence is moved by a motivation to find the truth, and this motivation gives her further motivations, say to be open-minded. However, since she is under the control of the evil demon, her beliefs arising out of the character traits we would normally call virtues are quite unreliable. Since her motivation for truth is quite reliably unsuccessful, and since Zagzebski claims that virtue is a success term, we are stuck saying that Prudence's admirable character traits like open-mindedness are not virtues. This goes contrary to our intuitions in the case.

We can see now that the problem for Zagzebski is a general one, which will affect any approach to the structure of intellectual which takes external success, or reliable success, to be a requirement for virtue. Even if, as with Zagzebski's view, these approaches require more than reliability for an agent to have an intellectual virtue, the mere inclusion of a requirement of reliability will leave us with cases in which an agent seems to be doing all she can to be intellectually virtuous, and yet fails to have that virtue. This result clashes with our intuitions from the new evil demon problem because it rules out the very possibility of there being a virtuous agent in a bad circumstance.

Implications of theories of virtue that contain an externalist element also clash with our intuitions in moral situations with a similar structure. Remembering that Zagzebski and others who are moved to virtue to explain the normativity of epistemology want to give a unified account of virtues in general, we can highlight the problem of requiring success in moral cases as well. Consider the case of Grace, who makes a large contribution to a local charity. Even though it means curtailing some of her own spending, she makes this donation because she is motivated to alleviate the suffering of others, and she frequently feels, and acts on, this type of motivation. Unfortunately, Grace's chosen charity is a well-hidden scam, and the money she donates to it never ends up helping anyone in need. So long as she would make the same type of donation in similar situations, Grace cannot be said to be reliably successful in bringing about the ends of her motivation. While she is motivated to help others, she does not do so reliably. Any virtue theory with a reliable success component will be forced to say that Grace does not have the virtue of generosity.

But this runs counter to our intuitions that Grace is indeed generous, even if, through the deception of others, she does not succeed in achieving the ends she is motivated to achieve. Thus the reliable success component yields results that clash with our intuitions about moral virtues as well.

The Historical Argument for Internalism

The reliable success component also clashes with the concept of virtue common to most of the ancient Greek originators of virtue theories. In her article “Structure of Virtue” (2003), Annas compares the structure of virtue as it appears in ancient virtue ethics, and as it has appeared in recent virtue epistemology, especially in the version of virtue epistemology offered up by Zagzebski. As a result of these comparisons, Annas concludes that virtue as used in ancient accounts differs significantly from the account of virtue needed in contemporary virtue epistemology. While I agree that there are striking differences between ancient virtue and Zagzebski’s modern version, I would argue that the difference is simply a result of the externalist reliable success component present in Zagzebski’s account. A more internalist approach to virtue, which moves away from requiring reliable success, would have a structure that more closely agrees with the ancient conception of virtue.

One element that Annas focuses on as a difference between virtues ancient and new is Zagzebski’s insistence that virtue is a success notion. Annas introduces the distinction between our *telos* or overall aim, and our *skopos*, or immediate aim, and uses this distinction

to clarify the claim that virtue is a success notion⁴². She claims that the general ancient view requires a virtue to succeed in its *telos*, or more specifically, that all the virtues acting together will ensure the success of our reaching our overall *telos*, which is living a good life for us, as well as in having the individual virtues which together constitute the good life. While on the way to achieving this *telos* though, we rightly decide to pursue many immediate goals, many different *skopoi*. Annas's example is that of a person trying to rescue those in a burning building⁴³. The agent's virtue of courage manifests itself in a particular instance as a decision and an attempt to save the people in the burning building. Thus saving those people becomes a *skopos*. However the virtuous agent may, through no fault of her own, fail to save the people—e.g. someone may shoot them as she pulls them from the building. If this were to happen the agent would (or at least could) still succeed in obtaining her *telos*, in this case having the virtue of courage, even though she has failed to obtain a specific *skopos*. One reason that the *telos* can be obtained without the *skopos*, is that success in reaching the *skopos* is not always up to the agent, while success in reaching the *telos* is up to the agent. While we desire success in reaching the *skopos*, on a virtue approach, we place more emphasis on the long term character of the agent, and so focus instead primarily on success in achieving the *telos*.

It is here that we have some tension between Aristotle, and most other ancient philosophers, especially the Stoics. While Aristotle claims that virtue is praiseworthy in its own right, he also, when discussing certain virtues seems to require that virtuous agents achieve their *skopoi*

⁴² Annas (2003) p 24.

⁴³ Annas (2003) p 25.

as well. An example of such a virtue is magnificence, the virtue of making appropriate choice in large civic projects. Such a virtue requires both the prior possession of wealth to make such a large scale project, and actual success in spending and planning appropriately, so as to impress without being ostentatious. But, Annas points out, including of this type of a virtue, which requires success in obtaining the *skopos* is inconsistent for Aristotle⁴⁴. For he claims that none of the virtues can depend on moral luck. Yet one needs luck both to acquire the riches needed for civic projects, and to succeed in the perfect balance in its implementation. Annas attributes Aristotle's inclusion of such problematic virtues to his desire to respect what were conventionally regarded as virtues⁴⁵. Annas regards as more appealing the view held by the Stoics and others on which it is an unacceptable position to say that whether or not I am a moral person depends on elements outside my control. This approach is more consistent with the overall project on which one is looking for virtues that can ensure success in reaching one's *telos*.

Sticking with this Stoic approach to virtue, Annas thinks that virtues are not directly helpful in giving an account of knowledge. Because virtue does not require success in reaching one's *skopos*, it cannot be used to explain the requirement of success in knowledge. Annas writes that,

...knowledge is a success term, and so a theory which defines knowledge in terms that feature virtue must take virtue to be a success term.⁴⁶

⁴⁴ Annas (2003) p 26.

⁴⁵ Annas (2003) p 27.

⁴⁶ Annas (2003) p 28.

As a result she thinks that the account of the virtues used in ancient virtue ethics cannot be used in modern virtue epistemology, and something more like Zagzebski's model of the virtues should be used instead. However this move, I think, is too quick. Even if knowledge is a success term, and is defined in terms of epistemic virtues, the success component need not be built into the virtues themselves. Rather the success component can be something extra that is added to the concept of justified or virtuous belief to make knowledge. Since the ancient conception of virtue does not contain the externalist reliable success component, it seems a reasonable approach to start with this general conception of virtue, and see if when applied to epistemic virtues, it yields a useful concept. The right way to approach this problem is to begin with an account of epistemic virtue and only after this project is completed to develop an account of knowledge. In developing an account of knowledge we will require success at some stage—after all knowledge does require truth. But the mere existence of a truth requirement in an account of knowledge does not mean that the epistemic virtue or justification that plays a role in the account of knowledge also require a special connection to truth like that of reliable success.

The ancient approach to virtue is a good fit for an internalist version of virtue epistemology. Such an approach will be more concerned with the concept of justified or virtuous belief than with the concept of knowledge. In developing an internalist virtue epistemology we will of course need to move beyond talk of epistemic virtues to make contact with traditional epistemological worries. Justification of a particular belief then will depend on the motivation that moves the agent in the particular event in question. If the agent is acting out of a virtue that she possesses, then her particular belief is justified, even if the belief is false.

This approach respects the ancient intuition that an agent's overall (epistemic) well being depends only on the agent herself. Each individual instance that makes up that epistemic life has the same structure—its being justified or not depends only on characteristics of the agent. Thus our *teloi* as epistemic agents, being justified, is something we can succeed at regardless of moral luck.

The Argument from the Sources of Value

Why does Zagzebski move away from the ancient conception of virtue and include a success component in her account of the virtues? One reason might have to do with her own two-stage account of epistemic motivation. According to Zagzebski, all the epistemic virtues can be identified by the unified primary motivation that brings them about. This primary motivation is to gain truth or “cognitive contact with the world”⁴⁷. This primary motivation then gives rise to more specific motivations that are characteristic of the virtues. Zagzebski does not elaborate on the way that these secondary motivations are developed, but clearly they could be the result of an agent believing that these characteristic ways of acting are good ways of achieving the goal of the primary motivation, and so adopting them as secondary motivations. The more specific secondary motivations are each characteristic of a virtue. For example, the motivation to stand by one's ideas and to give them their proper due is characteristic of the virtue of intellectual courage. Thus we develop specific patterns of behavior, and the motivations to keep us in these patterns, as a way to reach our primary goal of obtaining truth. We develop intellectual virtues as a response to our motivation for

⁴⁷ Zagzebski (1996) p.206.

truth, and so the virtues must actually succeed in reaching the truth. If they don't we should adopt other virtues that are really successful at reaching the truth.

Given this model on which there are two different stages of motivation playing a role in the generation of our virtues, we can reasonably ask from which stage the value of a virtue derives. A tempting answer, and one I think Zagzebski suggests is that the value of each of the intellectual virtues derives from the same shared source. Since all of the intellectual virtues are identified with secondary motivations, which are themselves based on the same primary motivation, the primary motivation for truth could be the source of the value of the virtues. We might say that a motivation to obtain truth is a necessary component of the good life, or that such a motivation has its own basic value. The value of the secondary, more specific, motivations might then be derived from the primary motivation for truth, insofar as the primary motivation is valuable, and the secondary motivation is a means to satisfying the primary motivation.

This approach to explaining the value of the virtues has the admirable characteristic of theoretical simplicity, since there needs to be only a single source of basic value. However this approach puts the value of the virtues on shaky ground, because it makes the value of the virtues depends on the contingent connection between the primary motivation and the secondary motivation. There is always a risk that the secondary motivation we think of as a virtue is not an efficient means to satisfy the primary motivation. If this is the case, then we are left in a position of not being able to explain the value of the virtue. So long as the virtue itself does not have a basic value, but only has instrumental value based on its effectiveness

bringing about something else valuable, there is a risk that the connection will be severed and our virtues will be left without value.

We might further question whether the motivation for truth ought to be the source of value for the virtues when we consider the ways that the motivation for truth itself might come into conflict with the epistemic virtues. If we think that we have the truth the motivation for the truth may lead us to become dogmatic and to ignore the objections of others. If we think we have encountered an oracle the motivation for the truth may lead us to place too much faith in the word of another, thus losing our intellectual courage. In both these cases, a desire for truth can lead us into an intellectual vice rather than a virtue. Thus we might well wonder why we should respect any secondary motivations that are adopted out of our motivation for the truth. If this motivation can lead to vice in some cases, it cannot be relied upon to be the source of value in others.

In addition to these worries about virtues having derivative value, I think there are more straightforward reasons to want virtues to have non-derivative value. If we are attempting to present a robustly virtue-based epistemology (or for that matter a virtue ethics) we want take seriously the idea that the value of specific acts derives from the value of the virtues that produce them. In order to honor this structure of analysis, we should not then give an account on which the value of the virtues depends on the effectiveness of the particular acts it produces in obtaining a set goal, such as truth. To analyze the value of a virtue in this way is to trivialize the role of virtue. If we think that a virtue is only valuable because it produces good acts, then we are really proposing an act-based, rather than virtue based theory.

Furthermore, within a virtue theory, we will want to use our conception of the virtues to explain why individual acts are virtuous or good. But if we base the value of the virtue upon the value of the acts that the virtue produces, we cannot then base the value of the acts that the virtue produces on the virtues themselves without descending into a vicious circle.

For these reasons, I think that it is better to treat the value of the virtues as a basic value. Such an account could simply treat virtues, or the motivations they are identified with, as having a basic value; holding that the virtue or motivation is good in itself. Or it could treat virtues as having value because they are constitutive of the good life. To treat the value of the virtues in this way is not to treat it as derivative—the virtues are not good because they are instrumental to bringing about the good life. Rather, by being constitutive of the good life, they have a basic value, which is part of the basic value of the good life. While this approach lacks some simplicity, by admitting many basic sources of value, some of the desired simplicity can be regained by recognizing that the good life (and its parts) is the only thing with basic value. The disparity of the basic constitutive parts of the good life will add some complexity, but this complexity will have limits. Since the good life is a unified whole, the elements of the good life need to fit together in a way to make up such a unified whole. No mere collection of desirable traits will make up a good life, since the desirable traits brought together run a great risk of not being jointly satisfiable.

The Proper Structure of the Epistemic Virtues

Putting together the above arguments, we come to the conclusion that the structure of the virtues used for epistemology should be internalist, that is, it should not contain a success component. Our concept of the intellectual virtues can retain something like Zagzebski's motivational component, while rejecting the success component. This amendment will ensure that the virtues described are something that individuals are capable of achieving, even if they are not easy to achieve. Indeed, if virtues are, as traditionally conceived slowly developed character traits, then obtaining them will be quite difficult, and require much practice and habituation. Thus virtues need not, and ought not, be easy to obtain, but they should be possible to attain, even for those in unfortunate epistemic circumstances.

One might worry though that an account of the virtues that depends on motivation alone is too weak, and that it places too few constraints on who gets to count as virtuous. For example we might consider the case of Sam, who has the characteristic motivations associated with a number of different virtues. These motivations are a result of his deep psychological make-up and have developed over time. However, though regularly having the motivations characteristic of the virtues, Sam never acts on them. While he has generous motives to help those in need, they never get him up from the couch. While he has courageous motives to protect those in danger, they never get him away from the TV. Despite Sam's admirable motives, it seems implausible to call Sam the couch potato a virtuous agent. While Sam has the correct motivations, it seems that he does not have them in the right way. Thus we might worry that having the correct motivation is not sufficient for having a virtue.

The correct response to this worry is to point out that we can retain some elements of the idea of success without holding that a virtue requires reliable success at obtaining external goods. While we have seen arguments above against requiring external success in a virtue, there might still be some type of success that we require—that of internal, or motivational success. In order to be virtuous, the agent needs not only to have the right kinds of motivations, those motivations also need to actually motivate the agent. This does not mean that the agent must act on them in all circumstances—in some cases more pressing concerns will cause an agent to ignore some of her motivations. Nor does it mean that the agent must act on her motivations in a way that is actually effective or successful—this would be to require external success. Rather the agent must have the characteristic motivations of a virtue and those motivations must lead the agent to act in ways that she sees fit, whenever this motivation is not interrupted by other concerns. This internal element of success is not troubling in the ways that an external success requirement would be. To be virtuous, an agent must be moved to some action. However if she is deceived, her actions on the basis of that motivation might not be the best possible ones. She might act poorly because of her unfortunately deceptive circumstances—this does not exclude her from having a virtue. But to properly respond to the motivation making up a virtue she needs to act, rather than remain passive.

With this clarification on the limited role of success, we are ready to give a definition of virtue that satisfies the demands of the arguments I have given. Virtue should not require actual success in reaching one's ends even if it does require limited success in that the agent

must be moved by his motivation into action. Thus we can give the following definition of virtue in general:

A **virtue** is an acquired character trait that is distinguished by its distinctive goals, and motivation to attain those goals, which is acted upon to the best of the agent's ability in the circumstances; having each virtue is also a constitutive part of having an overall good life.

This definition does not have any reliable success condition, but it requires success in a limited way by requiring that the agent not only be motivated, but also that the motivation leads the agent to act insofar as he or she is capable.

It might be objected that this account makes intellectual virtue too undemanding. For the person who is grossly negligent in his attempts to bring about the ends of virtue, failing in every instance, would count as virtuous. For example, Gene might have all the necessary motivations to have the virtue of generosity, however he is quite ill-adept at implementing his generosity. He gives bikes to the crippled, and televisions to the blind. While his motivations are perfect, the error of his implementation might lead us to think that he does not really possess the virtue of generosity. How can the proposed definition of a virtue accommodate this reaction?

There are two different responses that might be given in response to this case. The first is to question how Gene could have acquired this character trait. Virtues are developed by

first emulating the behavior of others, and then figuring out for oneself how to both act and be motivated as other virtuous agents are. Given this model of development, how could Gene develop his putative virtue of generosity? He would have to have found another agent to follow with as bad a track record as his own. Or he would have to have misunderstood the actions of the truly virtuous agent. In either case it seems likely that his actions would have been corrected during the time that he was developing his virtue of generosity. Perhaps someone might have pointed out to him that his actions were not very helpful. Or he might have noticed on his own that his gifts were rarely used. In any case it seems that Gene could not have developed his idiosyncratic approach to generosity without willfully ignoring evidence that his own efforts were not achieving the desired end of the characteristic motivation of generosity. Thus it seems that developing a putative virtue with such bad implementation is highly unlikely.

The second response to the case of Gene would be to admit that such an aberrant virtue could develop, but to look a bit closer at how it would develop. If Gene ignored the advice of others, urging him to change his ways of expressing generosity, then he can not have been very open-minded. If he ignored the evidence of his own experience, not considering the fact that his gifts were not used, then he could not have been very epistemically careful. Looking at this path of development, we notice that it is virtues other than generosity, such as open mindedness, or epistemic carefulness, that are lacking from Gene's development. If Gene were to possess these other epistemic virtues they would act as a corrective on his development and implementation of generosity. Since his actions fit his motivations so poorly, we can see that Gene lacks these other virtues. He is not an overall virtuous person.

But, while recognizing this, we can admit that Gene does have the virtue of generosity, even though he lacks the other virtues that would make his generosity more effective. In this way we can explain the feeling that Gene is lacking in something, while continuing to hold the other intuition that Gene is virtuous because he is correctly motivated.

By keeping the distinction between possessing an individual virtue and possessing them all clear, we can retain the definition of virtue that depends only on the agent's possessing the right motivation and acting on that motivation, while admitting that an agent who has this much might still be defective. The agent is defective because he does not possess all of the virtues that an agent might, even if he does possess the virtue in question.

The case of Gene the insensitive giver shows us that it is possible to have an agent with a moral virtue who lacks certain epistemic virtues, and hence is both not overall virtuous, and not efficient in achieving the desired ends in the individual cases to which he applies his virtue of generosity. We might also have an agent who is the opposite of Gene, who has epistemic virtues but lacks moral ones. An example of this might Sid, an amoral scientist working on developing weapons. This scientist is intellectually careful, and also intellectually courageous. He stands up for his theories about how to make more efficient weapons, and carefully looks for existing research and performs his own experiments to test those theories. Yet he is entirely unconcerned with the havoc his weapons might cause. He is not at all motivated by the virtue of beneficence—the negative outcomes for others simply don't move him. Although Sid has a great many moral failings, and we want to criticize him for these failings, I think that we would hesitate to say that Sid does not know that his well-

tested theories about how to make more efficient weapons are true. This example shows that that an agent might have epistemic virtues while lacking many moral ones. It also points out a need to make a distinction between the epistemic virtues and the moral virtues if we are going to use our account of the virtues to give an account of what it is to know. Otherwise we might end up counting an immoral or amoral agent as lacking knowledge simply because he lacks moral virtues.

An **epistemic** virtue is a virtue whose characteristic motivation concerns states of belief. That is, the particular acts that one is motivated to perform by an epistemic virtue are acts of belief or non-belief (or degrees of belief intermediate between belief and non-belief). In recognizing this, we can recover a limited but important role for truth to play in the epistemic virtues. While we have seen that requiring epistemic virtues to succeed in reaching the truth, even just reliability, leads to problems, we can still allow that the goal of all our individual beliefs is to be true. That is, our goal in each individual belief is to believe truths and avoid falsehoods⁴⁸. Since these are the goals of our individual beliefs, they would count as *skopoi* according to the distinction made earlier, not *teloi*. Thus truth plays a special role in our believing, and yet we can still believe well even if we end up believing false things. This is analogous to the courageous man who races into a burning house to save someone. His goal in so doing is to save the person inside, but even if he fails at this goal he is still acting courageously, and is still a courageous person. Similarly for the epistemic virtues. The *telos* of our epistemic lives is to be epistemically rational agents. This is the epistemic good for us

⁴⁸ See Lehrer (2000) for a similar formulation of the goal of acceptance. It is important to notice that it is not enough to only try to avoid errors, for this would be trivial task if we were not also concerned to believe truths.

as human beings. This *telos* is constituted by the particular epistemic virtues, such as intellectual courage, or open-mindedness. In particular instances our virtues lead us to particular epistemic actions. Our virtue of intellectual courage may lead us to believe our own theories even though others disagree with us. Our proximate goal or *skopos* in doing this is not just to believe our own theory, but to believe it correctly. Sometimes what we believe, even as the result of a manifestation of our epistemic virtues turns out to be false. In these instances we fail to achieve our *skopoi*, but we still have epistemic virtues, and are even acting in accordance with them⁴⁹.

However, given that our goal in each individual belief is reaching truth it may be cold comfort for the agent who learns that, although she believes virtuously, she also believes falsely⁵⁰. This is parallel to a case in which one discovers that one's acts have not brought about their intended consequences, say if Grace were to learn that the charity she has been giving to is a fraud. Upon discovering that we have been deceived, we as virtuous epistemic agents should work to uncover that deception and get to the truth of the matter. Thus it is not satisfactory for a virtuous agent to rest on her virtuous laurels. If an agent is truly epistemically virtuous and discovers that she has been deceived, her virtue will motivate her to find the truth. It is only from a third-person perspective that we can evaluate someone as both epistemically virtuous and deceived, and rest content that the agent is still virtuous.

⁴⁹ This role for the motivation for truth is similar to that offered by Foley in (1993). He holds that aiming at truth will lead us to certain epistemic standards, and following those epistemic standards is the best we can do, even if they do not give us the truth.

⁵⁰ Lehrer points out the potential conflict between the goals of achieving virtues and reaching the truth in his (2001) p 203.

From the first-person standpoint, the discovery that one has been deceived, or even simply made a mistake, should motivate one to seek the truth more assiduously⁵¹.

Thus we can see that that connection between epistemic virtues and the truth is a subtle one that is far less straightforward than simple reliability at getting to the truth. We as epistemic agents want to reliably get at the truth, even though we may be epistemically virtuous without a reliable truth connection⁵². So the desire for truth plays a special role in epistemic virtues, and can serve, along with the criterion of beliefs being the acts, to distinguish the special class of epistemic virtues.

Knowledge and Gettier: Undermining a Motivation for Externalism

Next I will turn to the tradition task of giving a definition of knowledge from within my version of virtue epistemology. Before I turn to this task however I would like to look at a general argument given by Zagzebski about the structure that any account of knowledge needs to have in order to avoid Gettier problems⁵³. I would like to examine this argument because I think that while it is correct, Zagzebski has not seemed to notice that it undercuts a possible motivation for externalism either in our account of justification or in our account of the epistemic virtues. I will show how this argument undermines a possible reason to give an externalism account of the virtues, and then present my own account of knowledge

⁵¹ This might be compared to the claim in virtue ethics that we do not aim directly and being virtuous, but rather aim at accomplishing the goals that our virtues motivate us towards (our *skopoi*). Indeed aiming only at being virtuous may be a self-undermining goal, like the goal of being happy. We do not want virtuous agents to be content in their virtue if they know they are not reaching the goals towards which the virtues motivate them. (Thanks to Dave Schmitz for pointing this out to me.)

⁵² Lehrer puts our goal in being epistemic agents in a very similar way in (2001) p 205.

⁵³ Originally introduced by Gettier in his (1963).

that conforms to the requirements on an account of knowledge given by Zagzebski's argument.

Zagzebski is looking for the features that generate the original Gettier problem, in order to give us a general way that we might avoid Gettier problems⁵⁴. She begins by looking at the traditional justified true belief accounts of knowledge as well as the modified accounts of knowledge that have sprung up as an attempt to avoid Gettier problems. Whether these accounts try to modify the account of justification to solve the problem, or add some other condition outside of justification to do the same job is immaterial. To get a general structure that covers both these kinds of cases, Zagzebski gives us a general formulation of all accounts of knowledge by lumping together whatever an account requires for knowledge over and above true belief under the title "the goodmaking quality Q"⁵⁵. Zagzebski then claims that a Gettier case can be developed for any theory of knowledge in which the goodmaking quality Q does not entail truth. To see that a Gettier example could always be constructed against any such theory, Zagzebski analyzes the general structure of Gettier problems and gives us a recipe for generating a Gettier case to fit other theories. She holds that Gettier cases all include an instance of double luck. First we think of a situation in which the agent is unlucky in holding a belief that, although it has goodmaking quality Q, is not true. We then change the scenario slightly to include a second bit of luck – in this case good luck.. Whatever this good bit of luck is, it must make the original belief, fortunately, true. Thus by double luck, the belief meets all criteria for knowledge. The belief is, through

⁵⁴ Zagzebski (1999) p 99-104.

⁵⁵ Zagzebski (1999) p 101 ff.

the bit of second luck, true. And the belief has goodmaking quality Q, since it was chosen that way. Thus by this belief constitutes a Gettier case for whatever theory of knowledge we were considering.

This argument given by Zagzebski is a general one and shows that any account of knowledge in which the goodmaking quality Q does not entail the truth of the belief. This feature of a theory allows us to start on her recipe by finding a belief that both has property Q and is not true. Without this feature, we cannot start generating a Gettier case for the theory. So only theories in which truth is entailed by quality Q can avoid being undermined by Gettier cases.

Given this argument by Zagzebski, we can now ask what we want goodmaking quality Q to be. In the traditional justified true belief (JTB) account of knowledge, quality Q would be justification. Given Zagzebski's argument, we can now see that if we have a JTB account of knowledge, it must give us an account of justification such that justification entails truth. Any such account of justification would be a very strong infallibilist one. On the other hand, we might not want justification on its own to be the quality Q, but rather justification taken in conjunction with some other qualities of the belief (the sometimes called the "fourth condition" to be added to a JTB account). In this case justification need not entail truth, this job can be handled by the other part of the quality Q. Indeed, if we take the quality Q to be more than justification we do not need an account of justification that is truth connected at all. This makes room for an internalist account of justification, or an internalist account of epistemic virtues which will play a role in the account of knowledge similar to that played by justification.

Let's examine the appeal of an externalist account of justification on each of these two options. The first is that where the quality Q just is justification. One of the reasons that we might have found externalism either in an account of justification or in an account of epistemic virtues attractive in the past is the worry that if these accounts did not provide a connection to truth by requiring something like reliable success, then our account of knowledge might be susceptible to Gettier cases. However Zagzebski's results show us that the only form of justification that is strong enough to do this is an account of justification on which the justification entails truth. Adding a reliable success condition will not be sufficient, because even if a method is reliable, it may not succeed in reaching the truth all the time. This gap between reliability and truth gives us a place to start generating Gettier cases using Zagzebski's recipe given above. Thus the only type of externalism that would prevent Gettier cases would be one that guaranteed success, not just reliable success. But this is contrary to most externalist accounts, and it is certainly much stranger than the externalist views we have seen so far. Furthermore, on an externalist account where justification guaranteed truth, the concept of justification would be redundant. Any justified belief would be true, and so any justified belief would be knowledge. But part of the reason for having a concept of justification is so that we can talk about cases in which people have justification without knowledge. If there is no difference between justified belief and knowledge then the concept of justification becomes redundant, and this is an undesirable consequence.

What about the appeal of externalism if we take the second option and hold that quality Q is made up of more than just justification, but rather justification plus some fourth condition?

In this case justification can be fallible, because the fourth condition will ensure that the entire quality Q entails truth. Thus truth and the connection of the belief to truth will be built into the fourth condition. If this is so then there is little need to build a truth condition, even a reliable truth condition into the concept of justification or epistemic virtue. Indeed to do so would be redundant, since the truth connection will already be made by the fourth condition. This undercuts the motivation to have an externalist element in the account of justification, because it is not needed in order to avoid Gettier problems.

Thus we see that Zagzebski's argument about the general structure of accounts of knowledge can be used to undermine the motivation for including a success condition in an account of justification or of epistemic virtues. This is a bit surprising since Zagzebski is committed to having a success condition in her account of the virtues. But taking this argument which undercuts the motivation for externalism, together with the three arguments against externalism presented earlier in this chapter we have good reason not to include an externalist success condition in our account of the intellectual virtues. When developing my account of knowledge then I will use the internalist account of the epistemic virtues given above, and make the truth connection needed to avoid Gettier cases by building it into goodmaking quality Q.

Now let us first begin with an account of justification that is based on an internalist conception of the virtues.

A belief is **justified** for S iff it arises from one of S's epistemic virtues by way of that virtue's characteristic motivations.

Since our account of epistemic virtues contains a version of what Zagzebski calls the motivation condition, but not what she calls the success condition, this account of justification does not ensure that the agent's belief is true, or even reliably true. But this I think is a desirable result for an account of justification. Just as it seems possible for an agent to be epistemically virtuous even in a bad circumstance where her beliefs are not even reliably true, it also seems possible that an agent can have justified beliefs in a bad circumstance where her beliefs are not even reliable true. By not including a success condition we leave this possibility open for epistemic virtue, and by building our concept of justification out of this version of epistemic virtues, we leave open the possibility for justification too.

Despite the plausibility of this account of justification, it is not a good candidate to be simply inserted into a JTB account of knowledge. Since justification does not require a belief to be true, simply adding internalist justification to true belief would not give a satisfactory account of knowledge. For the truth of the belief may be entirely accidental, and unrelated to its being justified. So something more than true belief plus justification is required for adequate account of knowledge. As we have considered above, this needed connection to truth might be built into the goodmaking quality Q. Thus my candidate for the goodmaking quality is the following:

A belief is a **successful application of an epistemic virtue** iff it arises from an epistemic virtue by way of that virtue's characteristic motivations, it brings about (or

sustains) a true belief, and is successful in reaching the truth because the belief is generated by that virtue⁵⁶.

So a successful application of an epistemic virtue contains both a requirement of justification and a success condition. Thus we can end up requiring success for knowledge, which is after all a success term *par excellence*, without requiring success for justification. Not only does successful application of an epistemic virtue require actual success, it also requires non-accidental success. This requirement meets criterion set up by Zagzebski for avoiding Gettier cases. This final requirement prevents there from being a belief that has quality Q and is accidentally true or true through an instance of good luck only. Rather we require that the truth of the belief not be the product of good luck but rather the product of the application of the epistemic virtue.

Finally with the above definition of the goodmaking quality Q that is to be added to true belief to make knowledge, we are now ready to give a definition of knowledge.

Knowledge is true belief arising out of successful application of an epistemic virtue⁵⁷.

⁵⁶ This definition follows Zagzebski in her general approach to avoiding Gettier problems. See Zagzebski's definition of an act of intellectual virtue (1996) p 270 ff.

⁵⁷ A possible counterexample to this definition of knowledge has been suggested to me by Keith Lehrer which is variation on his Nogot and Havit version of the Gettier case presented in his (2000) p 19. Imagine that there is another student in your class Devious Dora, who decides to convince you of the truth that there is a student in your class with a Ferrari, by convincing you of the falsehood that Nogot has a Ferrari; However she will only do this if you go about trying to discover the vehicular status of your students in a virtuous way. We might worry then that your success in reaching the truth meets the requirement of arising out of a successful application of an epistemic virtue – the success is just brought about in an odd way – but is not a case of knowledge. This counterexample however depends on the claim that the truth of your belief is dependant on its having been generated by an epistemic virtue, but the fact that you are using an epistemic virtue does not play an essential role in the explanation of why you arrived at the truth. Although Dora chose to reward your epistemic virtue, she might just as well have decided to reward your epistemic vice. So if we are careful about

Although I have included a requirement of truth in this definition, we should notice that it is strictly speaking redundant. Any belief brought about by a successful application of epistemic virtue must be true.

Thus, by correctly interpreting Zagzebski's argument about the structure of accounts of knowledge we can achieve accounts of epistemic virtue, justification, and knowledge that meet our desiderata. We have seen three arguments for having an internalist account of the epistemic virtues. These were the new evil demon argument, the historical argument and the source of value argument. Parts of these arguments can also give us reason to prefer an internalism account of justification. The accounts I have given of virtues in general, epistemic virtues, and justification are all internalist. None require external reliable success. Thus we can preserve the possibility of having a virtuous agent in an uncooperative world, or a justified belief in bad circumstances. However we also want an account of knowledge that is externalist in requiring truth and that makes the correct link between the epistemic virtues and reaching the truth. We do this by defining the notion of a successful application of an epistemic virtue. This notion can combine the internalist elements of epistemic virtues with the strong externalism element of requiring success. As a result this notion, of the successful application of an epistemic virtue, while based on the epistemic virtues is strong enough to give us an account of knowledge that avoids possible Gettier problems.

what we count as success that happens because it is generated by an epistemic virtue we can avoid this counterexample.

CHAPTER 4: CONTEXT SENSITIVITY

Worries about Skepticism

Given the accounts of epistemic virtues, justification, and knowledge presented in the last chapter, we might move on to a worry about how often these characteristics are found in the world. Any account of knowledge or justification that leads to the conclusion that there is little justification or knowledge in the world may rightly be challenged as being too restrictive, and hence of inviting the specter of skepticism.

Within the realm of ethics this charge has been leveled at virtue theories. Concerns are raised that standards set by virtue theories are simply too high. Rather than setting a standard for action and asking agents to live up to that standard, virtue theory requires not only that actions are correct but also that the motivations behind them are correct as well. Furthermore there is a worry that living a virtuous life is a standard that few can achieve because most of us do not have the time or freedom to engage in solitary contemplation. Rather, most of us find ourselves embedded in lives with already existing obligations, and as a result we might worry that living a life of virtue is not an option for these people.

This is a serious worry addressed by the ancient Stoics. In reasoning about how to live a life of virtue the Stoics were similarly concerned that virtue not be a standard that is out of reach of most people. And they realized that most people, particularly by the time they were considering how to live a life of virtue, are already embedded in a life with its own demands

and obligations⁵⁸. One starts out to live a virtuous life as the daughter of particular parents, and perhaps mother to particular children. One is also a member of one's country, state and community. And one is member in particular ways. In choosing a profession one has taken on a particular way to contribute to the community at large and to your own professional community. We might summarize these commitments and obligations as an agent's social role. This social role then interacts with what a virtuous life will look like for a particular agent. If one has living parents one needs to live as a virtuous daughter. If one has children then the virtuous life for that person will include being a virtuous mother. Thus there are many different types of virtuous lives that an agent might live, and these different virtuous lives might look very different from each other.

This variation will be particularly obvious when we look at the individual actions of the virtuous agents with different social roles. Faced with the same circumstance it might be consistent with the virtue of different agents to do very different things. For example if one witnesses a burglary and one is a police officer, then the courageous thing to do may be to pursue the criminal. However if one is not a police officer, then the courageous action may well be to call the police. Calling the police is not a cowardly action, rather for the ordinary citizen it is the exactly right action. Thus in this example we can see that the right or virtuous action to take in a particular circumstance depends on the social role of the agent in question.

⁵⁸ See Annas (2002).

One reason that virtue can be sensitive to social roles in this way is that the virtuous action is the mean between two extremes. This is to be contrasted with an alternate model of the virtues on which the virtues aim at producing some good, and more of that good is ways better. This model of virtues as an extreme is one on which they are the pinnacle of a particular characteristic; on this model it is hard to see how two agents who do different things can both be acting out of virtue. Even if there are details of the situation that mean that the agent in question can't really be expected to perform perfectly, this sounds like an instance in which the agent excusably fails to achieve virtue. So on a model in which virtue is an extreme, there cannot be two agents who are in the same circumstance and yet both act differently while still both being virtuous. This model of the virtues as extremes would result in a picture of the virtuous life that few if any could achieve. However a model of the virtues on which they are a mean between two extremes results in a much less demanding and more realistic model of the virtuous life. On such a model the mean between two extremes can vary both in response to circumstances and in response to social roles. The virtuous action first depends on the circumstances. The mean between the extremes varies depending on details of the circumstances. Depending on the threat a courageous action may be defensive or offensive, or may simply be to wait out the assault. In addition to variation in response to circumstances, the mean may also vary in response to the social roles of the individuals involved. As the example above, the circumstances that two people find themselves in can be the same, and yet the courageous action varies because of the social roles of the agent—that of a police officer or that of a bystander. I will call both of these types of sensitivity—both to circumstances and to social roles—context sensitivity.

Context sensitivity can be extended from the moral virtues to the epistemic virtues as well. Just as we take on specific moral obligations we also take on epistemic obligations. The most common example of this is in our choice of a profession. It is a desirable result of virtue epistemology that agents are left to choose their own epistemic areas of interest. It would be both counterintuitive and pragmatically unstable to insist that all agents ought to have the same epistemic interests. Rather virtue theory allows each agent to choose the epistemic responsibilities she takes on. On deciding to become a doctor, an agent takes on responsibilities to be particularly epistemically careful in evaluating claims about drug effectiveness. An agent who does not take on the social role of a doctor still should be careful about evaluation of such claims, but need not be as careful as one who has taken on a doctor's obligations. This is another example of there being two means between extremes for agents with different social roles. Between the extremes of being epistemically slipshod, and being epistemically too cautious, we find the mean of epistemic carefulness. But epistemic carefulness is different for a doctor and a layperson, even if they find themselves in the same circumstance. Given the same evidence the layperson may act virtuously by believing a new drug stops heartburn, while the doctor acts epistemically virtuously by withholding judgment and looking for more information. Thus virtue epistemology can offer virtues that are not too demanding on the agents by being sensitive to their own interests and to the epistemic social roles they choose to take on the basis of those interests.

Because many versions of context sensitivity or contextualism are externalist in nature, I want to stress here that the context sensitivity I am suggesting as part of virtue epistemology is internalist in nature. That is to say the social role that affects the virtuousness of particular

lines of action or thought is the perceived social role of the agent. This means that in a situation involving large-scale deception the actual role and the perceived social role may vary. For example think of the police officer in the example above whose virtuous action is to pursue the burglar. Now imagine that rather than being a bona fide police officer she is simply an ordinary woman who has been elaborately duped in to thinking she has passed through the police academy and is now a police officer. So long as the deception is careful and extensive enough we will not say that she lacks epistemic virtue by being deceived, rather that she is being epistemically virtuous in a bad circumstance. In this case her perceived social role is well epistemically grounded, even if not correct, and thus can serve as the basis for her chasing the burglar being virtuous. Indeed it would be an inappropriate response on her part to stop a pursuit and behave like an ordinary citizen, given that she believes that she is a police officer with special duties to uphold. Thus the social roles that I am talking about will be the agent's perceived social roles, which may diverge from the agent's actual social role in situations involving deception.

Other Versions of Context-Sensitivity

Including some element of context-sensitivity is not a new tactic to avoid skepticism. Indeed the tactic has been employed in several different ways. In this section I would like to isolate two main approaches to the application of contextualism to the problem of skepticism, attributor contextualism and inferential contextualism, and then compare them to the context sensitivity in my version of virtue epistemology. Both versions are introduced to solve the skeptical problems presented by arguments of the form:

I do not know that I am not a brain in a vat

If I do not know that I am not a brain in a vat then I don't know that I have hands

Therefore I don't know that I have hands

This type of argument can be used to attack any piece of our everyday knowledge. However this is a counterintuitive and massively skeptical conclusion which we would like to avoid.

Contextualism is introduced in two ways to avoid this type of argument. Both types of contextualism respond that in the context of a skeptical argument it is appropriate to say that I don't know I have hands. However this does not mean that in normal contexts it is OK to say that I don't know I have hands. In those contexts I really do know that I have hands. It is only when we engage in philosophical discourse, and particularly philosophical discourse of the skeptical sort that we lose our knowledge. As David Lewis pithily puts it, "epistemology destroys knowledge"⁵⁹. But not for good. When we move out of skeptical contexts, our knowledge returns⁶⁰. The two major versions of contextualism agree about this much. What they disagree about is what determines the context in question and how that context can be changed. These differences will give different results in particular cases, and can be so evaluated. They can also be evaluated on the basis of their theoretical simplicity and overall persuasiveness.

The first type of contextualism I will call attributor contextualism⁶¹. This type of contextualism focuses on instances of knowledge attribution and the truth conditions for those attributions. Although the context of the subject of the attribution is taken into

⁵⁹ Lewis (1996).

⁶⁰ The mechanism by which this occurs is different for different authors. See Lewis (1996), DeRose (1995), Cohen (1988), and Williams (1992).

⁶¹ Following Cohen (1998). This version of contextualism is also closely related to the version that Pritchard (2002) calls semantic contextualism.

account, what is distinctive of this type of contextualism is that the context of the attributor plays an important role in determining the meaning of the contextually sensitive knowledge attribution. This means that according to attributor contextualism, it is possible for two attributors, A and B, to both attribute knowledge to the same third person, say John, and for the knowledge attribution by A to be true while the knowledge attribution by B is false⁶². One way this can be true is if the attribution in question is “John know that he has hands” and while attributor A is in a normal non-skeptical context, attributor B is in a skeptical context, say discussing skepticism in an epistemology seminar. According to attributor contextualism, the context of the attributor can change the meaning of the expression “know”. In a normal context, when the attributor A uses the term “know” it has associated with only our everyday standards for knowledge. However when the attributor B used the term “know” in a skeptical context it has higher standards, perhaps impossible standards, associated with it⁶³. If John meets the low standards but not the high standards, then “John knows he has hands” can be true when uttered by A and false when uttered by B. This is the distinctive feature of semantic contextualism, and the result that shows just how radical this version of contextualism is.

Having determined that the context of the attributor determines the meaning of the term “know”, we still need to understand how features of the attributor’s context can affect this meaning. The structure for this is introduced by David Lewis (1996). This is the rubric of properly ignored alternatives. In evaluating a knowledge claim in a given context there are

⁶² Cohen points out this feature of attributor contextualism in his (1998).

⁶³ DeRose explains the mechanism by which standards can be raised and lowered in his (1996).

particular alternatives that we ignore. These alternatives are properly ignored because of certain background assumptions that are taking place in the conversation. Within a conversation the context can be changed by changing the focus of attention. The simplest way to do this is to simply introduce an alternative that has been properly excluded up to this point. Once an alternative is introduced it cannot be properly excluded. This is captured by Lewis as the Rule of Attention⁶⁴. However there are less direct ways to introduce an excluded alternative. For example if an alternative is quite similar to an alternative being considered, then the similar alternative cannot be excluded either. This is Lewis' Rule of Resemblance⁶⁵. What is distinctive about these rules for setting context is that they are easy to change—all I have to do to introduce an alternative that has been properly excluded up to this point is to mention it, or even to mention an alternative similar to it. Thus on the attributor contextualist view contexts are very flexible, and can be changed quite easily.

This feature of semantic contextualism is in contrast with the type of contextualism held by Michael Williams, which is called by Pritchard 'Inferential Contextualism'⁶⁶. This version of contextualism does not hold that the context is determined by the circumstances of the attributor, but rather by the circumstances of the subject—the person who's putative knowledge is under analysis. This means that it will not be possible for one attributor to truly attribute knowledge to a subject while another attributor truly attributes lack of knowledge to the same subject, as is possible in attributor contextualism. Furthermore conversational features play a far less important role for Williams. According to him a mere

⁶⁴ Lewis (1996) p, 434.

⁶⁵ Lewis (1996) p, 429.

change in conversation cannot change the context⁶⁷. Each of these inferential contexts carries with it certain parameters determining which things can be questions and which can be taken as basic within that context. For example we might think of the inferential context for doing history⁶⁸. In such a context we might wonder if a particular document is authentic or not, or whether the date we have attributed to a found object is correct. But within a historical context it is not appropriate to consider the “Russelian Hypothesis” that the world was created five minutes ago with all of its apparent historical traces. Although it might be reasonable to consider the Russellian hypothesis in other contexts, within the context of doing history even considering it is unreasonable. Since inferential contexts are connected to particular purposes, we can only change those contexts by changing our purposes and getting others to change theirs too. Thus for Williams it is possible to change the context from that of doing history to that of skeptical worrying, but it takes much more than simply mentioning a skeptical hypothesis to make this change.

Michael Williams’s contextualism also differs from attributor contextualism in that there is no hierarchy of contexts⁶⁹. In attributor contextualism conversational changes can raise or lower the epistemic standards by shifting the interlocutors to a different context. Speaking of ‘higher’ or ‘lower’ standards implies that there is some ranking of the contexts. Some are more rigorous than others and some more relaxed. Although there are different ways in which the epistemic standards can be raised—by introducing different alternatives—there is

⁶⁶ Pritchard (2002) p35 ff.

⁶⁷ Williams (1992).

⁶⁸ Pritchard (2002) p 40.

⁶⁹ Williams (1992).

a clear sense in which the different standards in these contexts can have at least a partial ordering. For any context A in which the alternatives to be properly considered are a subset of the alternative to be considered in another context B, we can say that the epistemic standards in the context B are stronger than the epistemic standards in context A. In order to meet the standards for knowledge in context B we have to rule out more alternatives than we have to rule out in context A, and thus it is a more rigorous epistemic context. The epistemic standards in the contexts used in inferential contextualism are not so ranked. Rather than thinking of one standard as being more demanding or rigorous than another we are to think of the standards of each context as simply being the appropriate standards for the context. This lack of a hierarchy is a result of a very strong commitment to contextualism on Michael Williams' part. All epistemic claims are to be judged within a specific epistemic context. This leaves no extra-contextual ground from which to compare different epistemic contexts or the standards used within them. If all is contextual then there is no ground outside of context to use to compare different contexts to each other⁷⁰. Of course it is possible to compare two contexts within a third context, but this will not give a definitive answer, since evaluations from other contexts might disagree with this evaluation.

All this leads to the inferential contextualist's answer to the skeptical problem. When one is in a normal context it is true that one knows that one has hands. When one is in a skeptical context it is true that one does not know one is not a brain in a vat or that one has hands. Epistemic skepticism is true, but only within skeptical contexts. Fortunately for those who want to avoid skepticism, there are many other contexts in which epistemic skepticism is

⁷⁰ Williams (1996) p 120 ff.

false and we know any number of things. Thus Williams' concedes that the skeptic is correct, but points out that his correctness, in a skeptical context, is not as bad a result as it originally seemed. Skepticism in a skeptical context does not endanger our knowledge in other contexts.

Comparing Inferential Contextualism to Attributor Contextualism

Duncan Pritchard has argued that inferential contextualism is superior to attributor contextualism in addressing skeptical worries. In this section I would like to look at these arguments, expand on them, and eventually agree with them. However, after agreeing with this conclusion I will go on to show why the context-sensitivity of virtue epistemology is superior to inferential contextualism. It retains the benefits of inferential contextualism, and further can explain why we are interested in using particular inferential contexts, in a way that rests on the value of the social role associated with an inferential context. Thus virtue epistemology is the strongest version of context-sensitive epistemic theory.

Attributor contextualism and inferential contextualism differ in two main ways.. These are, first, the hierarchical ranking of the contexts in attributor contextualism, and, second, the differences in the determiners of context and the way in which contexts can be changed.

The first objection to attributor contextualism is the worry that its hierarchical ranking of contexts grants too much to the skeptic. If the epistemic standards given by different contexts are ranked, then we might rightly ask what the "best" standard says. But the most

stringent standard, the one that will be ranked highest, is the most skeptical possible context. For this is a context in which every possible alternative must be considered. So according to the highest ranked epistemic standard, skepticism is true. And this seems like a victory for the skeptic.

Pritchard explains this worry by making a comparison between absolute terms such as flat or straight and our epistemic terms⁷¹. We may interpret absolute terms as having a single absolute standard. Thus the absolute standard for flatness is something that a table (or a stomach) could never hope to achieve. However we are permitted to use the term flat to describe tables or stomachs that approach the absolute standard, where “closeness” is determined by the context. But however appropriate it is to use ‘flat’ in these contexts, attributions of flatness here are not really true, because they don’t meet the highest standards of flatness. We might worry that a similar standard is applied to knowledge by attributor contextualism. Although it is fine to say that I have hands in a normal context, the knowledge attributed here hasn’t really met the highest standards. Thus we might worry that while it is OK to say that I have hands in a normal context or that a table is flat in a normal context in both cases there are higher standards that have not been met, so I don’t *really* know and the table isn’t *really* flat.

This potential worry for attributor contextualism is not a worry for inferential contextualism because it does not endorse any hierarchy of contexts. Since there is no “highest” context,

⁷¹ Pritchard (2002) p45. For an argument that seeing ‘knowledge’ as an absolute term leads to skepticism see Unger (1971).

we cannot worry that a claim has not met the standards of the highest context. Indeed it is a tenet of the theory that it is not possible to compare epistemic contexts. Without such a hierarchy we need not worry that the skeptical contexts will be ranked as best, or be candidates as the absolute “real” standards for all contexts. Thus inferential contextualism seems like a better response to skepticism on this score.

The second ground on which attributor contextualism and inferential contextualism can be compared is what they take to be the determiners of context and the way in which they think contexts can be changed. Attributor contextualism is committed to the claim that the attributor's situation plays a role in determining context, and to the claim that changes in the attributor's conversational situation can make a change in the epistemic context and hence in the meaning of knowledge attributions. The first type of objection to this feature of attributor contextualism is the claim that contexts aren't changed in this way. Attributor contextualism lays down specific principles that determine when the context will be changed and how. And yet these principles seem to get the results wrong in a number of circumstances. Take for example the Rule of Attention. As we have seen this rule holds that an alternative that is mentioned is not properly ignored. Yet there are circumstances where it seems that we do or would properly ignore these alternatives. An example here might be if a defense attorney were to ask a witness on the stand whether she really knew it was the defendant at the scene of the crime⁷². After all she might have been a brain in a vat and only thought that she was seeing the defendant. In making this assertion the defense attorney is clearly drawing attention to an alternative in an attempt to make it a relevant

alternative. But just as clearly he is failing. It would be perfectly appropriate for the witness to say that in light of her long acquaintance with the defendant and the good lighting conditions at the scene of the crime, yes, she did know that the defendant was there. Thus she might properly ignore the brain-in-a-vat alternative, even though it has been introduced.

This objection is only to one of the specific rules for determining the relevant alternatives given by the attributor contextualism. However it highlights a general issue. It seems that sometimes, no matter how hard we try, conversation cannot change context. Inferential contextualism can explain this phenomenon by its contention that it is difficult to change an epistemic context. If doing history determines a context, then we need to stop doing history in order to change contexts. In a court case the inferential context is determined by explicitly stated rules about standards of admissibility and burdens of proof. Introducing skeptical worries cannot change this context. However there are some conversational moves that might change this context. For example, if someone runs into the courtroom yelling “Fire!” this changes the context from one of trying a case to one of evaluating a potential danger. It would be inappropriate to ask of the person yelling, “Do you know that there is fire beyond a reasonable doubt?” because the epistemic context has been changed. So inferential contextualism can explain how some, dramatic, conversational changes can result in changes in context.

However one might think that the inferential contextualism cannot explain enough of the ways that conversation affects context. Many times subtle shifts in conversation do work to

⁷² This example is from Mark Kaplan (1991).

make changes in context. And this seems to give an advantage to the attributor contextualist. However this apparent advantage leads to another objection that Pritchard raises to attributor contextualism's mechanism for changing epistemic contexts. This is the objection that while changes in conversation affects something, it is not the epistemic context (and by extension the truth value of epistemic claims), rather changes in conversation affect only what it is warranted to assert⁷³. This objection is basically a suggestion that we treat attributor contextualism as a claim not about semantics and hence truth values, but rather about pragmatics. Treated this way attributor contextualism can explain the same data that it originally could because that data is intuitions about what it is appropriate to say in particular situations—that is how warranted it is to assert them in a particular context. By allowing the warranted assertability conditions can vary from the truth conditions of an utterance we can explain intuitions that changes in conversation can make a difference in what it is appropriate to say without granting that they always make a difference in what is true. For example the skeptic holds that assertions of knowledge are never true. Yet even the most dedicated skeptic will find himself making knowledge assertions on a regular basis. For to do otherwise would be to create false conversational implicatures⁷⁴. If he were to answer his spouse's question about the car keys with "I don't know where they are" despite having seen them moments before, he would create the false implicature that he had not recently seen them (and annoy his spouse). Thus pragmatics of knowledge attribution may come apart from the semantics on a skeptical view, and this division may be accepted within other theories. In particular the inferential contextualist may accept this

⁷³ Pritchard (2002) 46-51.

⁷⁴ For more on how false implicatures may be created see Grice (1989)

division, and furthermore accept attributor contextualism, but only as a pragmatic theory. This would result in a theory in which it is the inferential contexts that determine the truth-value, but the factors of attributor contextualism that set the warranted assertability.

This objection to attributor contextualism DeRose calls the Warranted Assertability Maneuver (WAM)⁷⁵. In defending attributor contextualism he is concerned to address this type of argument and to show why it is not applicable in this case. He agrees that some instances of arguments of this type are legitimate, but claims that others are not. Those that are legitimate are generated by appeal to a general principle of conversational implicature. For instance, while it is strictly speaking true to say “possibly p” when you know not P, this is a case where “possibly p” is not warrantably assertable⁷⁶. How do the standards for assertability and the standards for truth come apart in this case? Because the assertion would be a violation of the general rule “Assert the Stronger”. This, DeRose thinks, is a legitimate use of a WAM. However using a jerry-rigged version of a WAM that is tailor made to fit the case you are interested in is not a legitimate use. DeRose holds that in general WAMs are only legitimate when they depend on an instance of a more general rule that is not content specific. De Rose holds that the pragmatic rule that would support a WAM for attributor contextualism would have the form “Assert that someone knows whenever they are close enough to knowing”. But this rule is far too specific, and applies only to instances with a particular type of content—those containing “knows” and its cognates. And so DeRose holds that the WAM type objection to attributor contextualism does not work.

⁷⁵ De Rose (2002).

⁷⁶ De Rose (2002).

However, even if we accept DeRose's standards for a legitimate WAM, its use in the objection to attribution contextualism can still meet this standard of legitimacy. For there is a principle even more general than the Assert the Stronger that explains the pragmatic limitations on warranted assertability in these cases. This is the principle "Do not create false conversational implicatures". One way that we might do this is by asserting the stronger, but another way might be to use knowledge claims in ways consistent with the attributor contextualism interpreted pragmatically. For example we might consider the case above of the skeptic who nonetheless ought to assert that he knows where his spouse's keys are. To do otherwise would be to create the false implicature that he had not recently seen the keys, or did not justifiably believe the keys to be on the counter. In addition to forcing knowledge attributions in cases where they are not thought strictly speaking true, worries about creating false implicatures can also render inappropriate knowledge attributions that are strictly speaking true. Returning to the example of doing history as an inferential context, we might consider a case in which someone inappropriately brings up a skeptical worry in this context. In response to the claim that Jane brought this pot shard into the lab, the skeptic might point out that it might not have been Jane, but someone cleverly disguised as Jane. It would then be inappropriate for the historian to reply "I know it was Jane" even though strictly speaking in this context of inquiry she does know it was Jane. This is because her statement would create the false implication that that the historian had special evidence to refute the skeptic's worry. Since she has no special information that appropriate response would be to simply dismiss the skeptic's objection, rather than to assert knowledge and create a false implicature. These two cases show that the conditionals of warranted assertability can

come apart from the truth conditions for knowledge attribution, and do so in a way that is governed by a very general principle. Trying to avoid generating false implicatures will affect the assertability of particular utterances, while not affecting the truth values of the sentences those utterances contain. Thus we may use a WAM to show that attributor contextualism can be used as a theory of pragmatics rather than as theory of truth values.

These considerations taken together give us good reasons to prefer inferential contextualism over attributor contextualism. Attributor contextualism grants too much to the skeptic, and the benefits of attributor contextualism can all be preserved by accepting a more moderate interpretation on which the conversational context determines the pragmatic appropriateness of assertions not their truth values.

Comparing Virtue Epistemology to Inferential Contextualism

Having ascertained that inferential contextualism is a better version of context-sensitivity than attributor contextualism, we shall now compare it with the context-sensitivity of my virtue epistemology. This comparison will show that virtue epistemology has all of the strengths of inferential contextualism and in addition has a further strength that makes it preferable.

Let us compare virtue context-sensitivity and inferential contextualism on the basis of the two types of objections raised to attributor contextualism. Remember that the first type of objection to attributor contextualism was a worry that it ends up granting too much to the skeptic because of the hierarchical ranking of contexts it offers. By admitting that the

standards in the skeptical context are “best “we introduce a worry that those standards ought to be met in all contexts. This objection does not apply to virtue context-sensitivity because the social roles used in this version of context sensitivity are not ranked in a hierarchical way. There may be a division in social roles such that some possible social roles have less value than others. For example the social role of a con artist or a forger is one that would be judged to have less value on the basis of being incompatible with a number of moral virtues. The social roles of being a passive believer of television commercials or government propaganda would likewise be social roles that are less valuable because they are inconsistent with the epistemic virtues. However there are a large number of social roles that do not conflict with either moral or epistemic virtues, and these social roles can be taken to be on a par with each other. Thus, although virtue context-sensitivity does allow for some ranking by making a division between desirable and undesirable social roles it does not provide the hierarchical structure needed to give skeptical worries a toehold. The social role of the skeptic or of the philosopher in general will not be more valuable than the social role of the doctor, teacher, parent, or garbage collector. Each of these social roles will have subject matters that are appropriate to that social roles, and standards of proof within and outside of that subject matter that are appropriate as well. An agent will be doing better in a social role insofar as she conforms to these standards. Of course a person can be a better or worse doctor or teacher. Furthermore it may be true that for a particular person one social role is better than another. But this is because a particular person may be better suited to the demands of one role over another, and not because the role itself is more valuable. Thus lacking a hierarchy of roles virtue context-sensitivity is not subject to the worry about skepticism.

The second type of objection raised to attributor contextualism also does not apply to virtue context sensitivity because it is not offering an approach on which conversation can radically change the epistemic context. Like inferential contextualism, virtue context-sensitivity can explain why it is sometimes the case that dramatic conversational changes can change the epistemic context. In these cases the conversation does this by making another of the agent's social roles salient. When someone shouts "Fire" in a courtroom, this brings salience to the social role of a self preserving member of society, and cues looking for exits and trying to help others out of them in response to the epistemic and moral commitments of that role. Thus conversation can change the context, but only by calling attention to an already existing social role of the agent, and this means that such changes are quite limited. However, just like inferential contextualism, virtue context-sensitivity can take on the benefits of attributor contextualism understood as a pragmatic rather than a semantic theory. Thus virtue context-sensitivity's variation by social roles can be supplemented by an explanation of why it is sometimes inappropriate to make knowledge assertions even when someone does know, or appropriate to make such assertions even if the person doesn't know. Thus the power of attributor contextualism can be harnessed by virtue context-sensitivity without taking on its semantic baggage.

Thus far virtue context-sensitivity and inferential contextualism are on a par. The difference comes when we search for an explanation for why people would take on the epistemic burdens they do when they enter a particular inferential context. Why take on these

burdens, when one could avoid them by simply working within another inferential context?

Why not always operate in a inferential context with a very low burden of proof?

Virtue context-dependence can explain why we take on these burdens by appealing to the costs and benefits associated with a social role. We take on the burdens of a given social role because we want the benefits associated with that same social role and we cannot have one without the other. This is true in both the moral and the epistemic realm. When someone takes on the role of mother she takes on an immense amount of responsibilities towards her child. Why would anyone take on these burdens voluntarily? Because the burdens are part of being a mother and having a child of one's own. Thus in order to gain the benefits of being a mother one must also take on the burdens, as they are both inseparable from the social role. The same can occur in the epistemic realm. Take the example of a historian. Why would she voluntarily put herself in the inferential context of doing history rather than the inferential context of a casual conversation? To do the former is to take on more epistemic burden. We can explain this behavior by pointing out that the person in question takes on this burden because she wants to be a historian.

In order to be a historian she needs to operate within a particular inferential context. Thus our desire to take on certain roles and to obtain the benefits associated with them can explain why we take on moral and epistemic burdens that we might otherwise avoid.

This explanation according to social roles can explain not just why we take on roles in the first place, but also why we hold ourselves to these standards after we have taken on the role.

On a day when she has a headache, a historian might not want to hold herself to the

standards of the inferential role of doing history. And yet, she does so. Why? If we think of inferential roles as unconnected or something we can voluntarily take on this is hard to understand. However if we think of them as following from social roles to which we have already committed ourselves, it is easier to explain why we hold ourselves to these standards even when we would rather not do so.

Thus virtue context-dependence has an advantage over inferential contextualism in that it can use social roles and their relationship to the person's life as a whole to explain why we would take on epistemic burdens that it is possible to avoid. At the same time virtue context-dependence avoids the problems that beset the semantic interpretation of attributor contextualism, while retaining its usefulness as a pragmatic theory. Virtue context-dependence is thus the best model to accommodate the various data that lead us to incorporate context-dependence into our epistemological theories.

CHAPTER 5: VIRTUE, REDUCTION, AND TESTIMONY

Testimony, construed broadly as all assertions (not just those in a legal context), plays an enormous role in each of our epistemic lives. Think of a college campus on a typical day. Students attend lectures and come to believe many things on the basis of the testimony of the lecturer. But testimony does not stop there, for the lecturer is probably speaking of some things she has not experienced. She may be reporting the results of experiments performed in another country or events from long ago. The lecturer herself has learned these things from the written or oral testimony of others, and this testimony in many cases is a report from written or oral testimony. By carefully considering the source of our beliefs, we can see that a very large portion of the things we think we know we have come to believe through testimony.

Since much of the evidence for our beliefs comes from others through their testimony, we may well ask if testimony can give us a proper justification for these many beliefs. This project is important because if testimony is not a sufficient ground for justified belief, many, if not most, of our beliefs will not be justified. And yet the role of testimony in justifying our beliefs has been largely ignored in traditional epistemology. Instead the focus has been on the individual knower, and the things that she is able to learn through her own experiences. Testimony as a source of justification and knowledge has been left out of this picture. Happily however there has been a recent resurgence of interest in the role of

testimony in epistemology. Part of this resurgence has been spurred on by C.A.J. Coady's Book *Testimony: A Philosophical Study* (1992).

In this book Coady has two main goals. The first is to establish the importance of testimony. Coady does this by demonstrating that much of what we know about the world around us reaches us through testimony. Given the importance of this source of knowledge, he argues that we must understand how testimony can serve to justify our beliefs in order to understand the foundation of much of our knowledge. Coady's second goal is to argue in favor of his own account of the justification of testimony. To do this Coady first clears logical space by demonstrating the insufficiency of competing views of testimony. Coady takes a historical survey of traditional approaches to testimony, explaining and criticizing existing views. Although he discusses the positions of many authors, Coady's primary target is Hume. Hume takes a reductionist position on the justification of testimony, on which the justification given by testimony can be reduced to the justification given by past experiences with testimony. Since Coady's own position is an anti-reductionist one (on which testimony is a special source of justification), he is quite concerned to show why Hume's account is insufficient.

Coady's arguments are a microcosm of the general debate about the justification of testimony. I will respond to Coady's arguments against Hume, and then use this response as a springboard to discuss the benefits of a virtue epistemological approach to testimony. A virtue approach highlights the depth of our understanding of other human beings, and explains how we use this understanding to justify our trust in the testimony of others. In

some of his writings Hume hints at the idea that we can use our understanding of others in evaluating their testimony, but this idea can be given a fuller development and justification in the framework of the virtues. The virtue approach is a natural extension of the Humean position, and so I will first be concerned to defend that Humean position from Coady's objections. Finally, I will show how virtue epistemology can extend the Humean position to provide us with a proper justification for our numerous beliefs based on testimony.

It is clear that both Hume and the virtue epistemologist would agree with Coady's first point that much of what we know we know by testimony. Hume, in particular, recognizes the importance of testimony in our epistemic lives (E111)⁷⁷, and hence puts great effort into providing an explanation of the role that testimony can play. This explanation occurs both throughout the *Treatise*, and especially in the essay "On Miracles" from the *Enquiries*. The real debate here is about the way that testimony is justified. In order to support an anti-reductionist position, Coady presents three main objections to Hume's account. While the first of these three objections raises an interesting point of caution, Hume has an available response to this first objection, as well as to the remaining two objections. I will demonstrate that none of the three objections offered by Coady gives us a reason to reject the basic Humean structure of the justification of testimony. Rather we ought to accept this structure of justification, and supplement it with developments in virtue epistemology.

⁷⁷ All references to Hume's *Enquiries Concerning Human Understanding and Concerning the Principles of Morals*, are to the Oxford University Press Edition of 1975, and will take the above form.

The Humean Position

Before I introduce the Coady's objections it is important to understand the reductionist position he is attacking. The Reductionist Thesis that Coady finds in Hume is based on his claim that:

It will be sufficient to observe that our assurance in any argument [from testimony] is derived from no other principle than our observation of the veracity of human testimony, and of the usual conformity of facts to the reports of witnesses. (E111)

This claim is simply a special instance of Hume's general claim that when reasoning about matters of fact, our only source of justification is our experience. When applied to the case of testimony this means that the only available source of justification is our past experiences of testimony, along with our experiences of the correlation of those pieces of testimony to the world. Much as in other cases of causal reasoning, our past experience of this correlation leads us to an association of ideas, here between the testimony and the fact to which it testifies.

Elizabeth Fricker in her review of Coady's book makes a useful distinction by pointing out that the Reductionist Thesis can be broken down into two distinct claims: the claim that reduction is necessary, and the claim that reduction is possible⁷⁸. The quotation above indicates that Hume is committed to the claim that reduction is necessary. If we cannot ground our acceptance of testimony in past experiences then Hume would claim we are not justified in accepting and reasoning on the basis of testimony. However, the claim that

reduction is necessary on its own does not give us any evidence that we can actually perform such a reduction, only that such a reduction is needed if we are to be justified. Skeptical worries will loom until we establish that reduction is not only necessary but also possible. Those who think of Hume as a skeptic in many areas might be inclined to wonder if Hume need really be committed to the claim that reduction is possible. Perhaps he is setting up a standard that he believes cannot be met, but is unfazed by the skeptical consequences of employing this standard. Hume's belief that reduction is possible is demonstrated by his claim that in some cases testimony can give us entire proof for a claim (E115). If reduction were a standard that could never be met, testimony could never serve as evidence for, much less as a proof of, a claim. Furthermore, Hume himself is concerned that people will take his standard as evidence that we are not justified in our beliefs about history (T1.2.13.4)⁷⁹. Hume wants to show that skepticism about all our historical knowledge would be an absurd conclusion and spends some time explaining why this conclusion does not follow from his account of testimony (T1.2.13.6). Thus it is clear that Hume would not endorse taking a skeptical route out of the problem, but rather is committed to both the necessity and possibility of reduction. Both together make up his full position on the issue, the Reductionist Thesis.

Coady's First Objection

Coady's first objection to the Reductionist Thesis is that it is vague, and that Hume seems to understand it in two different ways. When we limit the source of our justification to

⁷⁸ Elizabeth Fricker, (1995).

experience, we must be clear what kind of experience we are limiting ourselves to. We may either be limiting ourselves to our own personal experience or we may be limiting ourselves to experiences shared in our community. Coady offers textual evidence to support the claim that, intentionally or not, Hume appeals to experiences that are not his own, but rather those of his community, in his arguments. One example of this is Hume's appeal to our experience that sometimes people in good health die unexpectedly:

It is no miracle that a man, seemingly in good health, should die on a sudden: because such a kind of death, though more unusual than any other, has yet been frequently observed to happen. (E115)

Coady points out that while it is not unusual for a healthy man to die, this is probably not an event that Hume ever observed on his own, much less "frequently observed". By appealing to this sort of observation Hume is allowing himself to use the experiences of others in his community, and not strictly limiting himself to his own experiences. However Coady argues that this interpretation of the Reductionist Thesis is "involved in vicious circularity"⁸⁰. For the evidence that we can obtain from the observations of others reaches us through testimony—the very type of testimony that we are trying to justify. Until we have given a justification for accepting the evidence of testimony, we cannot use the experiences of others as evidence, and as a result we must limit ourselves to our own experiences in our quest to justify testimony. Thus the only intelligible interpretation of the Reductionist Thesis is one that limits the experiences used in reduction to an individual's own experiences.

⁷⁹ All references to Hume's *Treatise of Human Nature*, are to the Oxford University Press Edition, edited by Norton and Norton (2000), and will take the above form.

While this is a legitimate objection to using the testimony of others in our fundamental justification of testimony I think that it is misinterpreting Hume. Using the testimony of others is appropriate in the instance cited above, as well as in others cited by Coady (T1.4.2.20), because Hume is not in these passages engaged in the foundational problem that Coady takes as his primary topic. There are times when it might be appropriate to engage in a Cartesian project of rejecting our questionable beliefs, and then building the justification for those beliefs up from the foundations, but in the passages Coady quotes Hume is simply using the already established practice, or the conventionally accepted practice, of accepting evidence from the testimony of others. Since in these passages Hume is instead concerned with specifying what characteristics are needed to make an event truly miraculous, and not focused on justifying testimony, his use of testimony here does not involve him in vicious circularity. Hume could also agree that the proper interpretation of the Reductionist Thesis should limit the experiences used in justification to those of a single individual. These limits are in play when we are engaged in a foundational project, and so in that context we may only appeal to our own personal experiences. This is more than a question of interpreting Hume's position. It shows us a limitation on any attempt to provide a reductionist justification of testimony. Such a theory must limit itself to the personal experiences of a single individual if it is to provide an adequate justification for that individual to believe the contents of testimony. However, we will see that this limitation is not fatal to a reductive account of justification. Hume works within this limitation to provide a justification for testimony, and a virtue approach can give an even stronger justification for testimony under the same limitation.

⁸⁰ Coady (1992) p 81.

Coady's Second Objection

Coady claims that, when properly limited to our own experiences, the Reductionist Thesis is “plainly false, because it seems absurd to suggest that, individually, we have done anything like the amount of field-work that [reduction based on individual experience] requires”⁸¹.

Coady objects that there are many types of events and facts that we accept on the basis of testimony, yet have no experience of ourselves:

...many of us have never seen a baby born, nor have most of us examined the circulation of the blood nor the actual geography of the world nor any fair sized sample of the laws of the land, nor have we made observations that lie behind our knowledge that the lights in the sky are heavenly bodies immensely distant nor a vast number of the other observations that [reduction based on individual experience] would seem to require.⁸²

Each of these is a piece of purported knowledge based on testimony that we cannot, given our limited experience, check up on. The heart of the objection here seems to be based on a comparison in the size of the reference class of instances of testimony, to the size of the sample which we are able to check on our own. Given that testimony plays an important role in our epistemic lives, the class of things learned from testimony is immense. We depend on the testimony of others because we are unable to learn many things first-hand. But this means that the fecund method of learning from testimony far outstrips the number of instances that we can verify using our personal experience. In essence the sample we can

⁸¹ Coady (1992) p 82.

⁸² Coady (1992) p 82.

check up on is too small to support any induction about the reliability of testimony in general. Thus the Coady claims that the Reductionist Thesis is false because reduction is not possible.

This objection from Coady is an interesting worry about induction, but seems to be based on ideas about induction that are not Hume's own. Hume is quite serious when he says that we must judge from past experiences alone. We must judge based on the on the proportion of confirmed to disconfirmed bits of testimony in our past experience. On Coady's approach to induction, past experiences count, but so do the past experiences that we could have had but did not. He seems to be concerned not only with the confirmed or disconfirmed bits of testimony in our past experience, but also with all the unconfirmed bits of testimony in our lives. This is why the number of bits of testimony we have not checked is relevant. From his comments above, it seems that according to Coady, appropriate sample size depends on a comparison of the number of past experiences with number of possible experiences. Coady holds that if our experience is dwarfed by the number of possible experiences then our inductive base provides less support for a general claim, even if every past experience in that inductive base supports the claim. On this account of induction, the fact that dependence on testimony is so prevalent in our lives means that for any induction to properly support the veracity of testimony generally that induction would need to be based on a very large sample of pieces of testimony that conform to our own experience. Based on this approach to induction, Coady's claim that none of us do enough field-work to support a proper induction is understandable, and perhaps justified.

However this is not the notion of induction that Hume is working with, nor is it a picture of induction that we ought to accept. Hume is quite serious about our judgments being based only on our past experiences. This means that we need to weigh our past experiences against each other, but we do not need to consider all the possible experiences that we could have had but are lacking. Rather we are to take the balance of our past experiences and use that balance to understand the total evidence of our past experiences.

In all cases, we must balance the opposite experiments, where they are opposite, and deduct the smallest number from the greater, in order to know the exact force of the superior evidence. (E111)

In cases in which all of our past experiences agree we can take this evidence to amount to a “full proof” of the future existence of a similar event (E110). However, as is more common, our past experiences may contradict each other. In this case the past experiences do not amount to a full proof but rather to a probability, which is proportionate to the number of past experiences on each side. As a result, both proof generated from agreement of past experience and probability generated from disagreement of past experience are generated in the same way. In each case we expect the future to resemble the past. In the case of agreement, we expect the future to resemble what has occurred in all the past experiences. When the past experiences are mixed we expect the same proportions to hold in the future:

Any of these past events may again happen and we judge that when they do happen, they will be mix'd in the same proportion as in the past. (1.2.12.10)

Thus all reasoning by induction follows the same pattern, and this allows Hume to give a unified account of induction, which he applies to the special case of reasoning about testimony.

Admittedly the assurance that an induction gives us is based in part on the number of past instances that we have experienced. Hume does not recommend that we make an induction from a single experience or from a very small number of experiences. Rather:

The first instance has little or no force: the second makes some addition to it; the third becomes more sensible; and 'tis by these slow steps, that our judgment arrives at a full assurance. (1.3.12.2)

So Hume does seem to be committed to the claim that we need a suitably large sample in order to ground a proper induction. However, it is important to note that this sample size need not be proportional to the number of possible experiences. Rather an induction based on a suitably sized sample is acceptable regardless of the proportion that sample bears to the total number of possible experiences.

Given Hume's own account of induction, we can now see why Coady's claim regarding the impossibility of reductive justification of testimony misguided. For his objection depends on the relative proportions of testimony confirmed, and testimony unconfirmed. Since we accept enormous amounts of testimony, we will always have many instances of unconfirmed testimony, and this proportion will always be small. So the induction will never serve as a proper foundation for the justification of testimony. However, since Hume is not concerned with this proportion, our extensive reliance on testimony, and resultant large number of unconfirmed cases of testimony, cannot serve to undermine the induction in his system. So long as we have a suitable sample of pieces of testimony that we have personally verified, we can make a proper induction on that sample which can support a claim that

testimony in general is reliable. So on a proper account of induction, the possibility of reduction is not closed off.

The possibility of reduction, furthermore, is supported by the fact, noted by Hume, that we have a potent source of information about the behavior of others in ourselves. Coady does not discuss this possibility, rather limiting the relevant experiences from which we might make an induction to observed instances of the testimony of others. But our own testimony in past instances can serve as a fairly large inductive base, each instance of which can be used (so long as we were not testifying about something we knew nothing about) in the sample of testimony which the individual has verified. For many instances of our own past testimony we know whether that testimony was true or not, and thus our past testimony serves as a source of information about testimony in general.

In addition to our own past testimony, we can also use information based upon our understanding of our own psychology. For, although the relation between past testimony and reality is one potent source of inductive evidence, the relation between our motives and our actions is another source from which we can also draw inductive evidence:

Were not the memory tenacious to a certain degree; had not men commonly an inclination to truth and a principle of probity; were they not sensible to shame, when detected in a falsehood: were these not, I say, discovered by experience to be qualities, inherent in human nature, we should never repose the least confidence in human testimony. (E112)

How can we use experience to discover these characteristics of human nature? We can examine our own psychology and notice that we are inclined to tell the truth and that we fear discovery if we tell a lie. We can also notice the effect these characteristics have on our own testimony. Once discovered in ourselves, we may infer from the behavior of others that they too have these psychological characteristics. And, since people's psychological characteristics determine their actions (as Hume points out in (T2.3.1)), we may infer that the same effects these characteristics have on our own actions will be experienced by others too. Thus, by looking at our own testimonial behavior, as well as at our own psychological make-up, we can gain further evidence which can be used in the inductive justification of accepting testimony.

While Coady has framed the problem as one of finding a foundation for all instances of testimony, our inductive basis may be stronger, and more properly sensitive if we focus not on testimony in general, but rather on types of testimony. Coady anticipates this move, but objects that there are no sensible divisions into types of testimony that can reasonably be used in our analysis of the testimony presented to us. Coady considers two possible groupings on which we might make more specific inductions: by kinds of speaker or by kinds of content⁸³.

Coady objects that trying to group testimony by kinds of speaker is unhelpful. It would certainly not be helpful to group speakers by their height, nationality, gender, etc. The only helpful grouping would be in terms of expertise. This would allow us to justify accepting

testimony, at least of the people who are most expert in the area in question. However, Coady thinks that breaking the class of testimony into kinds by the expertise of the speaker is not helpful because it is not a characteristic that is available to us in making most of our judgments. Although some speakers may be granted expert status by our community, we can only accept their expertise on the basis of further testimony, leading to a regress.

Dividing testimony into kinds by the content of the testimony is also not helpful, because it is simply too difficult to apply. Coady takes as an example receiving the piece of testimony “There is a sick lion in the Taronga Park Zoo”⁸⁴. What kind of content does this assertion have? It could be classed as an existence claim, as a geographical claim, as a medical claim, etc. Classing this piece of testimony as falling within these different kinds may lead us to have different evidence for that claim based on past experience. Perhaps classified as one kind it is supported by past experiences, but when classified as another kind the relevant past experiences do not support it. So it is important to discover which kind a given piece of testimony really should be classified as. But Coady argues, this is a hopeless enterprise, and as a result dividing testimony by content-type will not help us to justify even specific kinds of testimony.

While Coady’s objections to the kind distinctions considered above may show that they are not a useful way to divide testimony into kinds, Coady overlooks divisions suggested by Hume that have more promise of being helpful. When discussing the evidence that

⁸³ Coady (1992) p 83-85.

⁸⁴ Coady (1992) p 84.

testimony can give us, Hume presents some characteristics of testimony that may lead us to doubt it:

We entertain a suspicion concerning any matter of fact, when the witnesses contradict each other; when they are but few, or of a doubtful character; when they have an interest in what they affirm; when they deliver their testimony with hesitation, or on the contrary, with too violent asseverations. (E113)

The first two of these characteristics of testimony are already covered in the general considerations about induction. When we encounter contradictory testimony we know that only one piece of testimony presented can be true, and increasing numbers of witnesses improve our support for the testimony since each piece of testimony brings with it the inductive evidence of the justification for all testimony. The remaining four characteristics of testimony may however serve as a guide as to how to divide testimony into relevant kinds. Let us take the manner in which someone testifies as an example. When reasoning about a current piece of testimony received we may note, for example, that the speaker delivers her claim with hesitation. This fact about the piece of testimony can be relevant when deciding what past experiences of testimony to consider. Rather than forming an induction based on all our past personally verified instances of testimony, we can consider only those past instances of testimony delivered in a hesitating manner. If, as seems likely, our past experiences with hesitating testimony have a bad track record, we may place little faith in the piece of testimony at hand. Our lack of faith can be reasonable based on an induction on the proper sub-set. By dividing testimony into kinds by these characteristics we may then arrive at a more sensitive justification based on the kind of testimony presented. Note also that the characteristics offered by Hume are, for the large part, accessible to the agent who is

considering a piece of testimony. Unlike expertise which is hidden from the agent and which can only be confirmed by other testimony, manner of presentation, conflicts of interest, and the character of the witness are all in principle available to the agent without recourse to other testimony. Thus division of testimony into kinds can be helpful in fine-tuning our inductions, and as a result we may be able to justify accepting some types of testimony while rejecting other types.

Thus we see that Coady's second objection is wide of the mark. While Coady's approach to induction might undermine the possibility of justifying testimony inductively, on Hume's own account of induction this possibility clearly remains. In addition, while Coady is concerned to justify all instances of testimony with a single general support, Hume would endorse an approach that is more sensitive to differences between kinds of testimony. By dividing testimony into types, we may provide an inductive basis for some kind of testimony while finding no basis for other kinds of testimony. This interpretation of induction along with a division of testimony into kinds makes the claim that reduction is possible far more plausible.

Coady's Third Objection

Finally Coady's third objection directly takes on Hume's claim that:

The reason why we place any credit in witnesses and historians, is not derived from any *connection*, which we perceive *a priori*, between testimony and reality, but because we are accustomed to find a conformity between them. (E113)

This claim, Coady says, implies that the results of our induction might have been other than they in fact are. If our justification for accepting testimony is really based on experience alone, then it could have turned out, given other experiences, that “there was no conformity between testimony and reality”⁸⁵. Coady wonders what such an induction would look like, and by considering an example, concludes that there is no set of experiences that could lead us to this conclusion. Hence Coady concludes that the justification of testimony is not merely the result of experience, but rather depends on some *a priori* connection. This is a repudiation of the claim that reduction is necessary. If there is an *a priori* source of evidence, then not all evidence justifying the acceptance of testimony comes from experience.

To this end, Coady develops an example of a group of Martians whose apparent reports consistently misrepresent the world. They are constrained by some element of their psychology to never make a true statement. Interestingly this means that after one of them says falsely “There is a tree in the yard”, none of them can object, “No there isn’t”, or even “You are wrong” because this would be a true piece of testimony. In such a world, Coady maintains that the Martians have four good reasons not to accept the testimony given them by others⁸⁶. First, the testimony they receive never conforms to their own experience. Second, acting on testimony received leads them astray in practice. Third, introspection indicates that they are unreliable, and they can reason from this feature of themselves to features of others. Fourth, testimony received is often chaotically divergent whenever there is more than one way to lie.

⁸⁵ Coady (1992) p 85.

⁸⁶ Coady (1992) pp 86-87.

Given this evidence, it seems that the Martians should conclude that testimony is completely unreliable. However, Coady claims that the very information that would prove this claim also serves to undermine the claim that the Martians make any reports at all. Coady holds that reporting is a speech act that implies a certain intent, either to truly inform the hearer, or to make the hearer believe whatever is reported. But in the Martian world, there is never any evidence of such an intention. Persistent falsehood shows that the Martians do not intend to truly inform, and an environment of falsehood ensures that reports do not make the hearer believe their content. Thus the Martians cannot believe that all reports are false and untrustworthy simply because they have no reason to believe that any one is ever making a report.

Furthermore, while Coady's original formulation of this example assumed that the content of the Martians' language was fixed, upon dropping this assumption it seems that the Martian language has no content at all. Coady thinks it is untranslatable from our point of view. When translating we should implement a Davidsonian principle of charity, trying to make as many of their utterances true as is possible. However given that all of their utterances are false, any translation that we make of their utterances based on the principle of charity will be radically incorrect. Also, given that falsehoods may be quite divergent, we may not see enough patterns in their language to even begin a translation. Thus their language would be nonsensical, in that it would have no externally accessible content.

Beyond this problem with translation from the Martians' language into any other language, Coady holds that there is also a problem within in the community with language acquisition. We learn our first language by being presented with many true reports, either direct or indirect, about the terms that correlate with objects. A child has to hear the word "ball" associated with many different balls before she can learn the meaning of the word "ball". In the Martian community an infant will hear the words "chair", "truck", "dog", etc. associated with the ball, but never the word "ball" itself. In such a situation it seems very unlikely that she will learn the meaning of the word "ball". Even if she had some sort of mechanism that picked out only words that she never heard associated with an object as the name of that object, there are many other word besides "ball" which she might take to name the ball, since her experience will be finite.

Coady argues that these two worries, about translation and about language acquisition, taken together show that the Martians' language has no real content. If the Martian language lacks such content, then we cannot say that their reports (or pseudo-reports) are constantly false. Rather the most we can say is that their reports (or pseudo-reports) are meaningless. Thus even in an instance with perfect anti-correlation, reasoning from past experiences cannot lead us to the conclusion that there is no conformity between testimony (understood as reports with the specific intent to convince) and reality. It can only lead us to the conclusion that there are no reports with the appropriate intent at all, or that the utterances we took to be testimony had no meaning. Coady concludes that since we cannot discover a lack of conformity between testimony and the world, the connection between the two must be based not on experience, but rather on some *a priori* relation.

While Coady's argument points out an interesting connection between our utterances counting as reports and the truth of those utterances, it does not show that the justification of testimony must be *a priori*. The problem here rests in an ambiguity between two understandings of the word testimony in the goal of discovering, a lack of "conformity between testimony and reality". The first reading would be *de dicto*. On this reading discovering a lack of conformity would mean also discovering that the word "testimony" had no reference. Hence it would be impossible to discover that *testimony* fails to correspond with the world—there would be no testimony to so fail. This is the reading of the goal that Coady seems to be employing in his claim that experience cannot disconfirm the correlation between testimony and reality. However I believe that the correct reading here should be *de re*. That is, "testimony" should refer to the utterances that we currently take to be instances of testimony, regardless of whether they turn out to be true instances of testimony. In the Martian world these utterances turn out to have no conformity with reality, in part *because* they turn out to have no meaning. Coady's argument then shows us an interesting correlation between two empirical facts; if we are to have meaningful testimony reports, then those pieces of testimony must, to some extent, correlate with reality. But this does not mean that the correlation of testimony to the world is *a priori*. The most it could mean is that the claim, "If there is testimony then it is correlates to reality" is *a priori*. We learn the antecedent of this conditional by experience, perhaps at the same time that we discover the consequent also by experience. (Although the antecedent and consequent may have an *a priori* relation it is most unlikely that we should actually learn the consequent by way of the antecedent. If we did then Coady's argument here would be old hat rather than the

intriguing suggestion that it is.) Thus both facts, though *a priori* related are learned on the basis of experience.

On the correct reading, the claim that “testimony is correlated to reality” can be disproved by experience. As a result the truth of the claim can be discovered only by experience, and this is in keeping with Hume’s original position in the Reductive Thesis. Thus the last of the three objections to the Reductive Thesis offered by Coady fails, and Hume’s account of the justification of testimony remains as a viable alternative to anti-reductionism, undermining the reasons Coady gives for a move from Hume to his own anti-reductionist position.

Using Virtue Epistemology to Justify Reliance on Testimony

In Coady’s second objection we saw worries about reduction not being possible because of the limits of our induction base. The Humean response to this type of objection is to notice that if we believe that others are relevantly psychologically like us, then we have an extra source of evidence in our own behavior and psychological dependencies. We can look to ourselves and to our own motivations, and then project those motivations onto others. This helps us both to explain the behavior of others, and to predict the actions that other’s might take. In the realm of testimony this means that by understanding our own psychology we are able to justify our predictions about the testimonial behavior of others. If we know of ourselves that in a similar situation we would be motivated to tell the truth because of our fear of being caught in a lie, we can infer that the person from whom we are receiving testimony may well be motivated in the same way. Thus by making connections between

our psychology and the psychology of other we have a more potent source of information about their behavior, especially that which involves testimony.

Still there is a worry that other people might have psychological set-ups that are quite different from our own. We gain evidence of this by interacting with others, and sometimes being surprised at their reactions and behaviors. (We might also find a potent source of this type of information about others by observing the individuals on daytime talk-shows.) So we see that the information provided by a parallel between ourselves and other can be limited. Of course there are some basic motivations that we think all people have, but the variance in other motivations may impede our ability to predict the behavior of others on the basis of our own psychology.

Taking a virtue approach to both epistemology and ethics gives us a further source of information about the actions of others. Rather than making a connection between our psychology and the psychology of others, a virtue theory makes a connection between a given individual's moral behavior and epistemic behavior. The virtues interact and facilitate the development of each other. An individual developing the moral virtue of generosity needs to have at least some of the epistemic virtues to allow her to figure out how to best and most effectively be generous. An individual developing the epistemic virtue of open-mindedness needs to have the moral virtue of fairness, and needs to not be biased, in order to listen to the opinions of others in a fair manner. This characteristic is called the reciprocity of the virtues, and it is a feature of most ancient and many modern theories of

virtue⁸⁷. A more extreme version of this claim is called the unity of the virtues. This is the claim that one cannot possess any of the virtues without possessing them all. The unity of the virtues is based on the idea that if one lacks any virtue, then the remaining putative virtues had by the agent are less well developed than they could be; there is a vice in the agent that is waiting to potentially trip up the exercise of the virtues in question. According to the unity of the virtues then, if we know that the person we are receiving testimony from is a morally virtuous person, we can know that this person is epistemically virtuous as well, and are therefore justified in accepting their testimony.

This strong connection between the moral and epistemic virtues helps us in our search for justification for accepting testimony because it greatly enlarges the number of experiences with an individual that count as being relevant to our judgments about their epistemic competence. On more traditional accounts of testimony, moral considerations plays only a limited role, being confined to considerations about whether the speaker is telling what he or she believes to be the truth. We might be interested in an agent's moral track record to discover if he or she is intentionally lying to us. The virtue approach respects the relevance of these considerations, but further claims that an agent's past moral track record is also relevant to his standing as an epistemic agent. Since those with moral virtues have the epistemic virtues as well, moral virtues are directly relevant to trusting someone who gives testimony not just to say what she believes, but also to believe correctly in the first place. This direct relevance of the moral virtues allows us to make use of the information that we

⁸⁷ For a good discussion of the reciprocity of the virtues and the unity of the virtues in ancient Greek philosophy see Annas (1993) pp 73-84.

have about an agent's moral character in evaluating him epistemically. This connection can help us with our original problem of lacking enough evidence to make a Humean reduction possible. We can now use our evidence about moral character, and in many cases this information is far more accessible than information about an agent's epistemic track record. Though we rarely check pieces of testimony directly against the world in order to discover someone's epistemic track record, we more often observe an agent's moral behavior and come to conclusions about his character. In a virtue framework that employs the unity of the virtues, we can use these conclusions to further justify accepting the testimony of others.

However, one might object to the strength of the unity of the virtues claim. First one might think that although it might be difficult to develop a virtue while lacking other virtues, such development still might be possible. Perhaps you have developed a certain strength of will that will protect the developed virtue from the temptations of your other vices. Or perhaps some virtues are so unrelated to each other that one need not have one to develop the other. A second objection to the unity of the virtues is that the virtues described by this theory would be very hard to discern in others. Perhaps my co-worker seems to me to be a great epistemic agent. I admire his methods and skill in judgment. However, unbeknownst to me, he is quite cowardly, but only in a type of situation I have never seen him in. In such a case I would be tempted to think that my co-worker really did have the epistemic virtue, say, of open-mindedness. However I would be mistaken in assigning this virtue to him. Thus the unity of the virtues thesis would leave us frequently making mistakes in our attributions of virtues—So much so that we might doubt that this strong version of reciprocity captured our original notion of the virtues.

Fortunately for the use that we want to put it to, we can retreat from the unity of the virtues to the weaker thesis of the reciprocity of the virtues. This is to claim that while it is not impossible to have one virtue while lacking any other virtue, it is quite difficult. The plausibility of the reciprocity of the virtues can be supported by an analysis of the development of the virtues. Perhaps it is possible to develop the virtue of open-mindedness without the virtue of fairness. Perhaps the agent might scrupulously keep an accounting of who he should be open-minded towards, even if he is extremely biased against some sub-group. But even if this is possible, it is not very likely. The easiest and most natural way to develop a virtue is to develop it in parallel with other virtues, and to allow the interplay between the virtues to reinforce each other. Thus the presence of one virtue is a good indicator of other virtues, although not a perfect one. Likewise, the worry about being able to identify virtues stated above is a result of the unity of the virtues thesis and does not affect the reciprocity of the virtues thesis. It is only if possession of a single virtue requires possession of them all that we encounter worries that a vice in an unrelated part of someone's life might mean that the virtue we think they have cannot be a virtue at all. On the reciprocity of the virtues thesis our judgments about virtues need not be entirely undermined by extra evidence of vice that we might uncover.

Conclusion

In this chapter I have defended Hume's reductionist position on the justification of testimony for two reasons. First it deserves a defense against the objections presented to it by Coady. The objections given do not undermine Hume's position on testimony. Second,

I have defended Hume's reductionism as a stepping stone to the more comprehensive account of the justification of testimony offered by virtue epistemology. The advantages of development in this direction are suggested by Hume himself. In its full development, virtue epistemology, supplemented by the reciprocity of the virtues, can provide us with a satisfactory foundation for the justification of beliefs based on testimony.

CHAPTER 6: REASONING ABOUT PROBABILITY

Reasoning with probabilities is a fundamental epistemological capacity. (Just how fundamental a capacity is sometimes in dispute.) But we all use probabilistic reasoning in our daily lives, as we try to figure out the chances of a coin flip, our horse coming in first in the fifth race, or a rainstorm. Though this capacity is a basic one, it is also problematic. In order to be successful in reasoning about probabilities we need to be able to break up the possible alternatives into reasonable classifications. To do this is to *partition* the space of possible alternatives. For example we might break up the weather alternatives into rainy, sunny, or cloudy and then reason about their respective probabilities. There are other, less intuitive ways of breaking up the alternatives. For example that it is either hotter than 95 degrees and it rains, or it is between 94 and 95 and it is cloudy, or it is neither of these. This partition is one that we would never think of, as it is highly counterintuitive. However it is a possible partition we could start our probabilistic reasoning with. In order to demonstrate the importance of partitioning, I will begin by giving some examples in which partitioning makes a big difference in our calculations of probability. These cases are ones in which it is not clear which partition we should favor. Yet choosing one partition over the other will give us very different results. I bring up these puzzling cases both to show how important partitioning and to demonstrate that partitioning is really at the base of all of our reasoning with probability. We simply do it so easily in normal cases that we hardly even notice. Hopefully these examples will show how crucial partitioning well is, and motivate us to understand how we can perform this crucial task so easily in normal cases.

We will then move on to examine different answers that might be given to the question of how we are capable of partitioning well. After considering the answers that might be given by a Bayesian and an act-based epistemologist, we will turn to the answer that virtue epistemology gives. By comparing these answers we will see that the virtue epistemological story best explains our everyday capacity in partitioning, by looking at partitioning as a skill that must be developed over time.

Puzzling Cases

The Sleeping Beauty Problem⁸⁸: Beauty is kidnapped by scientists who keep her in their lab and put her into the following scenario. Beauty will go to sleep on Sunday night. On Sunday night a fair coin will also be flipped. If the coin comes up heads Beauty will be awoken by the scientists on Monday morning and sent on her way. However, if the coin comes up tails Beauty will be awoken on Monday morning and then sent back to sleep with a drug that will erase all her memories of her Monday awakening. She will then be awoken on Tuesday (not knowing when she awakes that it is Tuesday) and sent on her way. Beauty knows all the details of this set-up and trusts the scientists to carry it out exactly as described. So Beauty goes to sleep on Sunday and later she finds herself being awoken by the scientists. She does not know whether it is Monday or Tuesday. What probability should she assign to the claim that the coin came up heads?

⁸⁸ First presented in Elga (2000).

There are two main positions in responses to this puzzle. The first is that Beauty ought to assign probability $\frac{1}{2}$ to the claim that the coin came up heads. A simple argument for this position is that, since the coin flipped was a fair coin, she knew going to sleep on Sunday that the probability that the coin would come up heads was $\frac{1}{2}$. Since Beauty has not learned any new information merely by going to sleep and waking up, she should continue to assign probability $\frac{1}{2}$. The second position is that Beauty ought to assign probability $\frac{1}{3}$ to the claim that the coin came up heads. A simple argument for this focuses on Beauty's uncertainty. She does not know if it is Monday and the coin came up heads, it is Monday and the coin came up Tails, or if it is Tuesday and the coin came up Tails. These three situations are indistinguishable for her and she has no reason to believe one over the others, and so she ought to apportion her probabilities equally between these three alternatives. Since only one is an alternative in which the coin came up heads, she ought to assign probability $\frac{1}{3}$ to the claim that the coin came up heads.

Can the theory of probability or the theory of probability updating help us in deciding between these two answers? Not really. Both answers are consistent with the theory of probability and both answers violate the Bayesian theory of probability updating.

Conditional probabilities are understood as the probability you assign to an outcome if you were to learn new evidence—it is the probability of the outcome conditioned on the evidence. Conditional probability is also mathematically defined. The Bayesian theory of probability updating requires that upon actually getting this new evidence you change your probability assignment of an outcome to your original conditional probability of that outcome on the evidence. This principle is violated by both answers to the Sleeping Beauty

problem. A version of this concern has been given above as the argument for the $\frac{1}{2}$ answer⁸⁹. We might object to the $\frac{1}{3}$ answer that it violates Bayesian probability updating, because Beauty changes her probability assignment with no new information⁹⁰. When she goes to sleep on Sunday she thinks that the probability of heads is half. Since she knows that in either scenario she will be awakened, the conditional probability she assigns to heads given that she wakes up should also be $\frac{1}{2}$. So according to Bayesian updating, on awaking, and learning she has just been awakened her new probability of heads should be equal to her old conditional probability of heads given that she woke up, and that should be $\frac{1}{2}$ rather than $\frac{1}{3}$.

However, it has also been argued that the $\frac{1}{2}$ answer to the Sleeping Beauty problem violates Bayesian probability updating in a different way.⁹¹ Consider what would happen if Beauty after waking were to be informed that it is Monday. She would then know that even if she were going to take the amnesia drug, she has not taken it yet. The two scenarios based on the flip of the coin are exactly the same up to this point. Indeed it wouldn't matter if the scientists flipped their coin on Sunday night or after Beauty awoke on Monday, since it is only then that they need to decide whether to give the amnesia drug or not. But if the coin flip hasn't happened yet, and Beauty knows it is a fair coin she should assign probability $\frac{1}{2}$ to heads. But for her to assign probability $\frac{1}{2}$ before she learns it is Monday and to continue to assign probability $\frac{1}{2}$ after she has learned that it is Monday is to violate Bayesian updating.

⁸⁹ See Davis Lewis's (2001). Justin Fisher and I have expanded on this argument and presented others in Fisher and Wright (ms) in which we argue for the $\frac{1}{2}$ answer based on pragmatic grounds and considerations of what we want the concept of probability to be able to do for us.

⁹⁰ Though Horgan (2004) argues contra this claim that Beauty does have new information when she wakes.

⁹¹ Adam Elga's (2000).

For based on the mathematical definition of conditional probability if her probability for heads upon waking is $\frac{1}{2}$ her conditional probability that the coin came up heads given that it is Monday ought to be $\frac{2}{3}$ ⁹². But this is a pretty counterintuitive probability to assign to a fair coin flip that will occur in the future. So unless she ended with a very counterintuitive probability assignment for heads after learning that it is Monday, Beauty will violate Bayesian updating if she assigns probability $\frac{1}{2}$ to the claim that the coin came up heads when she originally awakes.

Thus using only probability theory and theories of probability updating won't help us to get a definitive answer in cases like the Sleeping Beauty problem. Something more is needed to help us decide between the two possible probability assignments. It may be objected that this case is too artificial to present a serious worry to the use of probability and probability updating as a model of rational reasoning. While not conceding that we can reject the lesson of the Sleeping Beauty problem, I will move on to present another problem that is not at all artificial, and presents puzzling and very pressing problems for us today. This is the Doomsday Puzzle, and by examining what is problematic about it we will be able to see that the source of the tension in both problems is the same and reveals an underlying question in how we can reason using probabilities.

⁹² The conditional probability of A given B, $P(A|B)$, is defined as $P(A \& B) / P(B)$. If Beauty assigns probability $\frac{1}{2}$ to the claim that the coin came up heads, and divides her probability equally between the claims that the coin came up tails and it is Monday, and the coin came up tails and it is Tuesday, giving them probability $\frac{1}{4}$ each, then the probability that she assigns to the claim that it is Monday is $\frac{3}{4}$. Thus her conditional probability $P(\text{the coin came up heads} | \text{it is Monday}) = P(\text{the coin came up heads and it is Monday}) / P(\text{it is Monday}) = \frac{1}{2} / \frac{3}{4} = \frac{2}{3}$. By Bayesian updating, when she learns that it is Monday, the probability she assigns to the claim that the coin came up (or will come up) heads is the conditional probability she had before the probability of the coin coming up heads given that it is Monday. So if she follows Bayesian updating she will arrive at a probability of $\frac{2}{3}$ for the claim that the coin came up (or will come up) heads.

Doomsday Problem⁹³: There are many reasons to worry that human kind may die out relatively soon. Worries about global warming and nuclear proliferation are only a couple of these. However the Doomsday argument purports to show us that humankind will die out soon, without appeal to empirical evidence about these types of threats. Rather it only uses very limited empirical information, about population dynamics and your own position within that population curve. Consider two different alternatives Short and Long. Short is the hypothesis that the human race will only last a very short time longer, and (to have a convenient number) the 50 billionth human will be the last one. Long is the hypothesis that the human race will last a long time, though eventually die out (or evolve to something else) after the 50 trillionth human being. How can you decide between these alternatives? Well here is some information to help you. Look at your own case. You are a human with a specific place in the birth-order of all human beings. Though it is hard to say when humans really began being humans, let us say that you are within the first 50 billion human beings, say 45 billionth. You might wonder which hypothesis this fact supports. According to Bayesian reasoning, your observation (of yourself with your particular place in the birth-order) supports the hypothesis that makes this observation more likely. Leslie encourages us to think of ourselves as typical; “Now *whenever lacking evidence to the contrary one should prefer to think of one’s own position as fairly typical rather than highly untypical.*”⁹⁴ Thus you should favor an hypothesis that makes you a typical observer. But on the hypothesis Long, you are a quite untypical observer, appearing in the first 10th of a percent of human history. Thus your

⁹³ This argument is due to John Leslie (1990), who attributes its original formulation to B. Carter.

⁹⁴ John Leslie (1990) pp 65-72.

observation of yourself in your birth order gives you reason to support the hypothesis Short, which makes you a much more typical human observer.

To see just how this argument works lets consider a simpler case with the same structure.

Let us imagine that you check into hotel, that either has 2 rooms or 2000 rooms, you don't know which. To give yourself more information, you open your hotel door and check your room number, and lo and behold it is #2. (And you know that the rooms are, or course, numbered sequentially.) So long as you have no reason to think yourself untypical, say a guest of honor bound to be assigned a low-numbered room, your finding yourself in room #2 gives you evidence that the hotel has 2 rooms rather than 2,000. For if it had 2,000 rooms it is very unlikely that you would end up in room #2, but if it has only 2 then you are very likely to have ended up in room #2. In just the same way, so long as you are a typical human, your ending up in one the first 50 billion birth-order slots, gives you evidence that the total number of humans is that hypothesized by Short rather than that hypothesized by Long.

One crucial assumption of the Doomsday argument is that you are a typical human observer.

In many of the responses the to Doomsday argument, this assumption is questioned by insisting that observer has other evidence that he is a typical in his rank placement.⁹⁵ This type of objection I think misses a deeper point on which the assumption that one is a typical human observer is not in and of itself a problematic assumption, but rather reveals a focus on the role of the observers and their birth-order, over other kinds of factors. We might

equally insist that, whenever we are lacking evidence to the contrary one should prefer to think of *any given year* as fairly typical rather than any observer. Focusing on years rather than on the birth-order of observers, will give us a very different conclusion. To have convenient numbers let us assume that the humans have been around for close to 1 million years. The hypothesis Short, will then be the hypothesis that the human race will last 1 million years, while the hypothesis Long will be the hypothesis that humans will last 2 million years. Taking only the information that we this year is a year very close to the end of the first 1 million years of human existence, we now have evidence that the human race will last a Long, rather than a Short time. For on the hypothesis Short this is a very atypical year, one that is incredibly close to the end of human existence. On the hypothesis Long, this is a far more typical year. Thus focusing on years rather than on birth-order of observers gives us an Anti-Doomsday argument.

The Partitioning Problem

How can the Doomsday argument and the Anti-Doomsday arguments arise out of basically the same set of information? This extreme difference shows us how sensitive our reasoning about probabilities is to the way in which we partition the alternatives we are considering. If we partition by years, then we get a very different answer than if we partition by individual observers. One of the things that makes the original doomsday argument work is the fact that while we are well along in the history of human kind, we are also very early in the birth order. This difference is possible because of the geometrical expansion of the human population. To see this graphically represented examine figure 1 below. Partitioning

⁹⁵ See Eckhart (1993), Deiks (1992) and Bartha and Hitchcock (1999)

according to human years means to focus on the x-axis, and to count every year along it equally. Partitioning according to birth order means to focus on the total area under the curve, giving each year more or less weight according to its population. In this diagram this means to count each of the circles.

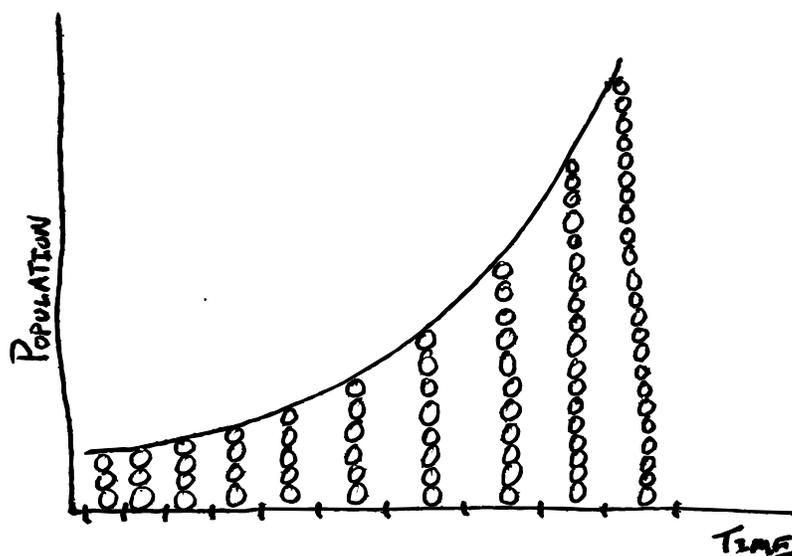


FIGURE 1, Alternate Partitions for the Doomsday Problem

Choosing a different partition will lead us to give very different answers in the probability assigned to different outcomes. This is because in reasoning in cases like this we make use of the principle of indifference. This principle is a more general statement of the idea that, without further evidence, every observer should be thought of as typical. In general that principle of indifference looks at a partition, and say, with no evidence that any member of this partition is special, we should assign them all equal probability. We use this principle whenever we reason from equal numbers to equal probability. If I tell you that this urn contains 10 red ball and 30 black balls, and ask you the probability that I will draw a red ball,

you use the principle of indifference when you assign an equal probability to my drawing each ball, and put the probability of drawing a red ball at .25. This is plausible enough when we think that the balls do not have any special characteristics. However when I tell you that the red balls are all coated with a slippery substance that makes them hard to grab, you are justified in moving away from the probability assignment given by the principle of indifference. Still, when we have no other special information we are justified in using a principle of indifference to assign equal probabilities to each element of a partition.

Uses of the principle of indifference are problematic when there are several competing partitions. This occurs in the Doomsday and Anti-Doomsday arguments, and it all so occurs in the arguments about Sleeping Beauty. Figure 2 below represents how the Sleeping Beauty problem might be differently partitioned. First, let us focus on the situations that Beauty might find herself in. These are represented by circles in the diagram. When she awakes, there are three different situations that Beauty might be in: Monday and Heads, Monday and Tails, or Tuesday and Tails. Beauty can't tell these circumstances apart, and has no special reason to think that she is in any one of these situations over another, and so using the principle of indifference, she assigns an equal probability $1/3$ to each one. This partition results in the answer that the probability of heads is $1/3$ ⁹⁶.

⁹⁶ To partition this way is to partition by centered worlds. To partition in the second way is to partition by possible worlds. For a discussion of this see Elga (2000) and Lewis (2001). It is interesting to note that the partitions in the Doomsday problem can also be characterized as being either on centered worlds (observers) or on worlds (years in which there are humans).

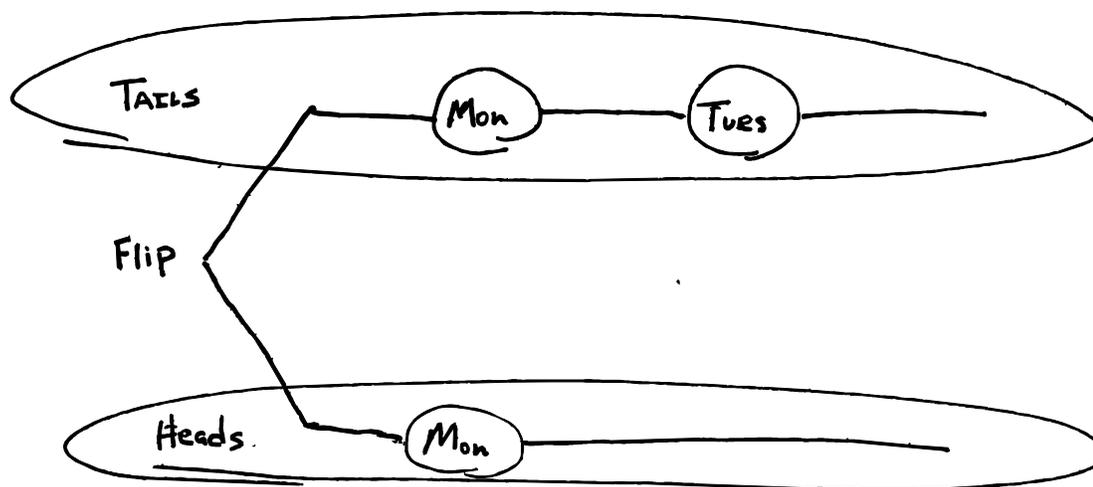


FIGURE 2, Alternate Partitions for the Sleeping Beauty Problem

However, there is an alternative partition, one that focuses not on where Beauty is, but rather on the overall ways that the world might be. These are represented by the ovals in figure 2. With this focus there are really only two ways the world might be: either heads came up or tails did. This partition yields two alternatives, and the principle of indifference assigns them equal probability of $\frac{1}{2}$. Of course, in this case we also have some special information, that is coin is a fair one, but this only serves to reinforce the probability assignment of $\frac{1}{2}$ to the claim that the coin came up heads. Thus this second partition results in the alternative answer that the probability of heads is $\frac{1}{2}$.

So in reasoning both about the Sleeping Beauty Problem and about the Doomsday Problem, we can see that partitioning plays an important role and can affect the probabilities that we

eventually assign to the outcomes⁹⁷. This presents a general problem for any epistemological theory. While these puzzling cases leave us with no clear cut answers on how to partition, normal cases are not like this. In fact in most ordinary cases the correct partition is so easy for us to see that it hardly even bears thinking about. The Partitioning Problem is to explain how it is that we seem to be able to come up with good partitions so easily in normal cases. Although in some odd cases we disagree about the appropriate partitions, in most cases we partition in uniform ways that are in agreement with others. How we can do this presents a problem for epistemological theories in general, for although making these partitions is easy and intuitive for us, it is difficult to see theoretically how we could take in all the information needed to make these partitions.

Bayesian Response to the Partitioning Problem

One way around the Partitioning problem is to deny that we use partitioning in many cases of reasoning about probability. This the option endorsed by Bayesians⁹⁸. Notice that in the second partition we considered in the Sleeping Beauty problem, the uses of a partition and the principle of indifference were not really needed, for we had special information that the coin being tossed by the scientists was a fair coin. This means that we do not need to partition and use the principle of indifference in this case. In fact, in any case where we already have a probability assignment, either given explicitly, or simply the probability assignment we happen to have, we do not have to use the partitioning method. Bayesians

⁹⁷ Dave Schmitz and I have given an argument for a similar importance of our initial decisions before probability calculations in the Newcomb problem. See Schmitz and Wright (2003)

⁹⁸ As an example see Mark Kaplan (1996). He suggests that Bayesianism can be taken as an independent view in epistemology, so that Bayesian updating might be the whole of epistemology.

can agree that partitioning and using indifference is a messy method, but they continue by arguing that this mess will come out in the wash as we continually update our probabilities based on new information. Even if we start with a really bad probability assignment, this bad beginning will be eventually be wiped out by new information and probability updating. This claim of Bayesians is backed up by the convergence thesis, which shows that probability assignments that are updated by Bayesian updating will eventually converge, no matter how disparate the original probability assignment⁹⁹. This is a very striking result, and seems to undermine our worry about the Partitioning Problem. If the probabilities that are determined by the partition that we choose don't matter in the end, it doesn't seem to matter what partition we originally choose. How we originally partition does not seem to matter in the end.

While the irrelevance of our partitions is true in theory, it is not true in practice. For there is an important constraint on applicability of the convergence theorem, and that is that in order to guarantee convergence, we need to allow for unlimited iterations of Bayesian updating on the probability functions. Though this is fine in theory, it is obviously not applicable in practice. Simply given the finite nature of our lives, no individual can have an unlimited number of iterations of Bayesian updating on their probability function. Though mortality provides an upper bound on the number of iterations, the problem for the Bayesian is worse than this. For it does not seem that individuals normally have a probability function that they retain across different instances of Bayesian updating. Individuals rarely have probability assignments that they bring with them to a probability problem. Rather to

⁹⁹ See de Savage(1954) p46 ff.

calculate a probability we normally begin with a partition and an application of the principle of indifference¹⁰⁰.

Even if we sometimes have a probability measure that we retain through different instances of updating, it is quite implausible that we have such a probability measure for all the claims that we might assign a probability to. Sometimes we move into an entirely new area, and learn about something we have never studied, or assigned a probability to before. For example if we were to take up pottery as a hobby I would begin with no idea of the probability to assign to the claim that a particular type of glaze will crack when fired to a given heat. Thus there are many cases, when we move into topics previously foreign to us, when we need to use the tools of partitioning and applying the principle of indifference. And unless we have a privileged way to make these partitions, the Bayesian will have a hard time explaining how we can successfully learn about a new topic quickly. Starting with some random partition we might be quite bad at learning to assign probabilities in a new field. Our success in new fields needs to be explained by our starting our studies of these fields with a good partition.

Belief-Based Theories and the Partitioning Problem

Most epistemological theories follow an belief-based model in which the primary question to be answered is, “Is this particular belief justified?”. To answer this question belief-based

¹⁰⁰ This approach is easy to see within articles about the proper probability assignments in puzzling cases like that of Sleeping Beauty and Doomsday. Authors of such articles frequently begin by presenting a partition, then giving reasons to think that we do not have special reason to prefer any element of this partition, and finally applying the principle of indifference to the case.

theories look for some general feature of the belief, and the circumstances, that makes it the case that the belief is justified. Belief-based epistemological theories are then looking for a general rule by which we can decide, for any given instance of belief, whether that belief is justified or not. When we apply this approach to the particular instance of partitioning, we find that belief-based theories need to give a general rule that differentiates the bad partitions for the good partitions. Only the probability assignments that are based on good partitions will then count as justified beliefs.

Looking for a rule that picks out the right partitions to which we can apply the principle of indifference we might look to the parallels between looking for a rule to pick out good partitions, and looking for a rule to help us pick out good predicates. This is Goodman's new riddle of induction¹⁰¹. The cases that Goodman gives point out that not all predicates or classifications are created equal. Some predicates like *grue* and *bleen* clearly give us an inferior classification scheme on which to base our inductions. This is because the predicates in question are not projectible. Although there are many obvious cases of predicates that are projectible, and equally obvious examples of predicates that are not projectible, the heart of the problem is coming up with a general principle by which we can tell projectible from non-projectible predicates in general. But this is not a simple task. Indeed it has spawned a cottage industry of different attempts at codification of projectibility, followed by objections and further refinements of these codifications. So far though there is no general agreement about the solution to Goodman's new riddle of

¹⁰¹ See Goodman (1955)

induction. Indeed there is not even agreement about the type of answer that should be given to this problem.

The difficulty of giving a general solution to the new riddle of induction gives us two reasons to doubt that the answers we have now will be the type of answers that epistemologists need in order to fit into their overall theories, and tell us which probability assignments are justified and which are not. The first reason for worry is simply that there is not much agreement in the field. Given this, it seems unlikely that any given purported solution to the new riddle of induction is the correct answer. There are so many competitors we may reasonably doubt that the particular answer we might choose to endorse is the correct one. The second, and deeper worry is spawned by the complexity of the theories given to capture projectibility. In response to many iterations of objections, these theories have become more and more complex to avoid the objections that have previously been raised. As these theories of projectibility become more complex, it is less and less plausible that they capture a rule that the agent is really following, either explicitly or implicitly, when that agent is making her partitions. Part of what an epistemologist is looking for is not just a rule that differentiates the good partitions from the bad partitions, but a rule that can also be used by the agent in her making those partitions. Otherwise the agent may make a good partition, but do so on the basis of a bad rule, or no rule at all. In these cases, even if the resultant probability judgment is true, it is not justified. Thus the more complex we think the rules of projectibility are, the less likely epistemic agents will be using them in their partitioning, and thus reaching justified beliefs about probability.

The problem with belief-based attempts to solve the partitioning problem by an appeal to a general rule about projectibility is that such rules are apt to be too complex for the agent to use in her reasoning. Therefore we should look for another way to solve this problem, one that moves away from general rules, since it is the complexity in these general rules that seems to cause problems.

Virtue Theories and the Partitioning Problem

Worries about the complexity of general rules are not just found in epistemology. They are also very common in ethics. Act-based theories of ethics have repeatedly been updated and emended in order to adequately reflect our intuitions about what actions are right actions. Still it is a very complex matter to come up with a rule that can be implemented to tell one, for example, when lying is acceptable and when it is unacceptable. This complexity makes it difficult to see how we could use rules in our moral reasoning. Either they are too simple and give us bad advice in many cases, or they are too complex and become unwieldy to use in our moral deliberation. Virtue theories respond to this worry by insisting that we do not use general rules in our moral reasoning. General rules are not the right sort of thing to guide out actions. Rather than using rules, we instead develop skills, and then in particular instances use those skills to help us decide what to do¹⁰². Explaining our capacities in moral reasoning, or in epistemological reasoning, by reference to skills allows us to avoid worries about rules. In fact the strongest versions of virtue theory claim not only that there are no

¹⁰² Although many virtue theorists think that virtues are not skills because they require something over and above what is required to develop a skill, virtue theorists in general agree that the way that we implement virtues requires at least as much as is required when we implement skills, even if more is required for virtues. For a discussion of the relationship between virtues and skills see Annas (1995).

rules of the sort that we could use in moral reasoning, but that in general moral virtues are simply not codifiable¹⁰³. For the argument at hand we need not go so far. We need only to hold that virtues are not codifiable in the sense that they could be given or learned through a set of rules.

Non-codifiability by itself does little to help us with our original partitioning problem. For we want to know how it is possible to learn to partition well and being told how we cannot learn it does little to answer this question. It is here that the skill analogy is helpful. We might well ask, how on earth does a carpenter ever learn how to make good cabinets? The procedure for making good cabinets is not codifiable in the weaker sense—that is it cannot be learned through a set of explicit rules. There are simply too many variables, too many types of wood, too many possible knots and whorls, too many variations in tools, and so on. Yet a carpenter can learn to make cabinets, despite the lack of rules to help him. The carpenter can learn this because he is learning a skill. Learning a skill does not require rules, though rules of thumb might be helpful. Rather learning carpentry requires apprenticeship. To learn carpentry one must first find someone who is already an experienced carpenter and then learn from that person. This form of learning will consist in observing the actions and choices of the master carpenter, and then trying on one's own to do the same things that he does. This is not mere mimicry though, for the apprentice carpenter who has truly learned the skill will be able to deal with novel situations, and novel pieces of wood, that he has never seen before.

¹⁰³ See the discussion of levels of codifiability in Hursthouse (1999).

The same type of learning then will take place as we develop the skill of being a good epistemic agent. This is true in general, as well as in the specific case of partitioning. Learning to partition well cannot be learned through learning a set of general rules. Rather in order to partition well we need to take others as our models, and watch the way that they partition the world. This does not mean blindly accepting the partitions of others, but rather picking out those we think do well, following their examples, but finally deciding for ourselves how to partition. This is similar to the carpenter who may develop a different style or technique from that of his mentor. The simplest example of this type of learning applied to partitioning, and hence to projectible predicates generally is the learning of categories by infants as embedded in our language. One of the reasons that Goodman's examples are so obviously wrong is that they undercut one of the classification systems that we learn earliest, that of color. As infants we learn these classifications from other adults, and over time we become very adept at applying them. However our adult application of the concepts we learn as a child is not automatic. For example, someone who grows up in a racist household may unquestioningly pick up certain categories as a child. However as they mature they develop the capacity to question these categories and to change them. Thus the categories and associated predicates that a developed adult uses are her own in the sense that she has chosen to retain them or to change them. This complicated process by which she arrives at the predicates that she thinks are projectible, and which she will use for her partitions, cannot be codified. Nor can her resultant skill be codified. To teach it to someone else she would need to "apprentice" them, rather than give them a set of rules to follow.

Thus the fact that a virtue has the same structure, both in development and in the final results, as a skill can explain how our capacity to partition well can be both learnable and non-codifiable. This capacity, which is an acquired skill, can explain how we are capable of using probabilistic reasoning. For before we begin reasoning about probabilities, we need to begin with an appropriate partition on which to apply the principle of indifference. Thus virtue epistemology can provide us with a way of explaining the capacity of partitioning that underlies our ability to reason about probabilities.

WORKS CITED

- Alston, W.P.(1993) "Epistemic Desiderata", *Philosophy and Phenomenological Research*. 53(3): 527-551
- Annas, Julia (1993) *The Morality of Happiness*, New York: Oxford University Press.
- (1995) "Virtue as a Skill", *International Journal of Philosophical Studies* 3(2): 227-243.
- (2002) "My Station and Its Duties: Ideals and the Social Embeddedness of Virtue", *Proceedings of the Aristotelian Society* 102: 109-123.
- (2003) "The Structure of Virtue" in *Intellectual Virtue: Perspectives from Ethics and Epistemology*, ed. Michael DePaul (ed), Oxford: Clarendon Press, 15-33.
- (ms) "Virtue Ethics and Social Psychology".
- Anscombe, G.E.M. (1958) "Modern Moral Philosophy", *Philosophy* 33: 1-19.
- Asch, S. E. (1956) "Studies of Independence and Conformity: A Minority of One against a Unanimous Majority". *Psychological Monographs*, 70 (Whole no. 416).
- Bartha, Paul and Hitchcock, Christopher (1999) "No One Knows the Date or the Hour: An Unorthodox Application of Rev. Bayes's Theorem", *Philosophy of Science* 66: S339-S353.
- Bonjour, Laurence (1985) *The Structure of Empirical Knowledge*, Cambridge MA: Harvard University Press.
- Chisholm, Roderick (1957) *Perceiving: a Philosophical Study*, Ithica: Cornell University Press.
- (1977) *Theory of Knowledge* second edition, Englewood Cliffs NJ: Prentice-Hall.
- Coady, C.A.J. (1992) *Testimony: A Philosophical Study*, New York: Clarendon Press.
- Cohen, Stewart (1988) "How to be a Fallibilist", *Philosophical Perspectives 2: Epistemology*, ed. James E Tomberlin Atascadero: Ridgeview Press, 91-123.
- (1998) "Contextualist Solutions to Epistemological Problems: Scepticism, Gettier, and the Lottery", *Australasian Journal of Philosophy* 76(2): 289-306.
- Darley, John and Batson, C. Daniel (1973) "From Jerusalem to Jerico": A study of Situational and Dispositional Variables in Helping Behavior, *Journal of Personality and Social Psychology* 27:1, pp 100-108.

- DeRose, Keith (1995) "Solving the Skeptical Problem", *Philosophical Review* 104(1): 1-52.
- (1996) "Relevant Alternatives and the Content of Knowledge Attributions", *Philosophy and Phenomenological Research* 56(1): 193-197.
- (2002) "Assertion, Knowledge, and Context", *Philosophical Review* 111(2): 167-203.
- Deiks, Dennis (1992) "Doomsday—Or: the Dangers of Statistics", *Philosophical Quarterly* 42(166): 78-84.
- Doris, John (1998) "Persons, Situations, and Virtue Ethics" *Noûs*, 32, pp. 504-530.
- Eckhardt, William (1993) "Probability Theory and the Doomsday Argument" *Mind* 102(407): 483-488.
- Elga, Adam (2000) "Self-Locating Belief and the Sleeping Beauty Problem", *Analysis* 60(2): 143-147.
- Fisher, Justin and Wright, Sarah (ms) "Sleeping Beauty and Two Kinds of Probability"
- Foley, Richard (1993) *Working Without a Net: A Study of Egocentric Epistemology*, New York, Oxford University Press.
- Foot, Philippa (1978) *Virtues and Vices*, Oxford: Blackwell.
- Fricker, Elizabeth (1995) "Telling and Trusting: Reductionism and Anti-Reductionism in the Epistemology of Testimony", *Mind*, 104: 393-411.
- Gettier, E. L. (1963) "Is Justified True Belief Knowledge?", *Analysis* 23: 121-3.
- Goldman, Alvin (1979) "What is Justified Belief?", in *Justification and Knowledge ed. George Pappas*, Dordrecht: Reidel
- (1996) "Epistemic Folkways and Scientific Epistemology", in *Empirical Knowledge: Readings In Contemporary Epistemology*, Ed Paul Moser, Totowa: Rowman-Littlefield.
- (1999) "Internalism Exposed", *Journal of Philosophy* 96(6): 271-293.
- Goodman, Nelson (1955) *Fact, Fiction, and Forecast*, Cambridge: Harvard University Press.
- Grice, H. P. (1989) *Studies in the Ways of Words*, Cambridge, Mass.: Harvard University Press.
- Harman, Gilbert (1999) "Moral Psychology Meets Social Psychology: Virtue Ethics and the Fundamental Attribution Error", *Proceedings of the Aristotelian Society*, 99, pp. 315-331.

- (2000) "The Non-Existence of Character Traits", *Proceedings of the Aristotelian Society*, 100, pp. 223-226.
- Hartshorne, H. and May, M.A. (1928) *Studies in the Nature of Character*, vol. 1 *Studies in Deceit*, New York: Macmillan.
- Henderson, David and Horgan, Terence (2000) "Iceberg Epistemology", *Philosophy and Phenomenological Research* 61(3): 497-535
- (forthcoming) *A Priori Naturalized Epistemology: at the Interface of Cognitive Science and Conceptual Analysis*.
- Horgan, Terry (2004) "Sleeping Beauty Awakened: New Odds at the Dawn of the New Day", *Analysis* 64 (281): 10-21.
- Hume, David (1975) *Enquiries Concerning Human Understanding and Concerning the Principles of Morals*, With analytical index by L. A. Selby-Bigge and text revised and notes by P. H. Nidditch, Oxford: Oxford University Press.
- (2000) *A Treatise of Human Nature*, edited by Norton and Norton, New York: Oxford University Press.
- Hursthouse, Rosalind (1999) *On Virtue Ethics*, New York: Oxford University Press.
- Kamtekar, Rachana (forthcoming) "Situationism and Virtue Ethics on the Content of our Character", *Ethics*.
- Kaplan, Mark (1991) "Epistemology on Holiday", *Journal of Philosophy* 88(3): 132-154.
- (1996) *Decision Theory as Philosophy*, New York: Cambridge University Press.
- Lehrer, Keith (2000) *Theory of Knowledge* second edition, Boulder: Westview Press.
- (2001) "The Virtue of Knowledge", in *Virtue Epistemology: Essays on Epistemic Virtues and Responsibility*, eds. Abrol Fairweather and Linda Zagzebski, New York: Oxford University Press.
- Lehrer, Keith and Cohen, Stewart (1985) "Justification Truth and Coherence" *Synthese* 55: 191-208.
- Leslie, John (1990) "Is the End of the World Nigh?", *Philosophical Quarterly* 40(158): 65-72.
- Lewis, David (1996) "Elusive Knowledge", *Australasian Journal of Philosophy* 74(4): 549-567. Reprinted in *Papers in Metaphysics and Epistemology*, Cambridge: Cambridge University Press.

- (2001) "Sleeping Beauty: Reply to Elga", *Analysis* 61(3): 171-176.
- Murdoch, Iris (1967) "The Sovereignty of Good over Other Concepts", London: Cambridge University Press.
- Nozick, Robert (1974) *Anarchy, State, and Utopia*, New York: Basic Books.
- Pollock, John (1990) *Nomic Probability and the Foundations of Induction*, New York: Oxford University Press.
- (2001) "Evaluative Cognition", *Nous* 35(3): 325-364.
- Pollock, John and Cruz, Joseph (1999) *Contemporary Theories of Knowledge* second edition, New York: Rowman & Littlefield.
- Pritchard, Duncan (2002) "Two Forms of Epistemological Contextualism", *Grazer Philosophische Studien* 64: 19-55
- Rawls, John (1971) *A Theory of Justice*, Cambridge MA: Belknap Press.
- Ross, Lee and Nisbett, Richard E. (1991) *The Person and the Situation: Perspectives of Social Psychology*, New York: McGraw-Hill.
- Savage, Leonard (1954) *The Foundations of Statistics*, New York: Wiley.
- Schmidtz, David and Wright, Sarah (2003) "What Nozick Did for Decision Theory", *Midwest Studies in Philosophy*, 27(1): 282-294.
- Sirigu, Angela et al. (2004) "The Involvement of the Orbitofrontal Cortex in the Experience of Regret", *Science*, Vol 304(5674), pp. 1167-1170.
- Sreenivasan, Gopal (2002) "Errors about Errors: Virtue Theory and Trait Attribution", *Mind*, 111, pp. 47-68.
- Sosa, Ernest (1980) "The Raft and the Pyramid: Coherence versus Foundations in the Theory of Knowledge", *Midwest Studies in Philosophy* 5: 3-26. Reprinted in *Knowledge in Perspective*, New York: Cambridge University Press.165-191.
- (1991) "Reliabilism and Intellectual Virtue" in *Knowledge in Perspective*, New York: Cambridge University Press.
- Unger, Peter (1971) "A Defense of Skepticism", *Philosophical Review* 80: 198-219.

Williams, Bernard (1985) *Ethics and the Limits of Philosophy*. Cambridge MA: Harvard University Press.

Williams, Michael (1992) *Unnatural Doubts: Epistemological Realism and the Basis of Scepticism*, Cambridge: Blackwell.

Zagzebski (1996) *Virtues of the Mind* Cambridge, UK: Cambridge University Press

----- (1999) "What is Knowledge" in *The Blackwell Guide to Epistemology* eds. Greco and Sosa, Malden, MA: Blackwell.