

# REASON DETHRONED; KNOWLEDGE REGAINED

by

James Arthur Moore

B.A., Indiana University (Bloomington), 1983

M.A., University of Pittsburgh, 1987

Submitted to the Graduate Faculty of  
Arts and Sciences in partial fulfillment  
of the requirements for the degree of  
Doctor of Philosophy

University of Pittsburgh

1991

© Copyright by James Arthur Moore  
1991

## **REASON DETHRONED; KNOWLEDGE REGAINED**

James Arthur Moore, Ph.D.

University of Pittsburgh, 1991

Hume held that we have no rational justification for our inductive beliefs. A more radical view is that we have no rational justification for any of our beliefs. This dissertation has two goals pertaining to this more radical view.

The first goal is to find a basis for constructive epistemology that is consistent with this view. This goal is first sought by considering externalist theories of knowledge (those that make justification a function of aspects external to awareness) since these do not require rational justification for knowledge. Externalist theories are defended against the usual objections, partly via a strategy of immunizing them from counterexample-based objections by arguing that epistemologies can be successful even if they fail to explicate ordinary epistemic notions. But a new objection to externalist theories is then brought to light. The objection begins as an attack against a dogma of contemporary epistemology, that the chief benefit of possessing knowledge is having a true belief. It is argued that there are many other benefits to having knowledge (e.g. that knowledge is a belief with which it is easy to "infect" others),

and that externalist theories are defective because externalist knowledge lacks these benefits. A mixed internalist/externalist theory, bilevel reliabilism, is then presented as a solution to this difficulty.

The second goal is to provide an explanation of the function and origin of human epistemic practices (e.g. the linguistic practice of knowledge attribution) that is consistent with the no-rational-justification view. Providing such an explanation is problematic for holders of this view because, if it is correct, it seems, *prima facie*, that there is no reason to have epistemic practices. This goal is achieved by arguing that epistemic practices, despite appearing (because of their misleading folk epistemology) to have rational justification as their goal, chiefly function to promote the existence of bilevel reliabilist knowledge, a very useful type of belief that is not rationally justified. It is then argued that the explanation of the origin of epistemic practices is that they arise from natural human inductive tendencies.

## FOREWORD

There are many people whose assistance in preparing this dissertation I would like to acknowledge. First, I would like to thank my dissertation director, Joseph Camp, for his many helpful suggestions and criticisms and for teaching me the difference between hard philosophizing and mere logic-chopping. I would also like to thank another member of my committee, John Earman, for his helpful comments and for alleviating some of the difficulties created by the untimely death of another committee member, Wilfrid Sellars. I would also like to thank Annette Baier and Alexander Nehamas for helping me to see, via Hume and Nietzsche, respectively, that life with "reason dethroned" was not as bleak as it at first appeared. And I would like to give a special thanks to my dear friend, Christopher Williams, for helping me to understand the limitations of analytic philosophy and for never letting me forget that philosophy is one of the humanities. Finally, I would like to thank my parents, Edward and Eileen, and brother, Michael, for their financial and other assistance that helped to assure that I got this done before the millennium was out.

# TABLE OF CONTENTS

<b>FOREWORD</b> .....	v
<b>PROLEGOMENON</b> .....	1
<b>1. OPTIONS FOR A CONSTRUCTIVE NEO-HUMEAN EPISTEMOLOGY</b> .....	4
<b>2. OBJECTIONS TO EXTERNALISM</b> .....	11
2.1 <i>Counterexample-Based Objections to Externalist Theories</i> .....	11
2.2 <i>Externalism, Awareness, and Norms</i> .....	19
2.3 <i>The Original Causation Problem</i> .....	20
<b>3. THE REFERENCE CLASS PROBLEM FOR PROCESS RELIABILISM</b> .....	22
<b>4. EXTERNALISM, OR, HOW TO GET NO MORE FROM KNOWLEDGE     THAN TRUTH</b> .....	32
4.1 <i>Vericentrism in Epistemology</i> .....	33
4.2 <i>The Extra-verific Utility of Knowledge</i> .....	35
4.2.1 <i>The Extra-verific Utility of T-knowledge</i> .....	36
4.2.2 <i>Externalism and Extra-verific Utility</i> .....	45
4.3 <i>Why R-knowledge and T-knowledge Differ in Extra-verific Utility</i> .....	48
4.4 <i>The Flaw in Externalism</i> .....	53
<b>5. BILEVEL RELIABILISM: AN ACCOUNT OF KNOWLEDGE AND     JUSTIFICATION</b> .....	56
5.1 <i>Mere Internalism Reconsidered</i> .....	56
5.2 <i>Mixed Internalist/Externalist Accounts</i> .....	59
5.3 <i>Bilevel Reliabilism, a Radical Mixed View</i> .....	62
5.4 <i>The Extra-verific Utility of Br-knowledge</i> .....	67
5.5 <i>Bilevel Reliabilism and Reasons-based Justification</i> .....	72

5.6 Bilevel Reliabilism and Alston's "Internalist Externalism" .....	73
5.6.1 Alston's Account .....	74
5.6.2 Alston's Account and Bilevel Reliabilism .....	78
5.7 An Externalist Objection.....	80
<b>6. EXPLAINING EPISTEMIC PRACTICES: PRELIMINARIES .....</b>	<b>83</b>
6.1 A Task for Bilevel Reliabilism.....	83
6.2 The Nature and Importance of the Project of Explaining Epistemic Practices .....	85
6.3 A Strategy for a Bilevel Reliabilist Explanation of Epistemic Practices .....	90
6.4 Human Folk Epistemology.....	96
6.5 The Content of Ordinary Epistemic Practices .....	106
<b>7. A BILEVEL RELIABILIST EXPLANATION OF EPISTEMIC PRACTICES 109</b>	
7.1 The Origin of the Practice of Knowledge Attribution.....	110
7.2 How Knowledge Ascriptions Could be of Br-Knowledge.....	113
7.2.1 Ascribing Knowledge to R-justified Beliefs .....	114
7.2.2 Ascribing Knowledge to Br-justified Beliefs.....	123
7.3 That the Practice of Knowledge Attribution Promotes Br-knowledge and that such Promotion is Useful .....	127
7.4 Explaining the Practice of Knowledge Attribution as it Exists Today.....	129
7.5 Summary of the Explanation and Consideration of the Generality of its Application.....	132
7.6 Implications of the Bilevel Reliabilist Explanation of the Existence and Function of Epistemic Practices .....	135
<b>BIBLIOGRAPHY.....</b>	<b>143</b>

## PROLEGOMENON

Hume held that reason alone could not be responsible for and could not justify our adoption of inductive beliefs. A more extreme view is not uncommon today: that reason alone does not justify the adoption of any of our beliefs.<sup>1</sup> In recent times the former view has often gone by the name of "inductive skepticism" and the latter by the name of "global skepticism". However, for the purposes of this dissertation, I want to introduce a new term, neo-Humeanism, for the latter view. There are several reasons for this move. First, there are a number of different reasons for being a global skeptic. Giving this view its own name differentiates it as the view involving a certain claim about the impotency of reason. Second, the view itself does not make any claim about the impossibility of knowledge, so the new name helps to differentiate it from full-blooded skepticism. Third, since the new name avoids the negativity implicit in "skepticism", it does not misleadingly paint out the positive aspects of the view,<sup>2</sup> e.g. that neo-Humeanism naturally leads to positive philosophizing in the form

---

<sup>1</sup> I would number among adherents of this view, or something very close to it, philosophers as diverse as Friedrich Nietzsche, Paul Feyerabend, Richard Rorty, the Scottish sociologists of knowledge, and the French deconstructionists.

<sup>2</sup> As Kemp Smith (1949) complained had happened with Hume's philosophy.



of determining what non-rational means we have for adopting and justifying our beliefs.

A fundamental assumption of this dissertation is the truth of neo-Humeanism. This assumption is made not, of course, because neo-Humeanism is uncontroversial but because the dissertation aims to explore some of the consequences of neo-Humeanism and there is not room in one dissertation to both explore its consequences and defend it.<sup>3</sup> One consequence that provides part of the motivation for undertaking the dissertation concerns epistemic practices--the practices consisting of attributing knowledge, assessing evidence, giving reasons, etc. These practices seem, for all the world, to have the pursuit of rational justification as their raison d'être, but, if neo-Humeanism is correct, this supposed raison d'être is a sham. The question then arises of just why these practices exist and what their purpose, if any, might be. Finding a satisfactory answer to this question is a primary purpose of the dissertation, and the reader would do well to bear this in mind even though the discussion often seems to veer far from this topic.

A second assumption should also be noted. This assumption is the viability of some sort of substantive theory of truth. I do not believe anything in the dissertation requires a robust realist and/or correspondence theory of truth (nor would I wish to

---

<sup>3</sup> Anyone interested in examining a defense of the view can consult a number of works, e.g. Rorty's Philosophy and the Mirror of Nature can be viewed in this light or, better yet, Hume's Treatise of Human Nature for the portion of neo-Humeanism pertaining to inductive and moral beliefs.

advocate one), but it does require a theory that can at least make some sense of the notion that cognitive<sup>4</sup> devices succeed, qua cognitive devices, by arriving at the truth.

---

<sup>4</sup> It has been pointed out to me that in contemporary psychological parlance "cognitive" has come to mean something like "rule-guided". In this dissertation, "cognitive" has the more traditional meaning, stemming from its Latin root, of "pertaining to or associated with knowing and believing".

# *1*

## **OPTIONS FOR A CONSTRUCTIVE NEO-HUMEAN EPISTEMOLOGY**

The epistemology most commonly associated with a Humean or neo-Humean position is the negative one of skepticism. A principal aim of this dissertation is to discover some basis for constructive epistemology consistent with neo-Humeanism. And this is not an easy aim to fulfill, for, as will be seen presently, most traditional approaches to epistemology are inconsistent with neo-Humeanism. It should also be noted that this aim has been chosen not because of an interest in epistemology for epistemology's sake but because human epistemic practices are an important and interesting element of human affairs, and it would be useful to have some viable discipline devoted to a positive examination of them. Let us consider, then, the suitability of the various types of epistemological theories as bases for such a discipline by examining the following taxonomy of epistemological theories. The taxonomy is not exhaustive, but it does include the options most important for present purposes. Since the taxonomy relies heavily on the internalism/externalism distinction and the notion of "rational privileging" to distinguish among sorts of theories, it will be useful to discuss these ideas before proceeding to the taxonomy itself.

Though it is widely used in the contemporary epistemological literature, the internalism/externalism distinction has achieved no fixed usage. It was first intro-

duced into the philosophical mainstream in Armstrong (1973) (though he only uses half of the contemporary terminology, "externalism", and credits its origination to an unpublished paper by Gregory O'Hair) and he explains the externalist half of it as follows:

According to 'Externalist' accounts of...knowledge, what makes a true...belief a case of knowledge is some natural relation which holds between the belief-state, Bap [i.e. *a*'s believing that *p*], and the situation which makes the belief true. (1973, p.157)

Contemporary versions of the distinction vary considerably from this original account and are applied to theories of justification as well as theories of knowledge. The version of the distinction to be employed here derives--but differs--from that in Pollock (1986). On the version used here, epistemological internalism is the view that justification is a function solely of the internal states of a cognizer<sup>5</sup>--where by "internal state" is not meant a state within the confines of the cognizer's body but a state of which he can be aware, e.g. a belief or perceptual state. Thus, an internalist account of justification will specify just what sorts and sets of internal states will justify a cognizer in a belief. An internalist account of knowledge will be defined as

---

<sup>5</sup> It might seem that some accounts that are usually thought of as internalist, e.g. Cartesian ones, clearly require more for justification than that a cognizer merely possess certain internal states. For example, a Cartesian account might seem to require not only that a cognizer believe the premises of a proof of his belief, the conclusion of a proof of his belief, and believe that the premises are known and lead inductively to the conclusion, but also that the premises are in fact known and in fact lead inductively to the conclusion. Otherwise, it might be argued, the Cartesian would risk making one justified in a belief, in effect, merely if one believed one were justified. However, it is possible to interpret such Cartesian views as avoiding this consequence without requiring anything that violates this definition of internalism: they do so by requiring that the cognizer possess a rational guarantee of truth for all such justifying internal states. This rational guarantee is supposed to be achievable by possessing internal states – if it is held to do so by other means the Cartesian becomes an externalist.

one that employs an internalist account of justification. Epistemological externalism is the view that justification is a function of factors external to the cognizer's awareness, e.g. the reliability of the process through which a belief is formed. Thus, an externalist account of justification will specify some set of external factors pertaining to a belief that make it justified or not. An externalist account of knowledge is one employing an externalist account of justification.<sup>6</sup>

I term a belief "rationally privileged" if a cognizer must prefer it to its negation in order to remain epistemically rational--where it is assumed that cognizers, qua epistemically rational beings, prefer truth or the probability of truth. I introduce this term because it seems to be a useful way to refer to the really important issue in the traditional debate between the skeptic and anti-skeptic. For example, Socrates's and Plato's epistemological debate with the Sophists can be viewed as a debate about the existence of rational privileging: the Sophists claiming that all is mere opinion (doxa) and Socrates and Plato claiming that there are rationally privileged beliefs (epistēmē) that are not the result of bias or upbringing and that the Sophists are irrational not to prefer. The advantage of framing such debates in terms of rational privileging rather than the more typical issues of the existence of knowledge or justification is twofold. First, in the current epistemological atmosphere "knowledge" and "justification" can mean very different things to different people, so it can be both problematic and unclear to describe the debate simply as one about the existence

---

<sup>6</sup> For an attempt to taxonomize the various kinds of externalism and internalism see Alston (1986).

of knowledge or justification. Second, the term and the notion of rational privileging are more descriptive of what is really at stake--especially in highlighting the key role played by reason.

We are now ready to consider the taxonomy of epistemological theories.

### Traditional Theories

Traditional epistemological theories are internalist and hold that their internal justification rationally privileges certain beliefs--some of which rationally privileged beliefs count as knowledge. Examples of such theories are those of Plato, Descartes, Locke, and many of the logical positivists.<sup>73</sup>

However, traditional theories are unacceptable as a neo-Humean basis for constructive epistemology, for in denying to any belief the sort of rational justification that Hume denied to inductive beliefs, the neo-Humean emphatically denies the existence of rational privileging; hence, traditional theories are unacceptable because they make knowledge a sort of belief that, according to the neo-Humean, cannot exist.

---

<sup>7</sup> My claims that traditional epistemological theories, as well as the particular theories of the historical figures I cite, have these features should not be interpreted as an exegetical claim about the history of epistemology. These claims only concern the way in which these theories have been widely understood as operating.

### Pragmatist Theories

Pragmatist theories differ from traditional theories in that they make knowledge, instead of a rationally privileged belief (and so one privileged in regard to truth), a belief privileged in terms of its usefulness.<sup>84</sup>

Such theories are also unacceptable to the neo-Humean because they implicitly rely on the existence of rationally privileged beliefs. For judgments about which beliefs are pragmatically privileged are presumed, by the pragmatist, to be rationally privileged in regard to the truth, i.e. the belief that one belief is more useful than another is to be believed not because it is useful, but because it is true. (Of course, a theory need not rely on rational privileging in this way. Such theories are not addressed in this category.)

### Mere Internalist Theories

Mere internalist theories attempt to solve the difficulties of traditional theories by abandoning rational privileging but retaining internalism. The mere internalist makes justification a matter of possessing internal states that are privileged not by reason but by some internal but more easily attainable means. Some remarks of Nelson Goodman's concerning the justification of inductive and deductive rules of logic are suggestive of one such means:

A rule is amended if it yields an inference we are unwilling to accept; an inference is rejected if it violates a rule we are unwilling to amend. The process of justification is the delicate one of making mutual adjust-

---

<sup>8</sup> As with my remarks about traditional theories, the claims made here about pragmatist theories should not be interpreted as exegetical claims about historical figures.

ments between rules and accepted inferences; and in the agreement thus achieved lies the only justification needed for either. (1979, p.64)

Such a process of "reflective equilibrium"<sup>9</sup> is a means of identifying privileged internal states, but the privilege involved is not rational privilege because it carries no guarantee of achieving truth or probability of truth. Another sort of mere internalist privilege might be employed by a social pragmatist epistemology: those states would be privileged that, according to the cognizer's judicious assessment, were sanctioned by his epistemic community.

We will go into some detail concerning the difficulties with mere internalism in a later chapter. But for now we can say that the difficulty with mere internalism as a basis for constructive epistemology is that it opens the door to epistemological nihilism; for, since the sort of privilege on which the mere internalist bases his epistemology gives us no guarantee of achieving truth or the probability of truth, it becomes unclear, on this view, why we should care about epistemic affairs.

### Externalist Theories

Externalist theories (e.g. those of Goldman, Armstrong, Nozick, McGinn<sup>106</sup>) try to solve the difficulties of mere internalism by replacing internalism with externalism. That is, such theories block the epistemological nihilism that threatened mere internalism by requiring, as do traditional theories, that justification helps to get

---

<sup>9</sup> The term is from Rawls (1971).

<sup>10</sup> Probably the first full-blown externalist theory of knowledge is presented in Goldman (1967). Less explicit precursors are F. P. Ramsey in 1929 (see Ramsey: 1978, pp. 126-27) and Watling (1954).



us to the truth. However, unlike traditional theories, this requirement is made by reference to external features of a belief. The advantage of appealing to external features is that trying to fulfill this requirement by reference to internal features is tantamount to attempting the impossible task of rationally privileging beliefs; but the requirement clearly seems fulfillable when only external features are required.

Since externalist theories do not rely on rational privileging and do not invite epistemological nihilism, they seem to offer the most promise of providing a basis for a constructive neo-Humean epistemology. Consequently, we shall examine their strengths and weaknesses in detail in the next few chapters.

## 2

### OBJECTIONS TO EXTERNALISM

This chapter examines the two most influential types of objection to externalism to determine whether they eliminate externalist theories as candidate bases for constructive epistemology. The type of objection examined in the first section concerns alleged counterexamples to externalist views. The type examined in the second concerns the alleged irrelevance of external features to the following of justificatory norms and, hence, to justification.

#### *2.1 Counterexample-Based Objections to Externalist Theories*

This type of objection uses counterexamples to show that externalist accounts of knowledge or justification are unacceptable because they violate traditional or ordinary intuitions about these concepts. For present purposes, it will be convenient to apply all the counterexample-based objections to be considered here to one paradigm theory, a simple form of the type of externalism known as process reliabilism. (Though some of these objections were originally applied to different sorts of externalist theories, the change should not alter their effectiveness since they were formulated to apply to externalist theories generally.) Reliabilist theories, generally, are externalist theories that make justification and knowledge a function of some external aspect of the reliability of cognitive processes. Reliable indication theories

(see e.g. Armstrong (1973)), the other chief type of reliabilism, make justification and knowledge a function of whether a cognizer's possessing a belief reliably indicates whether the state of affairs that the belief represents obtains. Process reliabilist theories (see e.g. Goldman (1986)) make justification and knowledge a function of whether a belief has been acquired via a reliable process. Our paradigm process reliabilist theory holds that a belief is justified just in case it was acquired via a reliable process (where "reliable" means the process results in true beliefs more than fifty percent of the time (a higher percentage might be required in some cases)) and that a belief constitutes knowledge just in case it is true and so justified. This theory, which closely resembles several of those offered by Alvin Goldman, will be termed simple process reliabilism.

Laurence Bonjour offers one counterexample-based objection to externalism. He describes the following case:

Samantha believes herself to have the power of clairvoyance, though she has no reasons for or against this belief. One day she comes to believe, for no apparent reason, that the President is in New York City. She maintains this belief, appealing to her alleged clairvoyant power even though she is at the same time aware of a massive amount of apparently cogent evidence, consisting of news reports, press releases, allegedly live television pictures, and so on, indicating that the President is at that time in Washington, D.C. Now the President is in fact in New York City, the evidence to the contrary being part of a massive official hoax mounted in the face of an assassination threat. Moreover, Samantha does in fact have completely reliable clairvoyant power under the conditions which were then satisfied, and her belief about the President did result from the operation of that power. (1985, p.38)

Bonjour claims this case demonstrates the insufficiency of externalist conditions for justification.

And, in the case of simple process reliabilism, Samantha's belief about

the whereabouts of the president does seem to meet the conditions for justification: the belief is formed by a reliable process, Samantha's reliable clairvoyant power. But, Bonjour claims, Samantha's belief is clearly not justified, for

Samantha is being thoroughly irrational and irresponsible in disregarding the evidence that the President is not in New York City on the basis of a clairvoyant power which she has no reason at all to think that she possesses; and this irrationality is not somehow canceled by the fact that she happens to be right. (1985, p.39)

Thus, Bonjour's objection to externalist theories seems to be that they provide insufficient conditions for justification because they fail to include rationality as a requirement for being justified, and rationality is an intuitive requirement for being justified.

Another counterexample-based objection to externalism is offered by Stewart Cohen. Cohen tries to show that externalist accounts of justification are unacceptable because they fail to provide necessary conditions for justification. He considers a possible world in which Descartes's evil genius hypothesis holds true, i.e. "our cognitive processes (e.g., perception, memory, inference) are not reliable owing to the machinations of ...[a] malevolent demon" (Cohen:1985, p.281). He then asks us to

imagine two inhabitants of this world, *A*, who is a good reasoner, i.e., reasons in accordance with the canons of inductive inference, and *B*, who engages in confused reasoning, wishful thinking, reliance on emotional attachments..., etc. Since the beliefs of *A* and *B* are both produced by unreliable processes (the evil demon sees to this), a reliabilist theory of justification must render identical epistemic appraisals of both sets of beliefs. (1985, p.283)

Cohen claims, in effect, that such a case shows that externalist theories such as simple process reliabilism fail to provide necessary conditions for justification because, while

neither *A* nor *B* meet these conditions, we intuit that *A*, in contrast to *B*, is clearly justified in his beliefs. Since what *A* possesses and *B* lacks can be fairly characterized as rationality, the upshot of Cohen's example is similar to that of Bonjour's. But while Bonjour's example shows that there can be externalist justification without rationality, Cohen's shows that there can be rationality--and with it justification--without externalist justification.

Even if the counterexamples alleged in this sort of objection are genuine, I would argue that such objections are much less conclusive against externalism than those who make them often assume. For these objections seem to assume that for an epistemological theory to be successful, its theories of justification and knowledge must constitute explications of ordinary or traditional notions. For example, the Cohen and Bonjour objections seem to make this assumption when they consider it a sufficient condition for rejecting externalist theories that they violate traditional intuitions that rationality and justification are intimately connected. Assuming these objections make this assumption, they would seem to fail as attacks on externalism in general for two reasons: many externalists do not make this assumption, and the assumption is false. I take up these two points in order.

To see that not all externalists assume that epistemological theories must explicate ordinary epistemic notions to be successful, we need only consider the variety of purposes for which such theories have been advanced. Though some externalist accounts have clearly been advanced as explications of ordinary notions (e.g. Goldman (1976) and Kornblith (1980)), others have been advanced for entirely

different reasons: some because it seemed a promising way to stop the regress of justification and build a tenable foundationalism (e.g. Armstrong (1973)); some because it was thought important to include animals, small children, and certain artificial cognitive devices among the realm of knowing beings (e.g. Dretske (1981), Alston (1983)); and some because it was thought it would be a good way to naturalize epistemology (e.g. Goldman (1978)). Since these theories are advanced with goals other than the explication of ordinary notions, it would seem clear that not all externalists make this assumption.<sup>11</sup> Thus, counterexample-based objections would seem to fail as general objections to externalism because they do not, as stated, present a difficulty for all externalist theories. However, let us term epistemologies that do not attempt to explicate ordinary epistemic notions radical epistemologies, and those that do, conservative epistemologies. We can say, then, that though counterexample-based objections are ineffective against radical externalist theories, they are effective against conservative externalist theories. Because of this effectiveness, and for other reasons, the remainder of our discussion will focus on the viability of radical externalist theories.

---

<sup>11</sup> Moreover, concerning the particular assumption made in the Bonjour and Cohen objections – that the intuitive connection between rationality and justification ought to be preserved – it is interesting to note that at least one externalist directly rejects it. Alvin Goldman has written in a discussion of skepticism that:

Another dimension of appraisal possibly orthogonal to skepticism is rationality. This is a widely cited epistemic desiderata, one that may be feasible for human beings even if both knowledge and justifiedness are not.  
(Goldman:1986, p. 40)

The second reason that counterexample-based objections fail as general objections to externalism concerned the falsity of the assumption that epistemologies must explicate ordinary notions to be successful. It is important to appreciate this reason as well, for if one does not, one might think that all a critic such as Bonjour or Cohen would need to make his objection work against radical theories would be to include an adequate defense of the assumption. To see why the assumption is false, consider an analogous and familiar example from the history of atomic theory.

An "essential feature" of the concept of the atom in the mid-nineteenth century was indivisibility. It is easy to imagine that many physical scientists at the time believed that so fundamental a tenet as the indivisibility of the atom could not be revoked without abandoning both the concept of an atom and anything like atomic theory. We know better. Atomic theory survived with the successor concept of a divisible atom because modern atomic theory serves an explanatory role and is related to other disciplines very much the way nineteenth century atomic theory was.

The parallel I want to draw with epistemology should be fairly obvious. Just as modern atomic theory has been successful despite employing a concept of atom radically different from the traditional concept, so an epistemological theory might be successful employing epistemic concepts radically different from--and so failing to explicate--traditional notions. It would achieve such success by being able to achieve,

while employing these radical concepts, many of the tasks epistemologies traditionally aim to achieve.<sup>12</sup>

Thus, it would seem to be false that epistemologies must explicate ordinary notions to be successful and, hence, impossible for counterexample-based objections to become effective against radical theories by defending that assumption. Moreover, it seems doubtful that any addition to counterexample-based objections would make them effective against radical theories, for merely pointing out, as such objections do, that the epistemic concepts radical theories employ differ from the traditional concepts, does not demonstrate a defect in such theories since they may be entirely successful despite these differences. Indeed, pointing out such conceptual differences is pointing out the obvious to the radical externalist, for his very strategy is to employ epistemic concepts that are significantly different from traditional versions in order to achieve greater success than traditional epistemology.<sup>13</sup>

Though it would seem doubtful that counterexample-based objections could be effective against radical externalist theories, a remark of Cohen's, occurring later in

---

<sup>12</sup> It might be objected that there is an important disanalogy between epistemology and atomic theory in that it is never considered a task of the latter to explicate ordinary notions about the concepts it employs, but it is often considered a task of the former. It is true that the two are different in this respect, but I see no reason this difference should undermine the overall point. The difference does mean that a radical epistemology, in contrast to a radical physics, will necessarily fail to achieve one traditional epistemological task, the explication of ordinary epistemic notions; but simply because a radical epistemology fails at one task does not mean it fails as an epistemology.

<sup>13</sup> In fairness to Bonjour, it should be noted that he does grudgingly admit that his counterexample-based objections are "ineffective" (1985, p. 57) against theories of the sort I have been calling "radical". However, his remarks and tone suggest that he doubts any radical theory could solve any important epistemological problem, so he would seem to believe that the counterexample is effective against any externalist theory worth considering.



the same discussion in which he presents his counterexample, can be interpreted as an attempt to supplement his objection so that it will be effective against such theories. After noting that a reliabilist might respond to his Cartesian demon criticism by saying that the distinction between *A* and *B* does not concern the justificational status of their beliefs but their being reasonable or rational, he replies:

If the Reliabilist wants to distinguish 'justified' from 'reasonable' or 'rational' he may do so. But clearly the important epistemic concept, the one epistemologists have been concerned with, is what the Reliabilist would call 'reasonability' or 'rationality.' ...the Reliabilist [is]...changing the subject. (pp.283-4)

For present purposes, we can interpret this remark as supplementing Cohen's counterexample-based objection with the claim that if the reliabilist "goes radical" and accepts the counterintuitive consequences of Cohen's counterexample, his epistemology will fail because it will constitute "changing the subject" and will no longer be an epistemology. But such a move fails to make Cohen's objection effective, for there is no reason given to believe that changing the notion of justified in this way will constitute a changing of the subject, and the case of the history of atomic theory clearly illustrates that radical changes in a discipline's fundamental concepts need not constitute a "changing of the subject". However, simply because there is no reason to believe it will constitute changing the subject does not, of course, mean that it will not. To determine whether a particular radical externalist epistemology will, requires examining the nature of a full-blown theory, i.e. considering what sorts of epistemological tasks it would perform, what interdisciplinary relationships it would have, etc. And here, clearly, the burden of proof is on the externalist. To show that he has a

viable epistemology he must show that it can accomplish a significant amount of epistemological work. To date, there seem to be no externalist theories that have clearly been demonstrated to have this capacity. However, this may simply be because the appropriate theory has yet to be developed. Thus, though it seems clear that counterexample-based objections do not refute radical externalist theories, determining whether such theories might be successful requires considering what tasks they might accomplish. After considering some other objections to externalism, we will begin this latter task.

## *2.2 Externalism, Awareness, and Norms*

Another influential type of objection to externalism concerns the alleged irrelevance of external features to justification. Versions of this objection have been presented in a number of places including Pollock (1987) and (1986) and Papineau (1985, p. 377). For our purposes, we can deal with the following generic version:

- (1) External considerations, by their very nature, are features of which we are unaware.
  - (2) In following norms, we can only concern ourselves with things of which we are aware.
  - (3) Hence, we cannot concern ourselves with external factors when following norms.
  - (4) Justification concerns the epistemic norms we follow.
- ∴ (5) Externalist accounts of justification must be wrong since they include in their accounts external factors that cannot be relevant to justification.

Pollock claims his version of this argument constitutes a "refutation" of externalism. However, it seems clear that the radical externalist can defend himself against this objection pretty much as he did against counterexample-based objections. He can

simply deny premiss (4) as a bias of traditional epistemology--one which his radicalized account of justification abandons. And Pollock actually makes reference to this option for the externalist:

The sense of epistemic justification with which I am concerned in this paper is the reason-guiding sense, and if it is acknowledged that...externalism bears only upon another sense of justification then my main point has been conceded. (1987, p.78)

Given this admission, Pollock's claim to have refuted externalism is clearly an overstatement.

### *2.3 the Original Causation Problem*

Before leaving this discussion of the difficulties facing externalist theories, a difficulty I have dubbed the "original causation problem" should be addressed. This problem, while not a significant factor in the internalism vs. externalism debate, is a difficulty that most externalist theories must solve. The problem can best be presented in the form of a counterexample:

Suppose Saddam comes to believe that the world is round via a reliable source, e.g. his high school geography teacher. Later, he comes to place absolute faith in the proclamations of his mentally deranged brother, whom he takes to be a religious prophet. In Saddam's presence, his brother, in the midst of assertions such as "electronic computers are demons from the underworld", "God exists and does not exist", "'Saddam' begins with a 'P'", etc., also proclaims "the world is round". At this point, Saddam's reason for believing the world is round is that his brother says it is.

Saddam's belief that the world is round seems to meet simple process reliabilist requirements for justification, for it has been acquired by a reliable process--Saddam learned it from his high school geography teacher--but he clearly seems to lack

justification for his belief in his later period. Thus, the counterexample would seem to show that simple process reliabilism gives insufficient conditions for justification.

This counterexample is not one which a radical externalist would do well to dismiss as merely reflecting a bias of traditional epistemology, for the intuitions it shows are violated are not merely traditional intuitions but also common-sense intuitions that any externalist wishing to present a consistent theory should respect. Probably the best way for the externalist to solve the problem is to specify that the belief-forming process that determines whether a belief is justified need not be the original process, rather, it is whatever process sustains the belief at a given time. Thus, Saddam's belief in his later period lacks justification because it comes to be sustained by an unreliable belief-forming process. HENCEFORTH, EXTERNALIST THEORIES PRESENTED HERE CAN BE ASSUMED TO INCORPORATE THIS MODIFICATION.

### 3

#### THE REFERENCE CLASS PROBLEM FOR PROCESS RELIABILISM

In the previous chapter we examined several criticisms of externalism and found none that ruled out externalist theories as a possible basis for a successful constructive epistemology. In this chapter we consider an influential criticism of a popular kind of externalist theory, process reliabilism. We pause to consider this more restricted criticism because we will be focusing on process reliabilist theories in coming chapters.

I call the difficulty presented for process reliabilism by this criticism--to supply it with the name it has heretofore lacked--the reference class problem for reliabilism. This name was chosen because the problem is closely related to a difficulty in the frequency interpretation of probability concerning the selection of appropriate reference classes for calculating the probability of individual events.<sup>14</sup>

The problem arises in trying to apply process reliabilist theories of justification to individual beliefs. Consider, for example, the simple process reliabilist theory of justification. This theory, presumably, gives us a recipe for determining whether a particular belief, a belief token, is justified: discover whether the belief token was acquired via a reliable belief-forming process (henceforth "*bfp*"). But to follow this

---

<sup>14</sup> For more on this difficulty for the frequency theory see, e.g. Salmon:1966, pp. 83 ff.

recipe we must resolve an ambiguity in the phrase "reliable *bjp*": does it refer to the particular series of events leading to the creation of the belief token--what we will term a *bjp* token--or to a *bjp* type? Since types are universals and could have no direct causal effect on a belief token, it seems the phrase would best be interpreted as referring to a *bjp* token. But now the really difficult question arises--what does it mean for a *bjp* token to be reliable? Two possible accounts suggest themselves:

- (1) A *bjp* token is reliable just in case it is subsumed by some reliable *bjp* type.
- (2) A *bjp* token is reliable under the same conditions a *bjp* type is reliable.

Unfortunately, neither of these accounts is acceptable for the process reliabilist. But to understand why, we first need to know what it is for a *bjp* type to be reliable. For our purposes we will consider a *bjp* type reliable just in case it tends to produce more true than false beliefs. (There are other plausible accounts: e.g., a *bjp* type is reliable just in case more of its actual and possible belief products are true than false. But we stick to this one for the sake of simplicity.) Now we can begin to see the difficulty with account (1). It turns out that any *bjp* token will be subsumed under many different *bjp* types.<sup>15</sup> For example, if we form a belief reading an article in Tuesday's newspaper, the *bjp* token leading to the belief will be subsumed under the following *bjp* types, among others: forming-a-belief-by-reading, forming-a-belief-by-reading-a-Tuesday-paper, forming-a-belief-while-engaged-in-another-activity, forming-a-belief-while-one-is-alive, etc. Given the great variety of *bjp* types subsuming a

---

<sup>15</sup> In fact there will be indenumerably many, but I will not trouble the reader with a proof of this claim.

token, it becomes clear upon further consideration that for any *bfp* token it will always be possible to gerrymander some reliable *bfp* type that subsumes it<sup>16</sup>; hence, account (1) has the consequence that every belief is justified and is thus an unacceptable interpretation for the process reliabilist. Account (2) does not fare better. Given our account of reliability for *bfp* types, account (2) yields "a *bfp* token is reliable just in case it tends to produce more true than false beliefs". Since *bfp* tokens are unique (both actually and possibly), there is only their actual belief product to indicate their "tendencies". But this means that all belief tokens producing true beliefs will be judged reliable and all those producing false beliefs unreliable, and, hence, that all true beliefs will be counted as justified and all false beliefs unjustified. Obviously, such a result is unacceptable for any theory of justification, so account (2) is also unacceptable for the process reliabilist. Thus, there does not seem to be ready-to-hand any account of the reliability of a *bfp* token that does not make the process reliabilist theory of justification vacuous, and this is the reference class problem for reliabilism.

A number of philosophers have pointed out, in one form or another, this difficulty, e.g. Foley (1985), Feldman (1985), Chisholm (1988), and Pollock (1986). Pollock in particular has been especially bleak about the prospects for a solution:

---

<sup>16</sup> A quick method of gerrymandering such a type for any token is as follows. First, construct a description of the type that picks out the token uniquely or almost uniquely (e.g. by specifying year, date, time, place, etc.). Then take any highly reliable *bfp* type, e.g. forming-a-belief-about-people's-telephone-numbers-by-consulting-a-phone-book. Finally, form another type description by disjoining these two descriptions. The result will be a reliable type subsuming the token in question.

The only way to avoid this is to impose some kind of restriction on gerrymandering, but there does not appear to be any non-ad hoc way of doing that. (1986, p. 119)

[These]... methodological problems seem to me to constitute a decisive refutation of process reliabilism. (1986, p. 121)

However, I suspect that, once again, Pollock has overstated the difficulties posed by his criticisms; for there is reason to believe that, contrary to Pollock, there is such a thing as the reliability of a *bfp* token--despite difficulties in articulating its nature in a philosophically acceptable way. The reason is that we seem to make meaningful judgments about the reliability of *bfp* tokens with relative ease in everyday life. For example, if we are eating on Tuesday at a Chinese restaurant and see that Sheila comes to believe she will receive a raise tomorrow simply because her fortune cookie presages she will, we judge, of course, that Sheila's belief is not reliably formed. And this judgment would seem to be easily interpretable as a judgment about the reliability of a *bfp* token: we judge that the particular series of events leading Sheila to develop her belief was not a reliable way to form the belief. It might be thought that one could also interpret this judgment as being equivalent to the conjunction:

- (a) Believing-what-one-reads-in-fortune-cookies is an unreliable *bfp* type
- and
- (b) The given token is an instance of that type.

Such an analysis would eliminate any reference to the reliability of *bfp* tokens and so provide critics such as Pollock with a means of explaining away apparently meaningful references to the reliability of *bfp* tokens. But the analysis cannot be right. Suppose we were later to discover that the following type subsumed the *bfp* token: believing-what-one-reads-in-fortune-cookies-about-receiving-raises-when-one's-boss-is-



owner-of-the-Chinese-restaurant-at-which-one-is-eating,-and-the-restaurant-makes-its-own-fortunes-and-fortune-cookies. This discovery might lead us, with good reason, to view our initial judgment that the token was unreliable as false, but note that the discovery would have no bearing on the truth of the suggested analysans--the analysans would still be considered (as well as be) true. This difference clearly indicates that the analysandum and the suggested analysans are not equivalent and, hence, that the analysis fails.

It would seem then that we make judgments about the reliability of *bfp* tokens that are not analyzable merely as judgments about the reliability of *bfp* types. Now it is open to Pollock to claim that these judgments are simply confused and not about any genuine feature of *bfp* tokens. But there are two reasons to resist such a move: (1) It is usually poor hermeneutics to interpret people as talking nonsense unless alternative interpretations have been exhausted and (2) there is a plausible alternative interpretation. This interpretation is that in making judgments about the reliability of a *bfp* token we are making a judgment about the reliability of that process type we judge to be critical for the causal production of the belief in question (henceforth I refer to this type simply as the "critical type"). We can put this interpretation in the form of the claim that the judgment is equivalent to the following conjunction:

- (a) The token is subsumed by some *bfp* type *b*
  - (b) *b* is reliable
- and
- (c) *b* is the critical *bfp* type for the token.

Obviously, an assessment of this interpretation depends on what account is given of the notion of the critical type for a *bfp* token. Intuitively, the idea is that the

critical *bfp* type is the one (or, at least, one of a class of types) that, given the nature of the cognizer's environment and psychology, is most relevant to assessing whether the belief token was reliably formed. To get a firmer handle on the notion of a critical type, let us recall the Chinese restaurant example. There are clearly several *bfp* types subsuming the token whose reliability we would judge irrelevant to assessing how reliably the belief was formed, e.g. forming-a-belief-on-a-Tuesday, forming-a-belief-in-a-Chinese-restaurant, forming-a-belief-in-close-proximity-to-food, etc. However, it is difficult to determine just what it is about these types that makes their reliability irrelevant. I suspect that we can see they are irrelevant because we use an intuitive "feel" for what is important in the situation in which the belief was formed in order to select relevant types and exclude irrelevant ones. I would speculate that this feel is the sort of understanding that it is nearly impossible to specify algorithmically (and hence the sort that makes artificial intelligence so difficult), perhaps because it is acquired in the midst of the complex process of becoming a person. However, I would conjecture that we can approximate some of this feel, and so give some definite content to the notion of a critical type, by requiring that the critical type refer only to features that would play a role in a correct psycho-sociological description of the cause of the belief. By "psycho-sociological description" is meant an answer to the question "what caused the belief" that is given in psychological and/or sociological terms. Though there is also no algorithm for determining what features would or would not be mentioned in such a description, relying on such a requirement is a significant improvement over relying on the aforementioned "feel" to give content to

the notion of critical type because the requirement does not merely rely on individual impressions to decide whether a type is critical, and, in most cases, determining whether a type meets the requirement should be relatively non-controversial, for, e.g. the requirement would clearly eliminate the three candidate critical types listed above from consideration because the day of the week, the ethnicity of the food, and the degree of proximity to food, respectively, would play no role in a psycho-sociological description of the cause of the belief.<sup>17</sup>

In addition to excluding types that refer to irrelevant features, there is another sort of type that should not be counted as critical. Examples in the present case are forming-a-belief-by-believing-a-written-fortune, forming-a-belief-by-believing-what-someone-else-has-written-about-you. The problem with these types is that, though both refer to features playing a role in the cause of the belief, they do not seem to include all the features one would think important for assessing the reliability of the cause of the belief. For example, neither mentions that the fortune is written by someone having no knowledge of the person for whom the fortune is predicted. Such types can be denied critical status by requiring that among types meeting other requirements for being the critical type, choose the narrowest.<sup>18</sup>

---

<sup>17</sup> The notion of a "psycho-sociological description of the cause of a belief" is, admittedly, less definite than one would like. However, this indefiniteness is not a major drawback, for its chief function in the account is to serve as a sort of theoretical placeholder for the content of everyday intuitions.

<sup>18</sup> It might be thought such a requirement risks making the critical type one that is so narrowly specified that it includes only the *bfp* token in question. However, I suspect any such narrow specification will include features that violate the previous requirement.

Thus far, we have given the following content to the notion of the critical *bfp* type for a *bfp* token:<sup>19</sup>

- (i) The critical type subsumes the token.
- (ii) The critical type refers only to features that would play a role in a correct psycho-sociological description of the cause of the belief.
- (iii) The critical type is the narrowest of those types meeting the other requirements for the critical type.

If we assign this content to the expression "critical *bfp* type" in the tripartite conjunctive analysis of the notion of the reliability of *bfp* tokens given above:

- (a) The token is subsumed by some *bfp* type *b*
  - (b) *b* is reliable
- and
- (c) *b* is the critical *bfp* type for the token.

we have the account of *bfp* token reliability that I want to recommend to the process reliabilist and now defend.

The first point to be made in its favor is that it solves the reference class problem in that it does not make the process reliabilist theory of justification vacuous in either of the ways described above or in any other way. The two vacuity problems discussed above were the difficulty of making every belief justified and the difficulty of making all and only true beliefs justified. The former difficulty is avoided because

---

<sup>19</sup> The inspiration for this account, as well as the source of the "critical process type" terminology, is the following remark of Goldman's:

Let me advance a conjecture about the selection of process types, without full confidence. The conjecture is: the critical process type is the narrowest type that is causally operative in producing the belief token in question. (1986, p. 50)

Apart from giving one example, Goldman says almost nothing about how we are to interpret his conjecture, so I cannot say to what extent his account and mine agree except to say that they differ in that Goldman wants to limit consideration of *bfp* types to those describing internal psychological processes.

the reliability of a *bfp* token is made dependent on the reliability of one type, the critical process type; hence, it is not possible to make any belief whatever justified by gerrymandering a reliable *bfp* type for it since the gerrymandered type need not be the critical type. The second difficulty is avoided because the reliability of the *bfp* token is made dependent on a process type that has multiple possible and (usually) actual tokens, and the various true-false belief product ratios for the type will lead to various reliability judgments for the critical type, and its corresponding token, regardless of the truth-value of the belief token; hence, both true and false beliefs will be justified and unjustified. Another advantage of the account is that it makes perfect sense of ordinary judgments about the reliability of *bfp* tokens. In particular, it avoids the aforementioned difficulty in making sense of such judgments that plagued the analysis which reduced such judgments merely to claims about the reliability of *bfp* types. That difficulty, it will be recalled, concerned the fact that our judgment about the reliability of the *bfp* token in the Chinese restaurant example would change if we were provided with certain additional information-- despite the irrelevance of that information to judging the truth of the analysans suggested by the account. The difficulty is avoided with the new account because the collateral information (that Sheila's boss owns the restaurant, and it makes its own fortunes and fortune cookies) is relevant to judging the truth of the new analysans, for the information should lead us to conclude that clause (3) was false since the information indicates a different *bfp* type is the critical one. One disadvantage of this account is that, since it fails to offer necessary and sufficient conditions for being the critical *bfp* type, it must be regarded as

incomplete and, hence, the boundaries of the concept of reliability for *bfp* tokens must be regarded as indefinite. Now, for philosophers trained in the analytic tradition (as am I), the inclination is to regard this as a near fatal flaw; however, it is wise to restrain this inclination. The notions of biological species, game, and chair (to name but a few) have so far defeated all attempts to provide necessary and sufficient conditions for their application (at least to my knowledge), yet all serve well in theoretical as well as ordinary contexts, and I see no reason why this account of the reliability of *bfp* types should not do likewise.<sup>20</sup> (It also might help to assuage analytic consciences to recall that the reason for the failure to specify necessary and sufficient conditions is not logical laziness, but the difficulty involved in articulating the "feel" we employ in selecting the critical process type.)

One final note concerning discussions of process reliabilism below. Those discussions frequently apply the phrase "formed by a reliable process" to beliefs. THIS PHRASE SHOULD HENCEFORTH BE UNDERSTOOD AS ABBREVIATORY FOR "FORMED BY A RELIABLE *BFP* TOKEN", WHERE "RELIABLE *BFP* TOKEN" IS UNDERSTOOD AS HAS BEEN EXPLAINED ABOVE. This convention is employed to minimize the literary awkwardness of the inherently awkward rhetoric of analytic epistemology.

---

<sup>20</sup> Indeed, in the many contexts in which this notion is employed in subsequent chapters, its indefiniteness causes no difficulties at all.

## 4

### **EXTERNALISM, OR, HOW TO GET NO MORE FROM KNOWLEDGE THAN TRUTH**

Thus far, we have succeeded in defending externalism from the most influential criticisms of it. It might be thought that the next step toward establishing radical externalism as a basis for a constructive neo-Humean epistemology would be to justify its radical nature by demonstrating the many tasks such an epistemology would accomplish. However, I will not now undertake such a step. The reason is that I believe the externalist approach contains a flaw that is not remedied by radicalization, and I do not want to spend time evaluating a program so flawed. However, I do believe an epistemology incorporating externalist elements avoids this problem, and it was towards the end of carving a niche for it that I engaged in the defense of externalism above. I will attempt to demonstrate the tasks this epistemology might achieve later. In this chapter the ultimate goal is to uncover this heretofore unnoticed flaw in externalism. However, an equally important goal, achieved en route to this criticism, is the explosion of a dogma of contemporary epistemology--that the primary end of epistemic endeavors is truth--and the bringing to light of the many other important goals there are for such endeavors in addition to truth.

#### *4.1 Vericentrism an Epistemology*

Contemporary epistemologists spend a great deal of effort trying to determine what knowledge is, but, somewhat surprisingly, they give little attention to the question of why we should bother obtaining knowledge or engage in epistemic endeavors. And this question would seem to be of no small import, for, for instance, educational institutions have as their announced aim the imparting of knowledge; so the question would also seem to be the question of the purpose of education.

When the question is addressed by contemporary epistemologists (of either externalist or internalist stripe), the answer given seems almost always to be that the goal of epistemic endeavors is truth or, more precisely, true belief. This view of the goal of epistemic endeavors will be termed epistemological vericentrism. More specifically, vericentrism is the view that all genuinely epistemic or cognitive endeavors have, with minor exceptions, true belief<sup>21</sup> as their only goal. This view is expressed commonly enough, and without argument, that it constitutes a dogma of contemporary epistemology. It is evidenced explicitly in the following quotations from the externalists Goldman and Alston and the internalist Bonjour. Goldman writes:

...I...wish to preserve the centrality of true belief in the network of intellectual values. ...We have various cognitive faculties and operations that seem to promote true belief...and this is their principal utility. (1986, pp.138-9)

---

<sup>21</sup> To avoid being diverted by problems with vericentrism that are of no concern here, I will include in this general goal of true belief considerations involved in the epistemic maintenance of one's belief corpus that are related to, though not literally a component of, the goal of acquiring true belief, e.g. the avoidance of false beliefs, the maximization of the number of true beliefs, etc.



Alston speaks of

...the basic aim of believing or, more generally, of cognition, viz., to believe truly rather than falsely. (1988, p.269)

While he is discussing the role of epistemic justification Bonjour writes:

...the goal of our distinctly cognitive endeavors is truth.... (1985, p.7)

A rough corollary of the general claim vericentrism makes about epistemic and cognitive endeavors is that the primary utility of justification is as a means to truth. For it is surely also true that the immediate goal of epistemic endeavors is obtaining justification, and since the vericentrist seems to hold that their ultimate goal is truth, it would seem that justification is beneficial only as a means to truth. Bonjour makes this claim explicitly:

If truth were somehow immediately and unproblematically accessible..., then the concept of justification would be of little significance and would play no independent role in cognition. ... The basic role of justification is that of a means to truth.... (Ibid.)

On this view justification acquires a role analogous to Wittgenstein's ladder--once it gets us where we want to go (truth) we no longer have any use for it since its only use is getting us where we wanted to go.

Of course, one does not point out the existence of a dogma unless one intends to object to it. My objection to vericentrism--including its more specific corollary--is that there are other goals for epistemic endeavors besides true belief; and these other goals derive from the fact that justification, at least as it is traditionally conceived, is more than just a means to truth. There are other benefits to justification--and hence to knowledge, since justification is a component of knowledge--and, it will be argued,

these other benefits, what will be termed the extra-verify utilities of knowledge, provide a significant goal for epistemic endeavors additional to true belief.

The relevance of the existence of these utilities to my criticism of externalism is that, according to it, externalist knowledge is relatively impotent because it lacks these extra-verify utilities. Obviously, however, it should first be established that there are such utilities before criticizing externalist knowledge for lacking them, and it is to this task that we move next.

#### *4.2 The Extra-verify Utility of Knowledge*

Since dogmas concern traditional views, it will be most appropriate to exploding the dogma of vericentrism to consider whether knowledge as it has been traditionally conceived possesses extra-verify utility. Thus, we turn first to an examination of the possibility of extra-verify utility for knowledge as it has been viewed by traditional epistemological theories (in the sense of "traditional" described in chapter 1); henceforth, we will refer to such knowledge as t-knowledge, and to the corresponding notion of justification as t-justification. Later, in connection with the criticism of externalism, we will consider the extra-verify utility of externalist knowledge.

In general, the question of extra-verify utility for knowledge is the question of whether some of the usefulness of an item of knowledge derives from a feature other than its being a true belief. That is, clearly truth is, in most cases, a useful feature of a belief--e.g. if one has true rather than false beliefs concerning the presence of lions in one's bedroom one will either avoid being eaten or avoid needless anxiety. Extra-

verific utility concerns what other benefits there might be to having knowledge apart from the having of a true belief. Idiosyncratic accounts of knowledge aside, knowledge is distinguished from mere true belief by justification; hence, if knowledge possesses extra-verific utility it must derive from justification.

#### *4.2.1 The Extra-verific Utility of T-knowledge*

Since extra-verific utility can only derive from justification, to determine whether t-knowledge possesses extra-verific utility we must consider whether t-justification provides any. This can be determined by performing a thought experiment in which we add t-justification to a particular case of mere true belief and see what utility is added. Since the original case is already one in which the belief is true, any added utility should be extra-verific utility.

Though it will be claimed here that the results of the thought experiment would be the same for almost any theory of t-knowledge or t-justification, it would be impractical to attempt to prove this claim by testing all such theories. Instead, one traditional account will be treated as paradigmatic, and I will attempt to conduct the experiment in such a way that the results will not depend on the account's peculiarities. The paradigm will be the view that a belief has knowledge-sufficient justification just in case it is either one of the kinds of (supposed) infallibly self-justifying beliefs (i.e. either a belief deriving from the (supposed) perceptually given or an a priori belief in (supposed) analytic propositions) or is infallibly derivable from such beliefs. Since this view bears many similarities to Descartes's epistemology, it will

be referred to as Cartesian foundationalism; however, it is clearly different from Descartes's actual view on a number of points.

The first step in the experiment is to describe a case of mere true belief. Suppose in ancient Rome Julius Caesar's brother Justinian comes by the mere true belief that

*m*: On the ides of March a group of men will attempt to assassinate Caesar.

We may suppose that his belief is merely true because he has come by it via a gypsy fortune teller.

Next, we will provide Justinian with Cartesian-style justification for his belief. Suppose that, in addition to seeming to hear *m* from a Gypsy fortune teller, Justinian acquires a variety of other infallibly self-justifying basic beliefs. For example, he eavesdrops on senators' conversations and acquires the belief that he seems to have heard Brutus discussing such a plan on two occasions, and he intercepts some of the senators' mail and acquires the belief that he seems to have seen a letter signed by several senators swearing themselves to such a plan, etc. And suppose that these beliefs, together with other knowledge possessed by Justinian, provide him with infallible justification for *m*, so that he now has available a demonstration that *m*.<sup>22</sup>

There seem to be at least five respects in which the addition of Cartesian justification increases the utility of Justinian's belief. First, since Justinian now has infallible justification for his belief, he will (typically) believe that he can't be wrong

---

<sup>22</sup> Since, in line with my neo-Humeanism, I believe no Cartesian-style knowledge or justification can exist, I will not attempt to fully describe a case in which he does – the reader must use his epistemological imagination here.

about it. This belief, in turn, should lead to an increase in the actual subjective probability of his belief in *m*. And this increase in subjective probability is beneficial because it is an increase in the subjective probability of a true belief.

Second, the justification for the belief can make Justinian aware of certain methods of acquiring justified beliefs and so help him acquire more justified beliefs. For example, he is now aware that he might get more justified--and very important--beliefs by surreptitious surveillance of senators' correspondence and conversations.

The justification provides three other benefits because it increases what will be termed the "infectiousness" of the belief. The infectiousness of a belief is the ease with which it may be transmitted to others.<sup>23</sup> The justification increases the belief's infectiousness because Justinian can now--at least with an ideal case of Cartesian justification--convince others of the belief by providing a conclusive proof of it using his justification.<sup>24</sup>

The first benefit relating to the increased infectiousness of Justinian's belief is that since the justification is, in this case, making a true belief infectious, it will have

---

<sup>23</sup> This notion of infectiousness bears some relation to recent discussions of the epistemology of testimony. (See e.g. Peacocke (1986), Fricker (1987), and McDowell (forthcoming).) These discussions focus on the question of how an assertion of a known proposition can transmit knowledge to a hearer of the assertion. However, since the infectiousness of knowledge concerns only the spread of belief – not knowledge, these issues are quite distinct and should not be run together.

<sup>24</sup> A "conclusive proof" is probably overstating the case; for self-justifying beliefs about how things appear that are involved in Justinian's justification for his belief will be self-justifying only for him and only when he is having the relevant perceptual state, so he will not be able to provide as strong a proof to others as he could have provided for himself, and so will not be able to provide a conclusive proof. But we can ignore this complication for the moment since it only affects the degree of the increase of infectiousness – not whether there will be any.

the advantage for the community as a whole of aiding the spread of a true belief through the community; and, for obvious reasons, it is generally better that community members have true rather than false beliefs. (Though, of course, true belief is not always beneficial for a community. For example, in the present case, if Rome will be better off with another emperor it might be better that no one else knew of the plot.)

A second infectiousness-related benefit of Justinian's belief is that it will give him greater power to influence the beliefs of others--especially since this increased infectiousness is under his control (he can offer the proof or not, as he wishes) and so pick and choose whom he wishes to infect with his belief. Clearly, such power will be of great advantage in motivating others to act against the conspiracy--especially those dubious of the reliability of fortune tellers.

A third infectiousness-related benefit of Cartesian justification is that since it increases infectiousness it will also tend to increase unanimity of opinion within the community. Unanimity of opinion is a benefit for the community when cooperative ventures are undertaken. (However, in some cases it is advantageous to have community members with diverse viewpoints, e.g. in a legislative body. Hence, this benefit is perhaps not as straightforward as the other two.)

In the case of Justinian and his belief in *m*, we seem to have discovered five extra-*verific* benefits for knowledge that are provided by Cartesian justification. We can summarize these benefits as follows:

- (1) Increase in the subjective probability for the cognizer of a true proposition.
- (2) Help in acquisition of more justified beliefs.
- (3) Increase in the belief's infectiousness, yielding the following benefits:
  - (i) Facilitating the spread of a true belief through the community (since the infectiousness of a true belief is increased).
  - (ii) Increasing the ability of the cognizer to influence the beliefs of others.
  - (iii) Increasing community unanimity.

We should consider now whether our discovery of extra-verify utility for this instance of Cartesian knowledge warrants the conclusion that Cartesian knowledge in general has each of these extra-verify utilities. We will consider two objections to that conclusion.

First, it might be objected that some of this extra-verify utility cannot be attributed to Cartesian knowledge in general because the utility is relative to the kind of community in which the belief occurs. For example, Cartesian justification will increase the infectiousness of a belief only in a community which has faith in the deliverances of Cartesian reason. (Perhaps community members would be irrational to question Cartesian reason, but no law of nature prevents human beings from being irrational.) Hence, the utilities deriving from increased infectiousness will not obtain for Cartesian knowledge in such a community. And, therefore, one cannot ascribe the utilities accruing to this case of Cartesian knowledge to Cartesian knowledge generally, for not all cases will occur in communities with congenial epistemic norms.

The initial reply to this objection is that both the infectiousness and the utility of Cartesian knowledge refer to potentialities. Certain communities or individuals may be unable to take advantage of them because of peculiar beliefs, e.g. that

Cartesian reason is untrustworthy, but nevertheless they remain and so exist even in societies with uncongenial epistemic norms.

But this characterization of utility and infectiousness opens the door to another difficulty. Since any mere true belief--in fact any belief at all--will have some characteristic in which some possible community's epistemic norms would place absolute faith, such an account of infectiousness runs the risk of rendering mere true beliefs and Cartesian knowledge equally infectious because of equal community-relative potential unless, e.g., some criteria for measuring the degree of infectiousness are included. And if all beliefs are equally infectious, it makes no sense to speak of increased infectiousness as some sort of utility.

The best way to solve this problem is probably to abandon the notion of a trans-community degree of infectiousness and, instead, make distinctions in our list of utilities in terms of kinds of infectiousness and degree of community-relative infectiousness to achieve utilities that are generalizable and avoid the preceding objection. Let us revise our list of utilities 3(i)-(iii) accordingly:

- (3) Provides a high quality and high degree of infectiousness for a belief in three respects:
  - (i) Provides truth-linked infectiousness and so facilitates the spread of true belief throughout the community.
  - (ii) Insures a high degree of infectiousness in a rational community, and so insures the cognizer a high degree of power to influence the beliefs of others in such a community.
  - (iii) Insures a high degree of infectiousness in a rational community, and so will promote unanimity in such a community.

Utility 3(i) now specifies a certain kind of infectiousness that is provided by Cartesian justification: "truth-linked" infectiousness. Infectiousness is truth-linked



just in case it derives from properties that are at least more often than not associated with true beliefs. Cartesian justification only accrues to true beliefs, so, clearly, infectiousness deriving from it is truth-linked. Specifying that Cartesian justification provides truth-linked infectiousness rather than increased infectiousness, solves two problems. First, it eliminates the objectionable mention in the old version of 3(i) of an increase in absolute infectiousness. Second, it specifies how, despite there being no increase in absolute infectiousness, Cartesian justification facilitates the spread of true belief. It facilitates it because it introduces infectiousness into the community that will only spread true beliefs. The sort of infectiousness accruing to mere true belief, on the other hand, may spread true or false beliefs indiscriminately. For example, the infectiousness deriving from Justinian's mere true belief that *m* being acquired via a fortune teller, would aid the spread of true belief in that one case, but, generally, this sort of infectiousness would not aid the spread of true belief because it is not truth-linked.

Utilities 3(ii) and 3(iii) now specify that Cartesian justification insures a high degree of relative rather than absolute infectiousness. The degree of infectiousness is made relative to a "rational" community. For a community to count as rational in this sense it is not required that its members always act rationally or only believe what is rational, but only that, at minimum, its members are, for the most part, susceptible to the various means of rational persuasion (though, for a variety of reasons, rational considerations may fail to persuade them in many cases). I take it it will be granted me that human communities are rational communities in this sense and

that living in a rational community is, generally speaking, preferable to living in a non-rational community. Granting these presumptions, the shift to rational-community-relative infectiousness solves difficulties similar to those solved by the shift to truth-linked infectiousness. First, it eliminates objectionable talk of increase in absolute infectiousness.<sup>25</sup> Second, it helps to show why, if Cartesian justification provides no increase in absolute infectiousness, it is beneficial in terms of power to influence others and promotion of community unanimity. It is beneficial because it promotes these features relative to rational communities, and such promotion is beneficial because that is the sort of community in which we reside and because rational communities are one of the best sorts of communities to reside in. Moreover, rationality is a feature of the human community as a whole; hence, if one ventures outside one's local human community, infectiousness relative to rational communities will certainly be applicable while parochial sorts of infectiousness may not.

These changes to our original list yield a list of utilities possessed by Cartesian knowledge in general, with one exception. Utility (1) (increase in subjective probability) suffers from the same defects we found above in the original utility 3(i), but these defects can also be remedied similarly. Below is a list of utilities that makes the

---

<sup>25</sup> The shift to relative infectiousness is alone sufficient to defeat the problems with talk of increase in absolute infectiousness noted above, but mention even of the increase of relative infectiousness was deleted because even it presents a problem: and makes clear sense only in cases in which one acquires a belief and later acquires Cartesian justification for it, as in the case of Justinian's belief in m; since all cases are not of this type, mention of increased infectiousness is omitted tout court.

necessary changes to utility (1) and which will count as our final version of the general extra-  
verific utilities of possessing Cartesian knowledge:<sup>26</sup>

- (1) Provides a truth-linked assignment of high subjective probability to the proposition believed.
- (2) Facilitates the acquisition of more justified beliefs.
- (3) Provides a high quality and high degree of infectiousness for a belief in three respects:
  - (i) Provides truth-linked infectiousness and so facilitates the spread of true belief throughout the community.
  - (ii) Insures a high degree of infectiousness in a rational community, and so insures the cognizer a high degree of power to influence the beliefs of others in such a community.
  - (iii) Insures a high degree of infectiousness in a rational community, and so will promote unanimity in such a community.

Thus, it can be seen that one type of t-knowledge, Cartesian knowledge, possesses extra-  
verific utility. Do all other types of t-knowledge possess such utility? It seems likely that most  
any other type of t-knowledge would also possess such utilities. To see why, consider how the  
chief differences among types of t-knowledge would affect their extra-verific utility. For  
example, consider a theory that required only fallibilist and coherentist t-justification for t-  
knowledge. The change to fallibilist justification clearly would have some effect on the extra-  
verific utility, e.g. the assignment of subjective probability would no longer be 1 and the degree  
of infectiousness of the belief relative to a rational society would no longer be so high; but it  
seems the extra-verific utilities would remain nevertheless. Nor would the switch to coherentism  
eliminate extra-verific utilities, for, since the coherentist fallibilist justification is t-

---

<sup>26</sup> The reader may note that some of these benefits are dependent on the belief's being true; nevertheless, they are additional to the utility one derives directly from the truth of the belief and so count as extra-verific.

justification, we know it will rationally privilege a belief in terms of its probability of being true; hence, because of the nature of rational privileging, it will provide infectiousness as surely as would foundationalist justification.

It would seem, then, that vericentrism is wrong, for justification is more than just a means to truth, and knowledge, at least as it is traditionally conceived, possesses extra-*verific* utilities which provide a goal for epistemic endeavors besides truth. Now that we have dispensed with vericentrism, I want to begin to show how its falsity can be used to mount a criticism of externalism. The first step in this criticism is to consider whether externalist knowledge possesses extra-*verific* utility.

#### *4.2.2 Externalism and Extra-*verific* Utility*

To determine whether externalist knowledge possesses extra-*verific* utility we can employ the same method used to determine whether *t*-knowledge possessed it--pick one view as paradigmatic and perform the relevant thought experiment on it. The paradigm will be the simple process reliabilist theory discussed in chapters 2 and 3. Henceforth, we will refer to knowledge on the simple process reliabilist theory as "*r*-knowledge" and justification of this sort as "*r*-justification".

In performing the experiment on *r*-knowledge we can use the same example of mere true belief used in our investigation of *t*-knowledge and then see if the addition of *r*-justification increases the utility of the belief. Adding *r*-justification means adding some reliable belief-forming process to the case. However, in choosing a process to add we must be careful. For it would be easy to choose a process that yielded extra-*verific* utility due to features additional to its reliability. For example,

considered as a belief-forming process, Cartesian justification is reliable, but it would be inappropriate to choose it to test r-knowledge since it adds much more than r-justification. Let us add then, the following facts to the case. Let us suppose that, unknown to Justinian, the fortune teller is a senator in disguise, who wants to warn Caesar anonymously of the plot. In this way we add reliability of belief-forming process to the example, but little else of relevance.<sup>27</sup>

With one small exception, the addition of r-justification to the belief does not seem to increase its utility. In particular, it fails to add the extra-verify utilities possessed by t-knowledge. The r-justification will not provide Justinian with utility (1), for he is not aware that his belief-forming process is reliable and so cannot adjust his subjective probability assignments accordingly. It will not provide utility (2) for similar reasons--Justinian is unaware of his belief's r-justification, so he cannot use it to acquire more justified beliefs. The infectiousness-derived utilities would seem to suffer similarly: since Justinian is not aware of his r-justification, he cannot convey it to others; hence, he cannot use it to positively affect the infectiousness of the belief to provide any of utilities 3(i)-(iii).

The exception to r-justification's lack of extra-verify utility is a slight increase in some of the infectiousness-derived utilities. For though, as was just noted,

---

<sup>27</sup> In fairness to the externalist, it should be noted that such an example would not count as knowledge according to what is probably the most sophisticated version of process reliabilism, that expressed in Goldman (1986) (concerning such examples see Goldman's reply to Putnam, p. 110). Nevertheless, I want to retain the example for a number of reasons. First, it draws the distinctions I want to draw more clearly than other examples. Second, the more sophisticated version is both extremely complicated and largely incomplete, so it is far from clear what cases it counts as knowledge. Third, I do not believe this version solves the problems which the present example illustrates.

Justinian himself cannot use his r-justification to positively affect the belief's infectiousness and so increase these utilities, his limitation does not prevent others from noticing its r-justification and so becoming more susceptible to infection; and the result of their increased susceptibility would be a utility-producing change in infectiousness.

To see how such third-person awareness of Justinian's r-justification could produce extra-verific utility, consider the following situation. Suppose Marcus, another senator, knows that a colleague plans to disguise himself as a fortune teller to warn Justinian of some plot. Justinian tells Marcus of his belief and its source. Marcus guesses the identity of the fortune teller and so becomes aware of Justinian's r-justification, i.e. becomes aware that Justinian's belief was acquired by a reliable process. Consequently, Marcus becomes infected with belief in *m*.

The possibility of such cases provides clear justification for ascribing extra-verific utility 3(i) to r-knowledge. For, clearly, any infectiousness deriving from r-justification will be truth-linked and will help spread true belief through the community. Utility 3(ii) will not be attained for it requires that the cognizer intend to make use of his infectiousness, and this Justinian cannot do. The presence of utility 3(iii) is arguable, for it is not clear that, e.g., Marcus is coming to believe in *m* because he is rational. But let us be charitable to the externalist and suppose that for present purposes it is because he is rational.

It would seem, then, that r-knowledge does possess some extra-verific utility; however, I would argue that, compared to the extra-verific utility of t-knowledge, the

extra-verific utility of r-knowledge is hardly worth noting. For it pertains only to cases of third-person awareness of r-justification as described above, and it seems clear that such cases are rare; moreover, even when they occur, r-knowledge possesses at most two of the five utilities of t-knowledge.<sup>28</sup>

We should now consider whether the extra-verific utility poverty of r-knowledge extends to externalist theories of knowledge generally. To determine whether it does, and also because it is an interesting topic in its own right, we will examine in detail the reasons for the utility disparity between r-knowledge and t-knowledge.

#### *4.3 Why R-Knowledge and T-Knowledge Differ in Extra-verific Utility*

Obviously, the utility disparity must be due to some difference between r-knowledge and t-knowledge. The most obvious difference is that t-knowledge requires internalist and r-knowledge externalist justification. But it will be useful to examine the differences in greater detail.

---

<sup>28</sup> A defender of r-knowledge might object that the case against extra-verific utility for r-knowledge has been unfairly conducted, for the typical case of r-knowledge will not be one in which there is no first-person awareness of r-justification – as in the present instance in which Justinian’s awareness is blocked by the senator’s disguise – but will be one in which such awareness is present, as when one is looking up some item in an encyclopedia. And since much of the criticism of r-knowledge has been predicated on this lack of first-person awareness, the criticism is unfair.

The reply to this objection is to point out that, as was noted above, the case of the disguised senator was specifically designed to add only r-justification to the case of mere true belief, and this involves excluding first-person awareness as well as Cartesian justification. The rationale for the former exclusion is that first-person awareness is an internal state of the cognizer and, hence, clearly not a component of r-knowledge. Thus, any contribution it might make to the utility of a belief is not to be attributed to r-knowledge.

Some of these differences are nicely detailed in what is perhaps the earliest criticism of a reliabilist account of knowledge. This criticism is made by Wilfrid Sellars in his 1956 lecture "Empiricism and the Philosophy of Mind". He there considers the following proposal concerning the nature of observational knowledge:

An overt or covert token of 'This is green' in the presence of a green item...expresses observational knowledge if and only if it is a manifestation of a tendency to produce overt or covert tokens of 'This is green'--given a certain set--if and only if a green object is being looked at in standard conditions. (1963, p.167)

This is a reliabilist account of observational knowledge since it makes knowing hinge on the possession of a reliable "tendency". Sellars's criticism of this view relies upon his claim (henceforth the "authority recognition condition" for knowledge) that

...to be the expression of knowledge, a report must not only have authority, this authority must in some sense be recognized by the person whose report it is. (Ibid., p.168)

And the discussion which follows implies that since the proposed view of observational knowledge does not require that the observer satisfy the authority recognition condition, it gives insufficient conditions for the possession of observational knowledge. Sellars says that to satisfy the authority recognition condition the observer

must know that tokens of 'This is green' are symptoms of the presence of green objects in conditions which are standard for visual perception. (Ibid.)

Moreover, Sellars implies that the proposed view also fails to include a related requirement for knowledge of any type:

...in characterizing an episode...as that of knowing...we are placing it in the logical space of reasons, of justifying and being able to justify what one says. (Ibid., p.169)



The reliabilist proposal fails to place knowledge in the "logical space of reasons" largely because it does not make having reasons for one's belief a requirement for justification.

I am not now concerned with the correctness of Sellars's criticism<sup>29</sup> but, as noted above, bring it up to delineate the differences between t-knowledge and r-knowledge. The first such difference is revealed in his claim that the reliabilist account is deficient because it ignores the authority recognition condition for knowledge. Clearly, r-knowledge has no requirement of authority recognition for knowledge. For, e.g., in Justinian's case of r-knowledge, he had no awareness or recognition of the special authority of his belief, for he has no idea that his belief was acquired via a reliable process rather than fortune telling. Though our version of t-knowledge does not explicitly require that the cognizer recognize the authority of her belief, it does make such recognition easily accessible, and it is typically the case (or would seem to be, if such knowledge existed) that the cognizer actually recognize the authority of her belief. For example, to have Cartesian t-knowledge, the cognizer must possess a sort of proof of his belief. And to the extent that one can possess a proof of a proposition and not recognize that it is a proof of a proposition, Cartesian knowledge does not require that one recognize the authority of one's knowledge. However, the possession of a proof clearly facilitates the recognition of such authority

---

<sup>29</sup> In fact, as it stands, Sellars's criticism is effective against a conservative but not a radical reliabilist. For a radical reliabilist can simply reply that "Yes, your criticism does note a discrepancy between traditional and reliabilist accounts, but I am not trying to create an account of knowledge that fits the traditional mold. So for what reason, then, must knowledge in my sense meet the authority recognition condition or be in the logical space of reasons?" Alston (1983) makes a similar response to Sellars's criticism of reliabilism.

and, in typical cases, it would seem that the possessor of such t-knowledge would recognize the authority of his belief, for, typically, when one possesses a proof, one recognizes it as a proof (at least if we make the requirements for possession of a proof reasonably strong). Thus, though authority recognition is not an explicit doctrinal difference between r-knowledge and t-knowledge, it would seem to be a difference nonetheless.

The second difference is revealed in Sellars's criticism of reliabilism for ignoring the logical space of reasons. Clearly, t-knowledge, but not r-knowledge, requires the having of reasons for knowledge; hence, only the former will be what will be termed a "reasons-based" account.

There is a third important difference not illustrated by Sellars's criticisms. Since r-knowledge does not require that the cognizer possess reasons for her belief, a fortiori, it requires no rationally privileged reasons. Thus, only t-justification requires rational privileging.

Let us consider which of these differences might account for the disparity in extra-*verific* utility between the two sorts of knowledge. The first difference noted concerned authority recognition. Clearly, authority recognition is an important factor in the extra-*verific* utility of t-knowledge. For it is authority recognition that allows the cognizer to treat his knowings differently from his mere true believings, and such differential treatment is important for obtaining each of the extra-*verific* utilities. For example, in the case in which Justinian possessed Cartesian knowledge of *m*, he recognized the authority of his belief, i.e. that he possessed Cartesian justification for

it and so could not be wrong. This recognition allowed him to see that his belief was especially trustworthy and adjust the subjective probability he assigned his belief accordingly (utility 1). Since it made him aware that his belief had t-justification, it made it possible for him to use that justification to acquire more justification (utility 2). Since the recognition made him aware that he possessed a proof of his belief, it made him aware that he had that resource to infect others (utility 3).

The second difference, making justification a matter of reason giving, also contributes to the utility disparity. Possessing justification of a reasons-based sort insures that the justification will, because it requires the possession of reasons, have a positive effect on the belief's infectiousness.

The third difference, rational privileging, also plays a role in the disparity and in two ways. First, it insures a high degree of infectiousness, for it insures (at least in the strongest form of rational privileging) that a full proof of the truth of the belief will be available, and, insofar as people are rational, they should be persuaded by a genuinely full proof. And, it would seem, human beings, although far from perfectly rational, are highly susceptible to persuasion by means of such proofs. Second, rational privileging insures a high quality of infectiousness and subjective probability assignment by insuring that both are truth-linked.

Thus, each of these three differences would seem to contribute to the utility disparity. Each points to some aspect of t-knowledge that contributes to extra-verify utility but is lacking in r-knowledge. But in addition to lacking these three attributes, r-knowledge also possesses certain features t-knowledge lacks, and we should consider

what effect these might have on the disparity. The primary relevant feature is the requirement that the belief be acquired by a reliable belief-forming process. The only effect on the disparity this requirement would seem to have is to narrow the gap slightly by providing the belief with an authority that others might recognize, as was noted above.

We began this inquiry into the causes of the utility disparity partly to determine whether the utility poverty of r-knowledge extends to externalist accounts generally. Each of the features of t-knowledge that we have found to be responsible for the utility disparity are features that no purely externalist account of knowledge could have, for each feature pertains to the cognizer's internal states. Thus, it would seem likely that the utility poverty does extend to externalist accounts generally. Now it is time to consider whether this utility deficit constitutes a problem for externalism.

#### *4.4 The Flaw in Externalism*

So far I have been content merely to point out the disparity in extra-verific utility between t-knowledge and externalist knowledge. However, I also believe externalist knowledge's lack of extra-verific utility constitutes a flaw in externalism. But so far the radical externalist could reply to this charge as to the charges of Bonjour and Cohen (see chapter 2): he could admit that there is this difference between t-knowledge and r-knowledge, but protest that a difference is not ipso facto a flaw.

But this charge is on firmer ground than the criticisms discussed above, for it does not merely concern the failure of externalist knowledge to satisfy our traditional intuitions of what knowledge is, but also concerns a practical deficiency in externalist knowledge--it does not do as much for us as we might expect knowledge to do. And the extra-*verific* utilities externalist knowledge lacks are, arguably, of substantial significance for the well-being of society, for they promote unanimity and so enhance the ability to engage in cooperative projects, and they promote unanimity of correct opinion and so increase the chances that such projects will succeed. The benefits accruing to the individual from these utilities are also significant. They give her power to influence the views of others (and only about correct opinions), and to have a high degree of confidence in correct opinions.

The radical externalist might respond to such a plea by saying that "it would be nice, for the reasons indicated, if knowledge, on one's account, had such extra-*verific* utility; but this is not the only desideratum for an account of knowledge. T-knowledge, for example, possesses extra-*verific* utility, but the traditional account cannot be the basis for a constructive epistemology simply because there can be no t-knowledge. The externalist account, on the other hand, can provide such a basis because externalist knowledge clearly can exist. Perhaps extra-*verific* utility is just what must be sacrificed if we are to achieve the greater good of making epistemology a constructive discipline."

This response would be fine except for one difficulty. It turns out that this sacrifice is not necessary to provide the basis for a constructive epistemology, for

there is a theory of knowledge that both makes knowledge a kind of belief that can exist and preserves extra-verify utility for knowledge. This theory will be the subject of the next chapter.

# 5

## **BILEVEL RELIABILISM: AN ACCOUNT OF KNOWLEDGE AND JUSTIFICATION**

Chapter 4 uncovered a defect in externalist epistemologies. In terms of the taxonomy of epistemological theories presented in chapter 1, the inadequacy of externalist approaches exhausts the options for a constructive neo-Humean epistemology. If we are to continue the search for a basis for constructive epistemology, we must either develop new options or somehow resurrect the old. In this chapter an attempt will be made to do both. First we consider whether an original option, mere internalism, might be resurrected. Then we consider a new option, mixed internalist/externalist epistemologies.

### *5.1 Mere Internalism Reconsidered*

Mere internalist theories, it will be recalled, make justification a matter of possessing some set of internal states that are privileged but not rationally privileged. Mere internalist theories are worth reconsidering at this point because they seem a natural way to avoid the difficulty encountered by externalism in the last chapter. The reason is that they avoid the difficulties of traditional theories by eschewing rational privileging, but retain the internalist character of traditional theories and so

offer hope that they might also retain some of the extra-verific utility of t-knowledge and so avoid one of externalism's chief difficulties.

But there are a number of obstacles to fulfilling the promise of mere internalism. First, rational privileging plays a non-trivial role in providing t-knowledge with extra-verific utility. It is what gives t-justification its infective "punch". Moreover, it is what links t-justification to truth, so it is what makes the truth-linked extra-verific utilities ((1) and 3(i)) possible. A mere internalist account must find some way to make up for the contribution of rational privileging if the level and quality of extra-verific utility is to remain the same. Second, rational privileging not only gives t-knowledge much of its extra-verific utility, it is also what makes t-justification a means to truth and so makes t-justification a provider of verific utility. A mere internalist account must also find some means of making up for this contribution of rational privileging if justification is to remain a provider of verific utility.

A natural way to attempt to overcome these difficulties is to try to specify some non-rational privilege in regard to truth that a belief must have and that will play some of the roles of rational privileging. Since we are now considering pure internalist theories, we can choose only from internalist sorts of privilege (in particular, we can't attempt to privilege reasons externally via, e.g., an externalist requirement that they be truth conducive). However, it is difficult to see how there could be any sort of purely internalist privileging in regard to truth that would both not be a form of rational privileging and would help mere internalism overcome these obstacles. For the most obviously helpful form of internalist privileging with regard to the



truth would involve the ability to "see" from our (internal) perspective that a belief is privileged over its competitors regarding truth, for that would allow us to make the best use of its privileged status. But any privileging that involves such "seeing" would seem tantamount to rational privileging. Thus there seems little hope of finding some sort of non-rational privileging in regard to truth that can play the aforementioned truth related roles of rational privileging.

The only avenue left for the mere internalist is to consider means of privileging beliefs irrespective of truth, and see to what extent these might achieve something akin to the (verific and/or extra-verific) utility of rational privileging. For example, one might privilege beliefs in terms of the cognizer's possessing justifying reasons of a sort that is socially approved. *Prima facie*, such a suggestion is promising since it seems a good way to obtain infectiousness without rational privileging--make one's belief infectious not because it is rationally superior but because one has reasons that the rest of society deem acceptable. But the rest of the utility of t-justification disappears with such a view. All truth-linked utilities (infectiousness-related and otherwise) disappear, as well as all verific utilities since such justification is no longer a means to truth. The only utility such knowledge would possess is the infectiousness-related extra-verific utilities pertaining to ability to influence others (3(ii)) and unanimity (3(iii)). Thus, such knowledge would seem to be even poorer than r-knowledge in overall utility.

Mere internalism, then, does not seem very promising as a basis for constructive epistemology. Its main advantage, that it does not require rational privileging for

knowledge, also seems to be its chief drawback; for without rational privileging mere internalist knowledge suffers from the same utility poverty as externalist knowledge.

### *5.2 Mixed Internalist/Externalist Accounts*

Reasoning that approaches that are inadequate separately might be adequate when combined, several philosophers have proposed mixed internalist/externalist accounts in recent years (e.g. Peacocke (1986) and Alston (1988)). There are two important respects in which mixed accounts can vary: by making different combinations of internalist and externalist elements and by constituting either conservative or radical epistemologies. We will consider first whether conservative mixed views might yield viable constructive epistemologies.

Since conservative epistemologies have different goals than radical ones, we must evaluate them on somewhat different terms. Since conservative mixed views are proposed in an attempt to develop an account of knowledge and justification that satisfies our intuitions about what these things are, we can use their adequacy on this count as a means of evaluation.

Two desiderata have been paramount in the quest to satisfy such intuitions: avoidance of counterexamples and preservation of the intuition that we do and can have considerable quantities of knowledge. The latter aim is achieved by these mixed views by eschewing rational privileging as a requirement for justification. Typically, proponents of such views begin their argument that their epistemologies achieve the former aim by considering pure externalist accounts. They claim that externalist

conditions for justification are clearly necessary, but, in addition, they claim (e.g. because of Bonjour type counterexamples) that they are not sufficient. The missing ingredient, it is then claimed, is the internalist requirement that the cognizer be acting "responsibly" from her epistemic perspective. (Such a requirement of epistemic responsibility constitutes the addition of an internalist element to the account of justification because responsibility pertains to the interconnections between one's internal states.) Typically, acting with epistemic responsibility will involve having reasons for one's belief that are, in some sense, good.

Let us see how this argument works with a particular example. Recall the Bonjour counterexample to externalism discussed in chapter 2 in which Samantha comes to believe (truly) via a reliable clairvoyant power--of which she is unaware--that the president is in New York when all the newspapers say he is in Chicago. Proponents of conservative mixed views claim that the reason we are reluctant to ascribe justification or knowledge to Samantha in this case is that she is being epistemically irresponsible--she has no evidence or good reasons for her belief, yet she is not reluctant to adopt it. Conservative mixed views avoid such counterexamples because they also have the internalist requirement that Samantha act responsibly, e.g. that she adopt the relevant belief only if she has substantial evidence for believing that she has a reliable clairvoyant power and that the belief resulted from it.

On the basis of such cases, proponents of these views conclude that their theory satisfies our intuitions about knowledge and justification. The chief difficulty with this claim becomes evident once we begin to inquire further into the nature of epistemic

responsibility and ask just what it is about so-called "epistemically responsible" behavior that makes it responsible. Two prima facie attractive answers are not available to the conservative mixed theorist: that it is responsible because it is rationally privileged and that it is responsible because it is externally privileged. The former answer is unavailable for obvious reasons, the latter because responsibility is a feature of internal states. The most likely remaining answer would seem to be that epistemically responsible behavior is just behavior that we are comfortable with and that is traditionally accepted by society for one reason or another. But once this answer is accepted, it seems increasingly unlikely that mixed views will ultimately satisfy our epistemic intuitions. For though they might initially be satisfied because the behavior required for the possession of knowledge is behavior with which we are comfortable, once it is made clear that this behavior lacks rational privileging our intuitions will be violated. Thus, e.g., suppose the above case is improved so that Samantha is acting epistemically responsibly because a reputable psychologist has told her that she has a reliable clairvoyant power and that it was responsible for the relevant belief. Our intuitions may initially be happy, but once we delve deeper and discover all the unjustified foundational beliefs on which the scientist bases his claim, our intuitions should react against the claim that Samantha has knowledge.

The problem, then, with conservative mixed views, (even, I would claim, with those that include some internalist element other than epistemic responsibility) would seem to be that their internalist element is too weak to satisfy our intuitions about

counterexamples but making it stronger risks making knowledge and justification unreachable, e.g. by requiring rational privileging.

### *5.3 Bilevel Reliabilism, a Radical Mixed View*

Let us move now to the category of radical mixed theories--the sole category of epistemological theory, of those we have so far mentioned, that we have yet to find unacceptable. But the approach in examining this category will be somewhat different from that employed in examining the others. No attempt will be made to come to a general conclusion about the theories so far proposed in this category, for I know of none. Rather, the primary goal will be to present a radical mixed theory, termed "bilevel reliabilism", that avoids the aforementioned difficulties of externalism--the poverty of extra-verific utility--and traditional internalism--the impossibility of knowledge--and so might serve as the basis for a constructive epistemology.<sup>30</sup>

To see what bilevel reliabilism is and how it avoids these difficulties, let us consider how we might, at this point in the dialectic, construct a theory that avoids them. The natural way of avoiding the difficulty concerning the impossibility of knowledge is to omit rational privileging as a requirement for knowledge, as with the mere internalist approach. But the mere internalist then tries to achieve extra-verific utility and build a positive epistemology with what remains of traditional internalism,

---

<sup>30</sup> Though there have been no radical mixed theories proposed as such, some of the conservative mixed theories discussed in the preceding section could serve double duty as radical theories. I believe bilevel reliabilism is superior to such theories, and towards the end of this chapter I argue for its superiority over one such theory that has attracted some attention.

and, as we have seen, this attempt is unsuccessful. But, once one has decided to eschew rational privileging, there is another natural path to follow to try to solve the second difficulty, and that is to try to "beef up" externalist theories with internalist elements (other than rational privileging) so that the resulting knowledge will possess extra-verific utility. Let us see how this second path might be pursued.

In chapter 4 we discovered three features of t-knowledge that accounted for its utility advantage over r-knowledge: an authority recognition condition, a reasons-based account of justification, and rational privileging. An obvious way to try to "beef up" the utility of r-knowledge is to try to add some or all of these features to r-knowledge. Since rational privileging would obviously be an unacceptable addition, we are left to consider the other two prospects. I first consider the possibility of adding an authority recognition condition to r-knowledge--a path that leads to bilevel reliabilism--and then consider whether reasons-based justification should be required.

Our present goal, then, is to see if some sort of authority recognition condition can be added to the notion of r-knowledge so that its extra-verific utility will be increased. Let us recall Sellars's version of the condition:

...to be the expression of knowledge, a report must not only have authority, this authority must in some sense be recognized by the person whose report it is.  
(1963, p.168)

In terms more appropriate for our present endeavor, we can state the condition as the requirement that a belief constitutes knowledge only if the believer recognizes the authority of the belief. The most straightforward way of adding the requirement to the account of r-knowledge would, of course, simply be to add it to the other

requirements for r-knowledge, that a belief be true and acquired by a reliable process. But the straightforward approach saddles our account of knowledge with two unacceptable unclaritys: unclarity about what constitutes recognition and unclarity about what authority is to be recognized. The unclarity concerning recognition is particularly worrisome because recognition is an epistemic term, and, in fact, for our purposes "recognizing the authority of a belief" is virtually synonymous with "knowing that a belief has authority". Thus, so construed, the condition would create a circular account of knowledge if it were merely appended to the requirements for r-knowledge. To avoid this circularity we must use some account of recognition that avoids epistemic terms. One easy means of doing this is to interpret recognition in reliabilist terms, i.e. to say that one recognizes a belief's authority when one r-knows it has authority. This interpretation can then be easily converted into an account without epistemic terms via the definition of r-knowledge. The unclarity concerning the nature of the authority to be recognized can be even more easily eliminated: we simply specify that the authority is what makes r-knowledge distinctive, its being acquired by a reliable process.

These clarifications result in the following account of authority recognition:

*S* recognizes the authority of his belief in *p* iff:

*S* r-knows that his belief in *p* was acquired by a reliable process.

Before putting this account to use there is one more ambiguity to be resolved: is the analysans satisfied if *S* merely r-knows that his belief in *p* was acquired by some process which (for all *S* r-knows) may or may not be reliable, but turns out to be

reliable? Or must *S* also r-know by which process he acquired his belief and that that process is reliable? I opt for the latter, stronger reading:

*S* recognizes the authority of his belief in *p* iff:

- (a) *S* r-knows that his belief in *p* was acquired via belief acquisition process *b*.
- and
- (b) *S* r-knows that *b* is a reliable belief acquisition process.

Hereafter, this sense of authority recognition will be referred to as the r-recognition of authority.

This notion of the r-recognition of authority can be used to produce the following version of r-knowledge "beefed-up" by the addition of an authority recognition condition:<sup>31</sup>

Knowledge is belief that is true, acquired by a reliable process, and whose authority is r-recognized by the cognizer.

This form of reliabilism is what I referred to above as bilevel reliabilism. I give it this name because it requires for knowledge reliably acquired belief at two levels--the belief constituting knowledge and the belief about that belief's authority. (Henceforth, when it is required for clarity, I will refer to the higher level beliefs as "meta-beliefs" and the lower level belief as the "object belief".)

In addition to an account of knowledge, bilevel reliabilism includes an account of justification. This account is a straightforward corollary to the account of br-knowledge (as it will henceforth be termed). One will be br-justified in a belief in *p*

---

<sup>31</sup> The reader should take care not to forget the convention announced at the end of chapter 3 concerning the interpretation of expressions such as "acquired by a reliable process" in interpreting this account.



when one's belief fulfills either all the requirements for br-knowledge or all of them except the requirement that  $p$  be true. More precisely:

$S$  is br-justified in his belief that  $p$  iff his belief is r-justified (i.e. acquired by a reliable process), and  $S$  r-recognizes the authority of his belief.

Another way to view the notion of br-justification is that it requires r-justification on two levels-- with the additional requirement that the meta-beliefs are true and, hence, r-known. Thus, the process of br-justification can be viewed as having two parts. First, one acquires an object belief via a reliable process, e.g. looks it up in an encyclopedia. Second, one r-recognizes the authority of the object belief, e.g. comes to believe, truly and via a reliable process, that looking up things in an encyclopedia is a reliable belief acquisition process and that the object belief was acquired by that process.

Let us consider a concrete example of br-knowledge. Suppose a librarian, Lester, acquires the true belief that

$h$ : Belize was formerly British Honduras

by reading it in an encyclopedia. Since it is true, Lester's belief that  $h$  constitutes br-knowledge if it is br-justified. It will be br-justified if it meets the two requirements for br-justification. It meets the first requirement because it is acquired by the reliable process of consulting an encyclopedia. It will meet the second if Lester r-recognizes the authority of his belief in  $h$ . He will do so if, in turn, he fulfills the two requirements for the r-recognition of authority. We can suppose he fulfills clause (a) because he r-knows, via reliable introspection, that he acquired his object belief by

consulting an encyclopedia. And we can suppose he fulfills clause (b) because his professional experience with encyclopedias allows him to r-know that consulting an encyclopedia is reliable.

#### *5.4 The Extra-verified Utility of Br-knowledge*

Having succeeded in developing a mixed theory based on r-knowledge but including an authority recognition condition, we should now consider whether this theory achieves its intended purpose, i.e. whether br-knowledge possesses a significant amount of extra-verified utility. Previously we determined whether a kind of knowledge possessed extra-verified utility by performing a thought experiment in which the relevant sort of justification was added to a case of mere true belief and then noting the effect of the addition on the utility of the belief. We can test br-knowledge in the same way.

Let us consider an example of mere true belief in which pathological liar Laurence makes claim *e* to his friend Eileen:

*e*: There will be a total eclipse of the sun tomorrow.

Suppose further that *e* is true and Eileen thus comes to have the mere true belief that *e*. Let us now add br-justification to Eileen's belief. Suppose that Eileen disbelieves Laurence's claim at first, and proceeds to verify it in a world almanac, and then believes *e*. We can suppose that she thus achieves the two components of br-justification. She achieves the first of r-justifying her object belief since it is now acquired by the reliable process of consulting a reliable source. The second component is the r-

recognition of the authority of the object belief, which, it will be recalled, also has two components: r-knowing that one's belief was acquired by reliable process  $b$ , and r-knowing that  $b$  is a reliable belief acquisition process. We will suppose she achieves the first because she uses reliable introspective and observational powers to notice that she acquired her object belief from the world almanac. And we will suppose she achieves the second because her past experience with world almanacs and similar reference sources reliably indicates that they are reliable sources on matters of this kind.

Now we should consider whether the addition of br-justification increases the utility of the belief. Likely suppliers of increased utility are the sorts of extra-*verific* utilities noted for t-justification in chapter 4:

- (1) Provides a truth-linked assignment of high subjective probability to the proposition believed.
- (2) Facilitates the acquisition of more justified beliefs.
- (3) Provides a high quality and high degree of infectiousness for a belief in three respects:
  - (i) Provides truth-linked infectiousness and so facilitates the spread of true belief throughout the community.
  - (ii) Insures a high degree of infectiousness in a rational community, and so insures the cognizer a high degree of power to influence the beliefs of others in such a community.
  - (iii) Insures a high degree of infectiousness in a rational community, and so will promote unanimity in such a community.

We saw in chapter 4 that the first component of br-justification, r-justification, brings about a small increase in utility 3(i). Whatever improvement br-justification is to make on this performance will hinge on the other component, authority recognition.

Authority recognition does seem to provide utility (1). Authority recognition provides Eileen with the meta-belief that her belief in  $e$  was reliably acquired, and, in

typical cases, this meta-belief should lead her to assign a high subjective probability to  $e$ . And this assignment will be truth-linked; for since her meta-belief that her object belief was reliably acquired must be true if she is br-justified, the assignment will be made only to reliably acquired object beliefs, which, of course, are mostly true.

Utility (2) also is obtained. Authority recognition provides Eileen with the information that almanacs are reliable sources of information. She can use that information by consulting almanacs to obtain more br-justified beliefs.

Utilities 3(i)-(iii) hinge on the quality and degree of infectiousness of a belief, so to see if they are obtained we should consider whether authority recognition improves the quality or degree of infectiousness. There seem to be two principal ways in which authority recognition can affect infectiousness. First, authority recognition provides Eileen with meta-beliefs that she can use to convince others of her object belief. For example, she can say to her philosopher friend Phillip "I read in the world almanac that there will be a total eclipse of the sun tomorrow", and also report, if he is unaware of the reliability of almanacs, the content of her other meta-belief, that almanacs are reliable sources. The reporting of meta-beliefs clearly improves the quality of infectiousness. It does so because the infectiousness provided is truth-linked, for to be a component of br-justification the meta-beliefs must be true, and if the meta-beliefs are true then the object belief must be reliably acquired and thus, usually true; hence, meta-beliefs of br-justification will be reportable mostly in cases in which the object belief is true, and so the infectiousness is truth-linked. The reporting of meta-beliefs can also be seen as contributing to the degree of infec-

tiousness. However, it would not seem to have any special effect on the degree of infectiousness in rational communities as did *t*-knowledge, for *br*-justification provides no sort of rational privileging. In the present case, the difference can be seen by noting that the fact that Eileen is *br*-justified provides no guarantee that she have any reason at all for believing that almanacs are reliable sources of information, and so there is no special rational persuasiveness that accrues to her reports of her meta-beliefs. However, there is a type of community in which it would be desirable in which to live and in which such reports would affect the degree of infectiousness--a community whose members had, to a significant degree, reliable belief forming processes. For the members of such a community would be prone to recognize the truth of Eileen's reports of her meta-beliefs, and so would be infected with the meta-belief and then would also be more readily infected with the object-belief. Thus, the reporting of meta-beliefs would contribute to the degree of infectiousness in such a community.

A second way in which authority recognition improves the degree and quality of infectiousness concerns the direction in which Eileen can pursue a defense of her object belief. Merely reporting meta-beliefs, as above, begins such a defense, but suppose these too are questioned. Her defense can continue in one of two directions: Eileen can attempt to defend *e* directly, by trying to convince the doubter that there will be a total eclipse of the sun tomorrow, or she can attempt to defend it indirectly by defending her meta-beliefs in the authority of her belief in *e*. The first means of defense is unaffected by the possession of the sort of authority recognition supplied by

br-justification, but the second is facilitated. To see how, let us consider how the second sort of defense might occur in the present case. Suppose Eileen mentioned "There is going to be a total eclipse of the sun tomorrow--I read it in the world almanac, which is reliable on such matters" to a visitor from Sudan who knew nothing of world almanacs. If the Sudanese doubts her claim that  $e$ , she might then defend her belief indirectly by defending her meta-belief in the reliability of world almanacs, e.g. by showing the Sudanese a world almanac and noting its accuracy on other matters. It can be seen that the r-recognition of authority facilitates her defense in two ways. First, it supplies Eileen with the meta-beliefs to defend. Second, since the authority recognition guarantees that the meta-beliefs Eileen is defending are true, the recognition increases the chances that the defense will be successful, at least among inquirers (such as the Sudanese and Eileen) who are capable of discovering the truth of the matter.

Since the defensibility of a belief is part and parcel of its infectiousness, the r-recognition of authority's contribution to the defensibility of Eileen's belief represents a contribution to its infectiousness. To determine whether this contribution yields extra-verific utility, we must consider whether the infectiousness contributed is an improvement in the degree and/or quality of infectiousness. It seems to be both. The infectiousness is truth-linked because it derives from br-justification, and br-justification is truth-linked. And the infectiousness contributed is of high degree in a desirable community type. As with authority recognition's other contribution to infectiousness, this sort of infectiousness will be effective in a community whose members

have, by and large, reliable belief forming processes; for the defense of meta-beliefs employed in the case of br-justification will always be of true beliefs and, hence, will have good chances of succeeding in such a society.

In sum, then, br-justification provides infectiousness-derived utilities in two ways: (1) it provides meta-beliefs the mere reporting of which can improve infectiousness and (2) it facilitates the indirect defense of object beliefs via the defense of meta-beliefs.

We see then that br-knowledge possesses extra-verific utility in abundance. In fact, apart from utilities 3(ii)-(iii) being relativized to a different sort of community, br-knowledge seems to possess all the sorts of extra-verific utilities of t-knowledge, though perhaps to a lesser degree.

### *5.5 Bilevel Reliabilism and Reasons-based Justification*

Above we noted that there are two ways in which one might try to "beef up" r-knowledge to increase extra-verific utility: add an authority recognition condition and require a reasons-based account of justification. Bilevel reliabilism is the result of pursuing the former option, and we should now consider whether the other option ought also to be pursued.

Since, as we have seen in the preceding section, pursuing the path of authority recognition has led to the successful theory of bilevel reliabilism, the other option seems most promising considered as an addition to bilevel reliabilism, i.e. the most attractive way to "beef-up" a reliabilist theory with reasons-based justification would

be to add it not to r-knowledge but to br-knowledge. But there are several reasons for not making this move. First, the primary upshot of such a move would be to exclude certain cases of br-knowledge--those lacking certain sorts of reasons--from the category of knowings. There would be a reason to exclude such cases if we were presently attempting to construct a conservative epistemology, for such cases clearly violate intuitions about the nature of knowledge. But since bilevel reliabilism is a radical epistemology, violations of intuitions are not a primary concern. Second, since there are no rationally privileged reasons, it is unclear if there is any real basis for selecting a sort of reasons to require for knowledge. Third, there are certainly a number of functions reasons serve, e.g. they seem to help us get to the truth, they seem to play an important role in our intuitions about what counts as knowledge, and they have an important semantic role as well. Since these functions must be recognized somewhere in one's theory of the world, why, it might be asked, not recognize them in one's theory of knowledge? In fact, each of these functions is addressed in my overall epistemology. The first is addressed simply in the bilevel reliability requirement, for if reasons help one get to the truth, that will be one more way in which one can reliably acquire a belief. The second and (to some extent) the third will be addressed in the account of folk epistemology presented in subsequent chapters.



## *5.6 Bilevel Reliabilism and Alston's "Internalist Externalism"*

Another mixed internalist/externalist theory, the account of justification presented in William Alston's "An Internalist Externalism" (1988), has attracted a good deal of attention recently. If this view were presented as a radical epistemology it would avoid the criticisms of conservative mixed views made above and would be a competitor to the bilevel reliabilist account. I argue below that if it is so construed, bilevel reliabilism is still the superior view. However, in addition to introducing Alston's view to provide a dialectical opponent for bilevel reliabilism to vanquish, I also want to consider it because it has many instructive similarities and differences with bilevel reliabilism the illuminating of which should aid in the appreciation of the special character of bilevel reliabilism.

### *5.6.1 Alston's Account*

Alston has generally written in defense of externalist theories such as reliabilism, though he has usually stopped short of endorsing a particular reliabilist theory. But in this article he confesses some dissatisfaction with the usual reliabilist account of justification:

We find something incongruous, or conceptually impossible, in the notion of my being justified in believing that p while totally lacking any capacity to determine what is responsible for that justification. Thus, when reliability theorists of justification maintain that any reliably formed belief is ipso facto justified, most of us balk. ...since it is possible for a belief to be reliably formed without the subject's having any capacity to determine this.... (1988, p.272)

He then asks why we have such intuitions about justification, i.e., why we think that for a subject to be justified in a belief, he must have "access" to the justification. In

his answer he notes first his view that the notion of "being justified in holding a belief" is important to us because we have

...the practice of critical reflection on our beliefs, of challenging their credentials and responding to such challenges--in short the practice of attempting to justify beliefs. (1988, p.273)

Granting this, Alston argues, it becomes clear why we intuitively require that a subject have access to his justification if he is to be justified:

But then it is quite understandable that the concept should include the requirement that the justifier be accessible to the subject. For only what the subject can ascertain can be cited by that subject in response to a challenge. (1988, p.274)

Thus far, Alston's line of argument should seem somewhat reminiscent of arguments presented above.<sup>32</sup> He notes an intuitive difficulty with reliabilist accounts of justification--that they guarantee the subject no access to the justification for his belief--and explains the intuition that such access is required by appeal to our practices of justifying our beliefs to others. Above, we discovered a defect in externalism in the form of a lack of extra-*verific* utility for externalist knowledge, part of this lack concerned the degree of infectiousness of a belief, and responding to challenges to a belief is, of course, a determinant of infectiousness.

Having presented a difficulty with reliabilism and analyzed its source, Alston proposes a solution. He proposes a view of epistemic justification which is fundamentally reliabilist and externalist, but includes an internalist element to solve the

---

<sup>32</sup> Just for the record, it should be noted that the theory of bilevel reliabilism as well as the line of thought leading to it were developed before its developer had read Alston's article.

problem concerning a subject's access to the justification for his belief. As I understand it his view is that

*S* is (prima facie) justified in believing *p* iff

(1) *S*'s belief that *p* is "based on an adequate ground". (p.265)

(2) The ground is "the sort of thing whose instances are fairly directly accessible to [*S*] on reflection". (p.275)

(To simplify matters I ignore the additional, and less interesting, conditions that must be fulfilled to be ultima facie justified on Alston's account.) Alston phrases his account in idiosyncratic terminology, so I will explain it before considering the account's plausibility.

To clarify Alston's position, we must consider what he means by the ground of a belief, being based on a ground, and being an adequate ground. Concerning "ground" he says:

I...use 'ground' for the psychological input to the belief forming mechanism, i.e., the belief or experience.... (p.267)

and

...the "ground" for a belief is not what we might call the total concrete input to the belief forming mechanism, but rather those features of that input that are actually taken account of in forming the belief, in, so to say, "choosing" a propositional content for the belief. (p.268)

One of Alston's examples of a ground is one's belief that the streets are wet when that grounds one's belief that it rained last night. It should be stressed that for Alston a ground is always a psychological state of the subject, either a belief or an experience. A belief is based on a ground only if the belief is causally dependent on the ground, but, of course, mere causal dependence is not sufficient. As to what conditions are both necessary and sufficient, Alston suggests:

Wherever it is clear that a belief is based on another belief or on an experience, the belief forming "process" or "mechanism" is taking account of that ground or features thereof, being guided by it, even if this does not involve the conscious utilisation of a belief in a support relation. To say that my belief that the streets are wet is based on the way they look is to say that in forming a belief about the condition of the streets I (or the belief forming "mechanism") am differentially sensitive to the way the streets look; the mechanism is so constituted that the belief formed about the streets will be some, possibly very complex, function of the visual experience input. (p.266)

Thus, a belief is based on a ground if and only if the belief is appropriately causally dependent on another belief or experience. Clearly, however, being based on a ground is insufficient for being justified in a belief. For example, one's belief that one will become rich today may be based on the ground that one believes a horoscope's prediction to that effect. Alston requires, in addition, that to be justified one's belief must be based on an adequate ground. And he says that the "adequacy of a ground...depend[s] on its being a sufficiently strong indication of the truth of the belief grounded" (p.270).

We now have a rough idea of the meaning of clause (1). If this had been all there was to Alston's view of justification, it would already have both internalist and externalist components. It is internalist in that it requires that the ground of a belief be a (internal) psychological state. It is externalist in its requirement that the ground be truth conducive.

But if Alston stopped with clause (1) he would have left the difficulty he ascribed to reliabilist accounts of justification unsolved. For though he has required that the ground be an internal state, he has made no provisions for insuring that this ground will be available for citation if the belief is challenged. This is the job of the

second clause, which adds a second internalist element. It insures that in most cases the subject will be able to cite (at least upon reflection) his ground in defense of his belief. (Alston's reasons for not requiring accessibility are irrelevant to the present discussion.) Thus, examples such as Bonjour's clairvoyance case--in which the believer has no means of defending his belief--do not come out justified on Alston's account, for there is no ground for the belief, and, so, a fortiori there is no ground accessible to the cognizer to defend his belief.

### *5.6.2 Alston's Account and Bilevel Reliabilism*

I want now to stress just one important respect in which bilevel reliabilism is superior to Alston's account.

The internalist element of br-justification, it will be recalled, is the authority recognition condition:

*S* recognizes the authority of his belief in *p* iff:

- (a) *S* r-knows that his belief in *p* was acquired via belief acquisition process *b*.
- and
- (b) *S* r-knows that *b* is a reliable belief acquisition process.

Similar to Alston's requirement that the subject have access to the ground of his belief (clause (2)) is clause (a) of bilevel reliabilism's authority recognition condition. In each case the subject has a belief with which he might conceivably respond if his original belief is challenged. Clause (b) also has a place in Alston's scheme, it is tantamount to recognizing that a ground is adequate; but Alston specifically excludes recognition of the adequacy of the ground from his conditions for justification (p.281). One of his reasons for doing so seems to be that he did not believe he could

include this condition and preserve the externalist character of his account. But, in any event, excluding it renders his account inferior to bilevel reliabilism.

The inferiority is reflected in the effect that excluding such recognition has on the extra-  
verific utility of knowledge. Let us consider the effect in the case of his example of believing it  
rained last night on the ground that the streets are wet. Suppose that this ground is sufficiently  
truth conducive to be adequate (broken water mains, etc. not occurring often enough to undercut  
adequacy), and suppose that the subject has access to the ground (i.e. if someone asks him why  
he thinks it rained last night he says "the streets are wet"), so clause (2) is also satisfied. But  
suppose further that the subject does not recognize the adequacy of his ground (the further  
requirement that bilevel reliabilism makes (clause (b)), i.e. that he does not recognize that the  
ground is truth conducive.

The effect of this omission on the extra-  
verific utility justification leads to knowledge is disastrous. Recall once again the extra-  
verific utilities of t-knowledge:

- (1) Provides a truth-linked assignment of high subjective probability to the proposition believed.
- (2) Facilitates the acquisition of more justified beliefs.
- (3) Provides a high quality and high degree of infectiousness for a belief in three respects:
  - (i) Provides truth-linked infectiousness and so facilitates the spread of true belief throughout the community.
  - (ii) Insures a high degree of infectiousness in a rational community, and so insures the cognizer a high degree of power to influence the beliefs of others in such a community.
  - (iii) Insures a high degree of infectiousness in a rational community, and so will promote unanimity in such a community.

The subject will not obtain utility (1). For he will not place a high subjective probability on his belief that it rained last night because he has an adequate ground,

for he has no idea whether the ground is truth conducive. Similarly, he cannot use his justification to acquire more justified beliefs, for he has no beliefs about whether this belief is epistemically better off than a mere true belief. Infectiousness (and the benefits that accompany it) is improved slightly, for the subject can defend his belief with "the streets are wet", and this defense will be both truth-linked and more effective in a society whose members have mostly reliable belief forming processes. Such infectiousness is similar to that achieved by the possessor of br-knowledge by the mere reporting of the content of meta-beliefs. But the infectiousness of br-knowledge pertaining to the defense of meta-beliefs is not obtained. For the subject has no idea whether his belief was acquired from an adequate ground, so he cannot defend his belief on the basis that it was reliably acquired.

Thus, in terms of extra-verify utility, Alston's account would seem to make only a small improvement upon the lot of the usual reliabilist accounts, whereas bilevel reliabilism succeeds almost as well as traditional accounts.

### *5.7 An Externalist Objection*

Because of its advantage in extra-verify utility, bilevel reliabilism would seem to be superior to any pure externalist theory. However, the externalist might contest our assumption that possession of extra-verify utility constitutes such a decisive advantage for bilevel reliabilism. He might claim that the extra-verify utility of br-knowledge comes at a price: the authority recognition condition included by bilevel reliabilism to achieve extra-verify utility excludes cases of mere r-knowledge--cases

of r-knowledge in which authority recognition is absent--from the bilevel reliabilist account of knowledge. And this exclusion is a problem, the externalist might argue, because non-human animals and pre-linguistic children seem to have lots of knowledge, but this exclusion would seem to prevent them from having any. The externalist might then claim that there is a pure externalist alternative superior to bilevel reliabilism. This epistemology would consist of the simple reliabilist view represented by r-knowledge, and would relegate to one of its branches other than the theory of knowledge and justification the issue of extra-verific utility. This epistemology would, the externalist might claim, be superior to bilevel reliabilism because it preserved the possibility of knowledge for animals and children while at the same time dealing with the issue of extra-verific utility.

The bilevel reliabilist's best reply to such an objection is probably just to point out that the pure externalist alternative is preferable only if the intuition that animals and children have knowledge is right. But there are, of course, well known reasons to doubt that that intuition is right. But, assuming these reasons are inconclusive, there would appear to be a standoff between bilevel reliabilism and the pure externalist alternative.<sup>33</sup>

If these views were developed into full-blown epistemologies, and we could see the explanatory roles the concept of knowledge played in each, perhaps there

---

<sup>33</sup> The reader should not conclude from the existence of this standoff that the criticism of externalism in chapter 4 comes to nought, for it has, even if the pure externalist option is embraced, forced the externalist to recognize the importance of extra-verific utility. And the recognition of the importance of extra-verific utility is, I think, the most important result of this criticism.



would be some firmer basis for preferring one to the other. But, at this point, we must proceed on the basis of educated hunches. My hunch is that bilevel reliabilism is the more promising of the two theories, and, having shown above that it avoids the difficulties of the usual epistemological alternatives, I want next to see what in the way of constructive epistemology might be built upon this foundation.

## 6

### EXPLAINING EPISTEMIC PRACTICES: PRELIMINARIES

In bilevel reliabilism we have discovered a radical mixed theory of knowledge and justification that escapes each of the objections that have so far been raised against externalist and internalist epistemologies. To establish it as a promising basis for constructive epistemology, there remains the all-important task of justifying its radical nature by showing that there are significant epistemological tasks it can accomplish. In this chapter the nature of one such task is discussed. In the next chapter the task is carried out.

#### *6.1 A Task for Bilevel Reliabilism*

There are some tasks that it would likely be fruitless for any radical or largely externalist theory of knowledge to attempt to undertake, e.g. explicating the ordinary concept of knowledge or, since external features are outside a cognizer's perspective, providing first-person guidance in evaluating epistemic norms. One sort of task that would seem suitable for both radical and purely or partly externalist theories, such as bilevel reliabilism, is the examination of a cognizer or group of cognizers from an "external" perspective, i.e., roughly, a perspective in which the examiner is privy to certain epistemically relevant features of a cognizer's environment that are external to

the cognizer's perspective. One example of such a situation would be a designer's evaluation of the performance of a robot with artificial intelligence. In such a case the designer would, typically, have access to many features of the robot's environment external to its perspective but crucial to evaluating its success, e.g. the design of the robot's hardware and software, whether the robot was operating as designed, whether the robot was finding the correct solutions to the problems it faced, e.g. whether it correctly identified the blue box as the blue box.

The particular task to which bilevel reliabilism will be applied is of this sort and will involve taking an external stance to a problem that is, in my view, one of the most important facing a neo-Humean epistemology (i.e. an epistemology that denies the existence of rationally privileged beliefs). The problem concerns the epistemic practices we engage in daily, our practices of inquiry and justification. These practices seem, for all the world, to be aimed at the rational privileging of beliefs. But if, as the neo-Humean epistemologist maintains, there is no rational privileging, then the question naturally arises as to just what the function, if any, of these practices is and just why they have arisen.

I want to apply bilevel reliabilism to this problem by using it to provide an explanation of the existence and function of epistemic practices. Such a project can be approached rather naturally from an external stance. We can consider the project from the viewpoint of an anthropologist who is attempting to determine the explanation of the existence and function of some cultural practice. Though in this case we are part of the culture being examined (human culture), the external stance can still be

taken since we will often be aware of many features of the epistemic environment that particular human subjects may not be, e.g. we will know details of human cognitive capacities and, when investigating past behavior, we will be aware of the future course of events.

Before presenting the explanation itself, we must first consider a number of topics in more detail: the nature and importance of the task, strategies for pursuing an explanation of epistemic practices, "folk" epistemology, and the content of human epistemic practices.

### *6.2 The Nature and Importance of the Project of Explaining Epistemic Practices*

We should first be more precise about just what project we are considering, for there are many sorts of practices that might be counted as epistemic and many aspects of epistemic practices that might be explained. As the term is used here, "epistemic practices" refer to practices peculiar to the production, maintenance, and evaluation of a society's body of knowledge. So construed, three of the chief epistemic practices are the practices of classifying beliefs as knowings or non-knowings, justifying and requesting justification for a belief, and engaging in inquiry.

As was indicated above, the aspects of epistemic practices to be explained here are their existence and function. By an "explanation of the existence of epistemic practices" is meant an account of how and why they have come to exist in human communities generally. That is, an account of the historical events leading to the development of such practices in a particular human community is not, of course,

what is meant here; but, rather, an account capable of explaining what general features of human communities have led to the development of such practices in nearly every extant human community. By "an explanation of the function of epistemic practices" is meant an account of their chief function, presumably a beneficial one, relative to human affairs. It is conceivable, though it would seem unlikely, that there is no such function, e.g. because, as with astrological practices, the practices are based on false assumptions and the chief function is not real. (Practices may, of course, have secondary functions even if there is no chief function, e.g. astrological practices may entertain people. But these secondary functions are not of concern here.)

This pairing of aspects to be explained is not, of course, random. The explanation of the function of a practice is often closely related to the explanation of its existence. For example, the function of placing broken limbs in plaster casts is to speed their healing and insure that they are in proper alignment. Though the explanation of the function of this practice does not explain exactly how it came to exist, it clearly goes a long way toward explaining why the practice exists. However, there are clearly cases in which the explanation of these two aspects of a practice are not closely related. For example, the practice among Catholics of avoiding the use of contraceptives has the function of contributing to overpopulation in some predominantly Catholic countries,<sup>34</sup> but clearly this function does little to explain the existence

---

<sup>34</sup> Actually, this is a secondary function of the practice, but it serves for illustratory purposes.

of the practice. We shall see below how the explanation of the function and the existence of epistemic practices are related.

Now that the explanatory project to be undertaken has been clarified, let us consider the reasons to undertake it. One reason is just that it would be nice to know about the origin and function of a practice so fundamental to human culture. Another reason is a reason largely for neo-Humean epistemologists. The explanation of the existence and function of epistemic practices is a rather straightforward affair for traditional epistemologies.<sup>35</sup> Traditional epistemologies identify knowledge with some sort of rationally privileged belief. Given this identification, it seems clear that the function of ordinary epistemic practices can be explained as the production, evaluation, etc. of such privileged beliefs; and the existence of such practices can be explained in terms of human beings recognizing the importance of such beliefs and developing practices to promote, evaluate, and identify them. However, as was noted above, because neo-Humean epistemologies abandon rational privileging, there is no such straightforward explanation for them to employ. And this lack creates a second reason for undertaking an explanation of epistemic practices: a neo-Humean epistemology that fails to provide an explanation of the existence of epistemic practices leaves itself open to an attack on grounds of explanatory inadequacy. This point can be illustrated by considering a recent neo-Humean approach to epistemology. This

---

<sup>35</sup> The straightforwardness of the explanation helps to explain why traditional epistemologists give little attention to the issue.

approach is called by Richard Rorty "epistemological behaviorism", and he characterizes it as the "method" of:

Explaining rationality and epistemic authority by reference to what society lets us say, rather than the latter by the former.... (Rorty:1979, p.174)

This approach lends itself to an epistemology according to which justification is merely a matter of having beliefs which are supported in socially approved ways. As it stands, such a deflationary epistemology, at least in its more sophisticated forms, is understandably appealing to those seeking alternatives to traditional epistemology. However, as it stands, it leaves us totally in the dark concerning an explanation for the existence of epistemic practices. Why should society bother to approve certain ways of supporting beliefs at all? Why not leave individuals to formulate beliefs however they choose? Such a deflationary approach gives no clue as to how such questions are to be answered. Indeed, it makes epistemic practices seem so trivial that one begins to wonder, if that epistemology were correct, why any society would develop them. Since not only some but all human societies have developed epistemic practices, this epistemology would seem to be explanatorily inadequate.

A third reason for undertaking the project of explaining epistemic practices is that it could provide the basis for arguments against certain epistemologies and epistemological attitudes. For example, suppose that the only adequate explanations are those that attribute to certain common epistemic practices a high utility. Such a result could be used as ammunition against epistemic nihilists, anarchists, and relativists. It provides ammunition against nihilists because it indicates that epistemic

practices have utility, against anarchists because it indicates that epistemic freedom should be restricted, and against relativists because it indicates that one society's epistemic practices may be superior to another society's. Such arguments would be of special relevance to the dialectic of contemporary philosophy since, on the one hand, deflationary accounts of epistemic practices are much in vogue, and, on the other, traditional epistemology seems to have few effective arguments against them.<sup>36</sup>

In addition to these three general reasons for pursuing such an explanation, there is also the aforementioned reason concerning support for bilevel reliabilism. However, the reader should be aware that the support provided by pursuing the explanation will not involve arguing that the bilevel reliabilist explanation is the best explanation. For undertaking such an argument would be beyond the scope of a philosophical work since it would involve gathering empirical data concerning the precise nature of actual epistemic practices as well as intensive evaluation of alternative explanations. Instead, the support for bilevel reliabilism will involve seeking a plausible bilevel reliabilist explanation. Plausibility is an appropriate goal for a project undertaken from the philosophical armchair, and is also a prerequisite for determining that further research on a theory is worth pursuing. However, it might be wondered whether merely finding a plausible explanation will provide much support for a radical theory such as bilevel reliabilism. In this case it should, for, to my knowledge, there are no plausible explanations of the existence and function of

---

<sup>36</sup> Of course, this reason as a motivating reason for undertaking the explanation only if one opposes such extremist views.



epistemic practices extant that are consistent with neo-Humeanism; thus, any theory that can produce one should command respect.

### *6.3 A Strategy for a Bilevel Reliabilist Explanation of Epistemic Practices*

One strategy for explaining epistemic practices was illustrated above in sketching a traditional epistemological explanation of epistemic practices. This strategy attempts to explain the function and existence of ordinary epistemic practices by showing that they approximate some ideal set of useful practices. Let us briefly consider how such a strategy could be applied by a traditional epistemology whose justificatory practices rationally privilege a belief by giving it a Cartesian justificatory foundation and also thereby entitle the belief to the honorific "knowledge". First, it would have to be argued that though we are certainly not perfectly Cartesian in our epistemic affairs, our epistemic conduct does approximate ideal Cartesian practice. Since Cartesian practices, the argument would continue, are highly beneficial (they help us get to the truth, etc.), practices that approximate them will tend to share in these benefits. Finally, it would be argued that ordinary practices do share in these benefits, and this is their function. This function, in turn, explains their existence; for, it would be claimed, humans began employing them because they recognized they could perform this useful function.

Henceforth, we will term this explanatory strategy, in which the function and existence of epistemic practices is explained by their approximation to a useful set of ideal practices, the approximation strategy. An important advantage of this strategy is

its reliance on a utility beneficial to human societies generally. This feature enables the approximation strategy to avoid a difficulty of some other approaches. For example, a historical approach might explain the existence of epistemic practices in a society by appealing to the need for a priestly class to maintain conformity of opinion after a sudden introduction of foreign ideas. However, it is clear that such an explanation is superficial at best, for it faces the immediate objection that since epistemic practices are ubiquitous (i.e. they exist in all human societies), a genuine explanation of their existence should appeal to features common to all societies. Since the approximation strategy relies upon a utility beneficial to human societies generally to explain the existence of epistemic practices, it meets this requirement.

Though the approximation strategy has this advantage, it would seem an inappropriate strategy for a bilevel reliabilist explanation. The reason is that, as will be shown presently, counterexamples indicate that the notions of br-knowledge and br-justification are not even approximations of their everyday counterparts. And if bilevel reliabilism and everyday epistemic views differ dramatically on such a fundamental point, it is difficult to see how ordinary practices could plausibly be claimed to approximate an ideal set of bilevel reliabilist practices.

A counterexample indicating the difference between the notion of br-knowledge and the ordinary notion of knowledge can be constructed similarly to the Bonjour counterexamples to externalist theories of knowledge simply by transporting the "irrationality" of the belief to the second level. For example, suppose Lisa believes truly that Gorbachev is in Leningrad, and that she has acquired this belief via

the reliable process of reading the newspaper. Suppose further that she r-recognizes the authority of her belief: she r-knows via a reliable process of introspection that she acquired the belief by reading the newspaper, and she r-knows that believing what one reads in newspapers on such matters is a reliable belief acquisition process. In such a case Lisa's belief fulfills all the requirements for br-knowledge. But suppose that her r-knowledge that newspapers are reliable on such matters was acquired via a reliable clairvoyant power of which she is unaware, and suppose that she has no grounds for believing that newspapers are reliable sources of information on such matters. In such a case it would be difficult to attribute knowledge, in the everyday sense, to Lisa. For if she were pressed to defend her belief she would soon have to admit she had no basis for believing newspapers were reliable, and her justification for her belief that Gorbachev was in Leningrad would seem to vanish. Similar considerations make it difficult to attribute justification, in the everyday sense, to Lisa even though her belief is br-justified.

This divergence between the notions of br-knowledge and br-justification and ordinary notions indicates that bilevel reliabilism would do best to employ some explanatory strategy other than the approximation strategy. But it might be objected that the divergence indicates an even more serious difficulty: that radical epistemologies generally are inappropriate for the task of explaining ordinary epistemic practices since a conservative epistemology<sup>37</sup> (traditional or otherwise) would, because of the

---

<sup>37</sup> The term "conservative epistemology" is explained in chapter 2.

similarity between its epistemic notions and ordinary ones, always seem capable of a simpler explanation via the approximation strategy.

However, I would argue that the opposite is true: that it is conservative epistemologies that are inappropriate for explaining epistemic practices. The reason is that, as will be discussed below, ordinary notions of knowledge and justification are fairly close to Cartesian--in a broad sense of the term--notions of knowledge and justification. And, like the corresponding Cartesian notions, knowledge and justification in the ordinary sense simply cannot exist; hence, the accounts of knowledge and justification of a successful conservative epistemology, because it aims to explicate ordinary epistemic notions, should also refer to things that cannot exist. Though this correspondence between the two sets of notions may allow a conservative epistemology to make a simpler explanation of epistemic practices, clearly an epistemology dealing with non-existent entities will have great difficulty producing a satisfying explanation of epistemic practices. (Thus, for these purposes, one might even argue that the lack of counterexamples to an epistemology constitutes an objection to it.)

Regardless of whether the claim that radical epistemologies are superior for such explanatory purposes is correct, the argument to support the claim does suggest a framework in which a radical epistemology such as bilevel reliabilism could play a role in an explanation of epistemic practices. But before showing how it could a pair of distinctions need to be drawn. First, let us distinguish everyday beliefs about what knowledge and justification are and what the purport of epistemic practices is, and

everyday epistemic practices themselves. The former (the set of beliefs) constitutes a kind of epistemology. Let us term it, adapting from the philosophy of psychology, folk epistemology. Second, in the philosophy of psychology it is common to distinguish folk psychology from scientific psychology; similarly, we can distinguish folk epistemology from what will be termed philosophical epistemology, the epistemological views of experts such as professional philosophers.

Using these distinctions, the difficulty concerning the use of a radical epistemology to explain epistemic practices can be restated as follows: How can a philosophical epistemology that diverges from folk epistemology successfully explain epistemic practices? Borrowing a bit more from the philosophy of psychology will enable us to see how to answer this question positively.

The issue of the extent to which folk and scientific psychology are compatible and the question of the import of the theoretical terms of folk psychology ('belief', 'desire', etc.) are much debated. Positions on this issue range from the view that scientific psychology is entirely compatible with folk psychology and will provide a systematization and theoretical underpinning for it, to the view that they are radically incompatible and that scientific psychology will not only show that folk psychology is largely false but also that the very ontology of folk psychology (i.e. the hypostasization of beliefs, desires, etc.) should be abandoned.<sup>38</sup>

It should now be apparent how to answer the above question positively. A philosophical epistemology divergent from folk epistemology could explain epistemic

---

<sup>38</sup> For more on these issues in the philosophy of psychology see, e.g., Stich (1983).

practices by standing to folk epistemology and its correlative epistemic practices in one of the range of ways that scientific psychology stands to folk psychology. That is, though philosophical epistemology might employ very different notions from those used in folk epistemology and epistemic practices, it nevertheless might be able to provide a systematization and theoretical underpinning for folk epistemology and epistemic practices, as in the case of compatibilism in the philosophy of psychology, or it might be able to explain the phenomena addressed by folk epistemology even though folk epistemology turned out to be false, as in the case of incompatibilism in the philosophy of psychology. In either case, a radical philosophical epistemology would be able to explain epistemic practices unhindered by its divergence from folk epistemology on key epistemic notions.

We have been discussing in general terms how a radical epistemology could successfully explain epistemic practices. Now we should consider what specific strategy to adopt for a bilevel reliabilist explanation. As the remarks above concerning the falsity of folk epistemology suggest, the strategy to be adopted is close to the radical incompatibilist approach. We will attempt to explain the function of epistemic practices by showing that though folk epistemology and bilevel reliabilism diverge radically, and so the utility of epistemic practices cannot be explained by the approximation of folk epistemology to bilevel reliabilism; nevertheless, epistemic practices promote both *br*-justification and *br*-knowledge (despite the fact that folk epistemology does not view them as having this function), and it is the utility achieved by the

promotion of these useful things that is the beneficial function of epistemic practices.<sup>39</sup> It might be thought that the explanation of the existence of epistemic practices would then follow as a matter of course from the explanation of their function by claiming that, as with the approximation explanation employed by traditional epistemology, epistemic practices have come to be because people recognized their utility. But this route is blocked in the present strategy because the falsity of folk epistemology means that humans do not correctly understand the utility of epistemic practices: they believe the practices are useful but for the wrong reasons and in the wrong ways. Accordingly, the bilevel reliabilist explanation of the existence of epistemic practices is one that is independent of their actual utility; instead, the existence of epistemic practices will be explained as a result of natural human inductive tendencies and erroneous beliefs in the utility of epistemic practices.

There are three major steps to carrying out this explanatory strategy. First, we must describe folk epistemology and actual epistemic practices. This descriptive step will occupy us for the remainder of this chapter. Second, we must explain the origin of these practices by showing how natural inductive tendencies have given rise to them. Third, we must explain the function of these practices by showing that they promote *br*-knowledge and are useful because of it. These latter explanatory steps will occupy us in the next chapter.

---

<sup>39</sup> Thus, in a sense it might be said that the approximation strategy will be applied indirectly because, though folk epistemology and bilevel reliabilism are incompatible, they end up recommending approximately the same epistemic practices.

#### 6.4 Human Folk Epistemology

Before considering the content of human folk epistemology, a few remarks about folk epistemology in general are in order. Above, folk epistemology was defined as the everyday view of knowledge, justification, and the purport of epistemic practices. When we speak of folk epistemology or folk psychology, we mean of course, the folk epistemology or psychology typical of humans. However, though this implicit restriction recognizes the possibility of non-human folk theories different from ours, this way of speaking does seem to imply a singularity of human folk theories. But surely this is wrong. Surely folk psychological and epistemological views have not remained static through human history nor have they been undifferentiated across the globe. Nevertheless, I want to retain this manner of speaking because it is a useful idealization. It is useful because, despite there being a plurality of human folk epistemologies and folk psychologies, there seems to be a common core of beliefs to which there are few exceptions (at least if we discount prehistory), and it seems appropriate to call this core simply folk epistemology or folk psychology. The description of folk epistemology below is a rough approximation of this core of beliefs.

The sense of 'everyday' employed in the definition should also be specified. By 'everyday' is meant the uncritical beliefs people have about what epistemic practices amount to, i.e. what they believe about them when they are not engaged in the philosophical task of considering what the best account of them is. Hence, 'everyday' view could also be rendered as 'working' view.



How does one determine the content of human folk epistemology? Prima facie it might seem a straightforwardly empirical task--one simply interviews people concerning their epistemological beliefs. But such a method of inquiry would be unsuitable for the same reasons it would be unsuitable in linguistics to use such methods to determine the meaning of a word or the grammar of a language--people are awkward at expressing themselves about such matters and often make mistakes. Alternatively, one might suggest that looking at epistemic practices in action would be a better method since that would eliminate difficulties with limitations of eloquence. But the problem of mistakes would remain, and the difficulty of the gross underdetermination of folk epistemology by actual epistemic practice would render such an approach of limited value. A more helpful approach to determine the content of folk epistemology is to look at the behavior and pronouncements of official bodies of inquiry, e.g. courts, congressional committees, etc. Their behavior is an excellent source of information concerning folk epistemology, for such bodies are supposed to have as their goal the impartial pursuit of truth, and they can function in society (at least democratic ones) only so long as they are seen as successful at this task; and they will be seen as successful only so long as their procedures are in general conformance with a society's general beliefs about proper epistemic practice, i.e. its folk epistemology.

Let us consider, then, the typical behavior and pronouncements of such bodies. Often they have the stated goal of "proving beyond a reasonable doubt" an answer to a question posed. The usual method to achieve this goal is to gather evidence and

then draw conclusions based upon that evidence. In both the evidence-gathering and conclusion-drawing stages of this method, rhetoric concerning the intent to deal only with "facts" and not "speculation" is often exhibited. Ideally, the outcome of the process is the "proof", in the sense of the provision of conclusive evidence, of some answer to the question posed, e.g. the prosecutor will have proved that the defendant is guilty or a congressional committee will have proved that the president had full knowledge of the illegal activities of certain members of his administration. Of course, the proceedings are not always held to have succeeded in proving anything conclusively, but it is not uncommon that they are.

The investigative procedure just sketched will probably sound banal and commonsensical to the philosophically untrained ear, and this is just as it should be when we are hunting for folk epistemology. But to the trained ear these pronouncements should reflect an epistemology bearing some similarity to Cartesian foundationalism, and which repeats some of its errors. First of all, there is implicit the Cartesian assumption of a foundation for knowledge, a class of "facts" with which we start the investigation, and which are "given" in the sense that we need not infer them from other evidence and that they come to us from the world and are beyond dispute. Such an assumption is the point of Jack Webb's famous refrain on the television show Dragnet, "Just the facts ma'am". The error in this assumption is, of course, thinking that the world gives us any such set of facts that are beyond dispute, that there is any special class of privileged beliefs which need not be critically examined and need no grounding beyond themselves. A second feature of the epistemology reflected in such

pronouncements is the view that we can obtain from such an investigation proofs of a claim, e.g. that the defendant is guilty of murder. I encountered examples of this view recently: a sports commentator spoke of waiting for the courts to "prove one way or the other" whether Pete Rose bet on baseball games, and Peter Jennings of ABC news spoke of there at last being "scientific proof" that aspirin can reduce the risk of heart attack.<sup>40</sup> This belief in the possibility of proof, in the sense of providing conclusive evidence, is also a feature of Cartesianism, and the error is similar also. The error is in supposing anything like conclusive evidence can be provided in such contexts.<sup>41</sup>

Above, folk epistemology was characterized as having three elements: views about the purport of epistemic practices, views of knowledge, and views of justification. So far we have addressed only the first element: Folk epistemology holds that epistemic practices (at their best, as in the case of investigative bodies) consist of gathering facts to prove claims. The rest of folk epistemology is more difficult to

---

<sup>40</sup> It might be objected that talk of "proof" in science should not be treated as tantamount to talk of conclusive proof because "proof" in this context just means "strong evidence". This synonymy claim is clearly wrong, for even in ordinary parlance one can have strong scientific evidence for  $p$  even though  $p$  is false, but one cannot have scientific proof of  $p$  when  $p$  is false.

<sup>41</sup> It might be objected that American civil court proceedings, which require only a preponderance of evidence for a favorable ruling, provide evidence against the claim that folk epistemology insists upon the possibility of proof. However, this fact does not tell against the claim that has been made for two reasons. First, the claim is not that folk epistemology recognizes the possibility of proof in every instance but only in a significant number of them. Second, the absence of a commitment to obtain proof in civil rather than criminal proceedings can be explained on practical rather than epistemological grounds. It is held as a great sin in criminal cases to convict an innocent man; thus, for the prosecution to obtain a conviction not merely strong evidence but proof is required that the defendant is guilty. In civil cases an erroneous verdict has less undesirable consequences; hence, there is no requirement of proof to guard against such verdicts.

determine, for only philosophers consider what knowledge and justification are. However, we can probably infer with a fair degree of accuracy such views from the partial folk epistemology we have so far uncovered. Knowledge, for folk epistemology, is just a true belief that one proves (or is able to prove) in the manner just described. Thus, if Rose is convicted of betting on baseball games we are entitled to claim that we know he did. Knowledge-sufficient justification is just possessing the proof of such a claim. But in addition to a belief in knowledge-sufficient justification, folk epistemology also seems to include a belief in justification insufficient for knowledge and having little to do with "proof". An example of such fallible justification might occur in the case of a congressional committee formed to assess the environmental impact of the Alaskan oil pipeline. In such a case evidence is gathered to support various claims. It is commonly thought that this evidence does not prove any of the claims--no one can be sure just what will happen when so many variables are at play, but the inquiry may give us a best guess--and fallible justification--for a claim as to what will happen. The belief in such justification raises two issues concerning folk epistemology--the folk view of the purport of such justification and the view of its exact nature. The view of its purport seems plain: it increases the likelihood of being right. Clearly, such a view is erroneous if, as we have been assuming throughout our discussion, Hume's attack on induction is successful. I don't believe there are any developed folk epistemological views concerning just what counts as justification of this type--such justification is usually evaluated on a case by case basis.

We have sketched accounts of the chief elements of folk epistemology and made clear the ways in which they are erroneous. Much more could be said about folk epistemology, but this brief sketch suffices for present purposes. For, as with the bilevel reliabilist explanation of epistemic practices itself, no claims are being made about the definitiveness of this account of folk epistemology; the chief aim is to provide a plausible account. However, to increase its plausibility, let us consider a few objections.

First, it might be objected that I have unjustly incorporated one of the least plausible aspects of Cartesian-style epistemology into folk epistemology: the requirement of conclusive evidence and, hence, epistemic certainty, for knowledge. For clearly, the objection might continue, considerations of evil geniuses and whether we can tell when we are dreaming are too far-fetched to be of folk epistemological concern, but they ought to be of concern if folk epistemology required certainty for knowledge.

There are several points to be made in reply to this objection. First, the fact that skeptical problems related to evil geniuses, etc., are ignored in day-to-day affairs can be adequately explained while maintaining the present account of folk epistemology. It may simply be a question of discounting fanciful threats to certainty in day-to-day affairs, while equally remote but less fanciful threats are taken seriously. An example of the latter would be that one would not be said to know that a recently purchased lottery ticket would not win the lottery even though the chances of winning were one in thirty million. Second, the usual reason for arguing that folk epistemolo-

gy does not require certainty for knowledge, is that it will transform folk epistemology from a false to a true theory. However, even changing the present account of folk epistemology in this way will leave a view that asserts the possibility of rational privileging, and so it will still be erroneous. Thus, unless there is clear evidence to the contrary, there would seem to be no reason to avoid what seems a natural interpretation of folk epistemology and retain the certainty requirement.<sup>42</sup>

J.L. Mackie has advanced an "error theory" concerning the nature of ethical statements:

...the denial of objective values will have to be put forward...as an 'error theory', a theory that although most people in making moral judgments implicitly claim...to be pointing to something objectively prescriptive, these claims are all false.  
(1977, p.35)

Similarly, the present account of folk epistemology might be viewed as an error theory concerning the nature of epistemic claims. Two sorts of objections that have been raised against Mackie's error theory might also be raised against this epistemological error theory. First, it might be objected that this error theory saddles the non-philosophical public with the errors of philosophers, and that is unfair because such persons do not have philosophical views on these matters. Second, it might be noted that throughout this work, in addition to making various claims about knowledge, I have also been claiming to know various things. Presumably, I do not believe these claims are false, but the error theory I propound implies that they are.

---

<sup>42</sup> Another reason to retain this requirement is that it immunizes everyday knowledge from the Gettier problem. (See Dretske (1971) for a discussion of why a certainty requirement provides immunity to the Gettier problem.)

The first objection would seem to stand or fall on the plausibility of two claims: (1) that the folk epistemology just presented attributes to ordinary people epistemic beliefs at a level of sophistication that their beliefs about this topic do not reach and (2) that it is bad hermeneutics to interpret a practice as thoroughly erroneous unless there is very good reason to do so and that no very good reason has been given in the present case. I suppose that much of the plausibility of claim (1) rests on the suspicion that, for example, a Texas ranch hand would have little to say if asked his views on the purport of epistemic practices or the nature of knowledge and justification. I will freely grant that such persons would respond in this fashion, but this fact alone does little to support claim (1); for all it demonstrates is an inability to articulate such beliefs, not the lack of them. Moreover, it should be noted that this folk epistemology, as was noted above, does not attribute to ordinary people theories about the nature of knowledge and justification; it only states that they are implied by their views about the purport of epistemic practices. The crucial issue in determining the plausibility of (1), then, would seem to be whether ordinary people possess (at least) inarticulate views of the sort included in this folk epistemology about the purport of epistemic practices. There are several reasons to believe they do. One reason is that such persons engage in quite sophisticated epistemic behavior, e.g. a ranch hand might evaluate the soundness of another's reasons for believing a prize steer has foot-and-mouth disease; hence, it does not seem implausible to suppose they have equally sophisticated inarticulate beliefs about the nature of epistemic practices. Another reason is provided by the typical response of students in introduction to

philosophy courses to Cartesian and Humean skeptical challenges. If they had no views on such matters, one would expect such students to react passively or with mild interest. But, of course, they typically react with shock or amusement that anyone could hold such views, and, when challenged, try valiantly to overcome them. What more natural way to explain such a reaction than that these challenges contradict views these students already have? There would seem, then, to be good reasons to believe ordinary persons possess at least inarticulate sophisticated views about epistemic practices; hence, claim (1) would seem to be on shaky ground.<sup>43</sup> Claim (2) is likewise on shaky ground. First, because there are numerous cases in which whole categories of human practices have come to be seen as thoroughly erroneous, e.g. magical, religious, astrological, alchemic, etc. Second, because there is good reason for supposing humans might arrive at erroneous beliefs about epistemic practices: men as brilliant as Descartes, Locke, Berkeley, etc. have arrived at systematically erroneous views about the powers of epistemic practices, so why couldn't the ordinary person? Third, because there are good reasons for supposing ordinary views about epistemic practices are erroneous--e.g. the pronouncements of official bodies of inquiry, television commentators, etc.--and these have already been discussed above.

My reply to the second of the error theory objections, concerning my own use of "know", is just that to say that I do not intend my tokenings of "know" to be

---

<sup>43</sup> The reader who remains skeptical of the claim that ordinary persons have even inarticulate sophisticated views about such matters might want to consider that all that is really needed to establish the legitimacy of this folk epistemology is the plausibility of the (in some respects) weaker claim that erroneous views of this sort be implied by the practices people engage in.



interpreted in the ordinary sense. I am uncertain exactly what interpretation they ought to have, but, for present purposes, they can be safely treated as at least being claims to know in the deflationist sense specified in the next section. (The reader may be surprised that I choose not to have my claims to know interpreted as claims to br-know. The reason for this choice is that br-knowledge seems too complex and theoretical a notion to serve smoothly in everyday contexts.)

### *6.5 The Content of Ordinary Epistemic Practices*

We now have a description of folk epistemology, but knowing what people believe about epistemic practices doesn't necessarily give us a full or accurate picture of the epistemic practices they engage in; and, in the present case, folk epistemology gives a highly inaccurate picture since it implies, for example, that knowledge is attributed to a kind of belief which cannot exist. A more accurate description will be attempted here, but, to keep this descriptive project as well as the explanation as a whole manageable, our attentions here and in the rest of the explanation will be restricted to the epistemic practice of knowledge attribution. By "knowledge attribution" is meant the practice of classifying beliefs as knowings or non-knowings.

A description of the practice of knowledge attribution involves specifying what beliefs we classify as knowings. However, for present explanatory purposes, the description ought to be applicable to the knowledge attribution practice of most any society; thus, though a detailed description is needed, we should not be so specific that features peculiar to a society are included. Also, a description is needed that will

not beg the very epistemological and explanatory questions at issue by being couched in theory-laden terms, e.g. a description that claimed "knowledge is attributed to br-knowledge" would not be very helpful. The basis for a description meeting these requirements can be obtained from the deflationist epistemology Sosa (1983) terms "conventionalist foundationalism" and ascribes to Richard Rorty. Roughly, conventionalist foundationalism holds that a claim is justified when it can be shown to follow, according to socially approved inferential and argumentative norms, from the class of assertions which receive direct social approval, i.e. assertions and kinds of assertions which society deems it ridiculous and irrational to question, e.g. first-person reports of psychological states, "France exists", etc. The deflationary component of this view is the claim that this is all there is to epistemic practices--they are just a matter of convention.

Many of the assumptions underlying conventionalist foundationalism seem to be clearly true. Society does sanction certain beliefs, classes of beliefs, and inferential norms, as ridiculous to question. And we do come to know which of these are sanctioned as part of the process of socialization. For example, every child learns that he can make deceitful use of the special authority granted to first-person pain reports to avoid another painful day of primary education. The controversial claim of conventionalist foundationalism is that societal sanction is just a matter of convention. But, if we excise this claim from conventionalist foundationalism the result will be the sort of description of knowledge attribution we are seeking:

The title of knowledge is conferred on just those beliefs which receive direct social sanction or which the believer can infer by socially sanctioned norms from claims which receive direct social sanction.

Although somewhat rough, such a description--henceforth the "conventionalist foundationalist description"-- meets the requirements set forth above. For it applies to human societies generally, and it is couched in terms that do not beg the epistemological and explanatory questions that are about to be addressed. To see that it does not beg such questions, consider that such a description would be applicable to a society that employed the knowledge attribution practice recommended by an epistemology as different from bilevel reliabilism as Cartesian foundationalism. All that would be required for the description to be true of a Cartesian society would be that the society sanction, for knowledge attribution purposes, the basic beliefs and inferential norms recommended by Cartesian foundationalism.

Our explanandum, the epistemic practice of knowledge attribution, thus isolated, we can proceed to the explanation.

## 7

### **A BILEVEL RELIABILIST EXPLANATION OF EPISTEMIC PRACTICES**

We are now ready to begin constructing the bilevel reliabilist explanation of the existence and function of epistemic practices, or, to be more precise, practice, for the explanation will be limited to the practice of knowledge attribution. The details of this explanation will be generated in reference to a hypothetical human society, the Gawanans, that will be used as a paradigm human society. The first step in developing the explanation will be to explain how a practice of knowledge attribution might originate among the Gawanans. The second step will be to explain the function of their knowledge attribution practice in terms of its promotion of br-knowledge. The final step will be to consider to what extent the explanation of the Gawanans practice can be used to explain the knowledge attribution practice of human societies in general.

Throughout the presentation of the explanation, the reader should bear in mind that the aim is only to present a plausible explanation--not necessarily the best--so no arguments that it is best will be attempted.

### *7.1 The Origin of the Practice of Knowledge Attribution*

The Gawanans are a prehistoric community of hunter-gatherers with a moderately rich spoken language consisting of declarative sentences, commands, questions, expressions of emotion, but no knowledge attribution practice. Gawanan development of a knowledge attribution practice might begin when, in their daily affairs, they come to believe certain things about their own beliefs and those of their associates. For example, when Ogg comes to believe that there is a saber-toothed tiger in the area and reports his belief to the others, the rest of the group usually comes to believe that he is correct, but on a few occasions they wind up believing he is wrong. They discover that when they believe he is right, he usually believes (and so do they) that he has directly observed the beast in good viewing conditions. When they believe he is wrong he usually believes he has seen its pawprint, caught its scent, etc. Through many generations such beliefs become the basis for inductions to generalizations about how beliefs acquired in certain ways can be relied upon whereas others cannot; and these inductions, in turn, become the basis for societal sanctioning of belief acquisition processes. (The transition from merely believing that certain processes are trustworthy to sanctioning them, being marked, in part, by the belief-process favoritism becoming part of the cultural heritage Gawanans transmit to their offspring.) Moreover, Gawanans discover that it is often useful to inform other Gawanans of how they acquired their beliefs. For example, they discover that if one Gawanans believes he acquired his belief about the proximity of saber-toothed tigers by one of the means held to be trustworthy, it will aid him in motivating the others to act appropriately if

he informs them how he acquired his belief. Because of such informings, though no word for knowledge has been introduced, it seems fair to say that something qualifying as a practice of knowledge attribution is largely in place, for Gawanans are classifying beliefs as being of a trustworthy kind or not. A word for knowledge would be introduced, perhaps, when the Gawanans begin to tire of having to make detailed explanations of how they acquired a belief whenever they wished to claim it was one of the processes thought to be trustworthy. Such a word could be substituted for the explanation whenever a Gawanans wanted to make such a claim about one of his beliefs.

The above explanation of the origin of knowledge attribution can be summarized as follows. The practice of knowledge attribution arises in Gawanans society because Gawanans inductive abilities, combined with desires to be right about states of affairs, lead them to develop views about the trustworthiness of various belief-forming processes. And these views, in turn, become part of the societal heritage and lead to certain belief-forming processes receiving social sanction. A word for knowledge is then introduced as an abbreviated way of claiming that a Gawanans's belief was acquired via one of the sanctioned processes.

However, this explanation cannot be our final explanation because the practice of knowledge attribution whose existence it explains is not the conventionalist foundationalist practice that is our ultimate explanandum. The practice just explained differs from the conventionalist foundationalist practice in that a "pedigree" rather

than an "evidentialist" view of justification operates implicitly. A pedigree<sup>44</sup> view of justification is that typical of reliabilists, and is distinguished by the relevancy of the causal pedigree of a belief to its being justified. An evidentialist<sup>45</sup> view holds that being justified is a function of the evidence one possesses. Later, we will consider how a conventionalist foundationalist practice might evolve from this pedigree practice, and so complete the explanation of the origin of the present-day practice of knowledge attribution. But for the next few sections, which deal with the function of knowledge attribution, I want to treat the pedigree view as the explanandum. The reason is that since br-justification is of the pedigree type, the details of the bilevel reliabilist explanation of the function of knowledge attribution will be simpler if we first consider the function of the pedigree view.

Before considering the function of knowledge attribution, I want to briefly consider an alternative explanation of the origin of this practice that I suspect many philosophers, especially non-epistemologists, might give if asked. This explanation is that the practice was developed to serve a certain performative linguistic function. The explanation claims with Austin that

To suppose that 'I know' is a descriptive phrase, is only one example of the descriptive fallacy, so common in philosophy. (Austin:1961, p.71)

In the case of avowals of knowledge, the performative function is just that, as Austin says, "when I say 'I know', I give others my word: I give others my authority for

---

<sup>44</sup> This term is borrowed from Levi:1980, p. 28.

<sup>45</sup> This term is borrowed from Feldman and Conee (1985).

saying that 'S is P'" (Ibid., p.67). Thus, roughly, on this view knowledge ascription in general performs the same linguistic function as promising--but with propositions instead of deeds. The explanation of the origin of knowledge attribution to which such a view leads is that the practice was developed to enable us to make such propositional "promises". Such an explanation is clearly at odds with the explanation given above, for according to it the chief linguistic function of knowledge attribution is the descriptive one of communicating the special sort of belief one has, i.e. a belief acquired via a process socially sanctioned as trustworthy.

This explanation is an intriguing alternative, but it does not work, at least as it stands. The reason is that "know" does not have the performative function the Austinian account claims it does. The easiest way to see this is just to note, as does Jonathan Harrison in a discussion of the topic, that:

...someone saying 'I promise...' is thereby promising, whereas someone saying 'I know...' is not thereby knowing, but simply claiming that he knows. (Harrison: 1962, p.449)

For further argument against the Austinian line I refer the reader to Harrison's article.

## *7.2 How Knowledge Ascriptions Could be of Br-Knowledge*

We are now ready to begin developing a bilevel reliabilist explanation of the function of Gawan epistemic practices. As was noted in the previous chapter, the gist of the bilevel reliabilist explanation is that the practice of knowledge attribution functions to promote br-knowledge and is useful because it does. The first step is to



establish the plausibility of three claims about the Gawanans' pedigree knowledge attribution practice:

- (I) That their knowledge ascriptions will, typically, be made of br-knowledge.
- (II) That their practice of knowledge attribution promotes br-knowledge.
- (III) That the promotion of br-knowledge is useful.

Claim (I) is addressed in this section, claims (II) and (III) in the next.

### *7.2.1 Ascribing Knowledge to R-justified Beliefs*

In the Gawanans' knowledge attribution practice, knowledge is ascribed to beliefs acquired by *bfp*'s that have come to receive societal sanction via inductive generalizations concluding that they are trustworthy. Since, by definition, a majority of br-justified beliefs will constitute br-knowledge,<sup>46</sup> one way to show that beliefs to which knowledge is ascribed will typically constitute br-knowledge--and hence that (I) is true--is to show that they will typically be br-justified. A first step toward establishing the latter claim is showing that such beliefs will be r-justified, and that is our concern here.

An obvious way to show that these beliefs will be r-justified is to show that the sanctioned *bfp*'s are reliable. The best way to show this is to show that the inductive

---

<sup>46</sup> The reason that this is true is that all br-justified beliefs are r-justified, and, by definition, a majority of r-justified beliefs are true. And since a br-justified true belief constitutes br-knowledge, majority of br-justified beliefs will be br-knowledge.

(Note: For the sake of simplicity, I have employed some claims in this argument that are not literally true. For example, it is not literally true that "by definition, a majority of r-justified beliefs are true". By definition a belief is r-justified iff it is acquired by a reliable process, and that process is reliable iff it tends to produce more true than false beliefs. But a particular group of r-justified beliefs might be all false if they were products and processes that tended to produce more true than false beliefs but, in the case of this particular group, produced only false beliefs. The above argument (as well as similar claims below) ignores this possibility because it would constitute a statistical freak.)

inferences which largely determine which processes receive sanction are, in some appropriate sense, good ones. Let us consider an inference that might play a role in the sanctioning of a *bf* in Gawan society:

On 10 occasions Ogg has believed he directly observed a saber-toothed tiger in the vicinity.

On these 10 occasions there was a saber-toothed tiger in the vicinity.

Therefore, usually when Ogg believes he directly observes a saber-toothed tiger in the vicinity, there is one in the vicinity.

Inductions such as this might lead to the sanctioning of direct observation as a trustworthy *bf* not only for Ogg's observances of saber-toothed tigers but also for observances of events, medium-sized physical objects, etc., and eventually for the sanctioning of direct observation as a trustworthy *bf* for all normal Gawanans (i.e. sane, with good eyesight, etc.) What would have to be the case for such inferences to be an effective means of sanctioning reliable *bf*'s? An obvious answer is that the premisses would have to be true and the regularities the inferences project would have to be realized. Let us see if there is reason to suppose these two elements obtain for such inductions.

Reason to suppose the first element obtains can be garnered from the fact that Gawanans, being humans, are the product of a long evolutionary history. Since having reliable cognitive machinery has obvious survival value, it seems plausible to suppose that evolution would reinforce whatever hereditary tendencies humans might have in this direction. Moreover, it seems plausible that cultural pressures would, ceteris paribus, tend to preserve and reward reliable belief-forming methods. These

two factors might combine to give the Gawanans belief processes reliable enough to produce sufficient true premisses for successful inductions concerning the reliability of belief-forming processes. (Though these reasons are not as strong as one would like given the breadth of the claims we are considering (e.g. that the Gawanans possess generally reliable cognitive processes), I see no way of strengthening the reasons or narrowing the claims and still getting the epistemic ball rolling for the Gawanans--they must start out being largely right about things before they can increase their accuracy by employing epistemic practices.)<sup>47</sup>

Providing an argument that the second element, the realization of the regularities that the Gawanans' sanctioning inductions project, obtains is not so simple. One might even suspect that providing such an argument will require a Herculean effort since it might seem to require constructing a robust theory of projection that will circumvent Goodman's paradox and other difficulties. However, fortunately, our task is not so daunting. Let us consider the nature of a theory of projection more closely.

Most directly, a theory of projection is required for the construction of an adequate system of inductive logic, for in addition to inference rules such a logic

---

<sup>47</sup> Richard Boyd, in his attempt to support scientific realism by arguing that the best explanation for the "instrumental reliability" of scientific methodology is that the theories associated with it are "approximately true" (see e.g. Boyd 1984), has noted a similar difficulty that he calls the "start-up problem" (mentioned in a lecture c. 1988). The start-up problem, if I recall correctly, is the problem that since scientific methodology is theory-laden, one cannot use scientific methods to get true theories unless one's present theories are already largely true; for if one's present theories are false then the methodology employed will be unsuccessful because it will be laden with false theories. Thus, according to Boyd, the engine of (supposed) scientific progress cannot be started unless it is first infused with a number of non-scientifically acquired true theories. Since Boyd's project is also a sort of explanation of epistemic practices, this start-up problem may be one faced by most such explanations.

would require rules dictating which predicates were or were not projectible.<sup>48</sup> Secondly, a theory of projection could play a role in the t-justification<sup>49</sup> of induction, since it might help to show that our choice of predicates to project is not epistemically arbitrary. I do not believe there is a theory of projection that can succeed at either of these projects (at least in any non-trivial way) and, of more present relevance, we do not need to complete such projects to solve our present problem. For that problem concerns showing that inductions are effective, and they may be effective even if they are totally devoid of logical, rational, or epistemic privilege; for it is easy to imagine possible worlds in which inductions are effective regardless of their rationality. For example, a world in which the conclusions of all inductions made are true "by accident". (Though, of course, it might be objected that induction is not really effective in such a world since "by accident" implies that counterfactual applications of induction would not be successful.) To see how to construct an argument, without recourse to logic or rationality, that the regularities Gawanans' project inductively are realized, let us turn to some of the ideas of W. V. Quine.

In "Natural Kinds" Quine claims:

To trust induction as a way of access to the truths of nature...is to suppose...that our quality space matches that of the cosmos. (1969, p.125)

---

<sup>48</sup> The classic source for arguments for the last claim is, of course, Nelson Goodman's Fact, Fiction, and Forecast.

<sup>49</sup> T-justification, it will be recalled, is traditional justification in the sense discussed in chapters 1 and 4.

This remark provides the basis for the argument we will pursue: the Gawanans' inductive projections are realized because "our [i.e. human beings, including the Gawanans] quality space matches that of the cosmos". To present this argument we must first unpack some of the obscurer notions it employs: the notion of a quality space, what it would mean for our quality space to match that of the cosmos, and why a match would promote inductive success. I take up these points in order.

Quine borrows the notion of a quality space from psychology. Our quality space (also called similarity space) consists of a ranking of things as similar to each other or not as reflected by our dispositions to judge them as similar. Expanding metaphorically on the spatial reference in "quality space", we can imagine this ranking as represented in a three-dimensional space occupied by object types, tokens, and properties (e.g. shades of colors) where there is a direct correlation between the proximity of entities and the degree of similarity between them. Some features of a quality space may be innate. For example, Quine cites animal experiments in which:

A response to a red circle, if it is rewarded, will be elicited again by a pink ellipse more readily than by a blue triangle.... (1969, p.123)

It may well be that such a response is, at least partly, genetically programmed. But even apparently innate quality spacings can be altered. For example, pre-theoretically we would classify a whale as more similar to a shark than a seal, but post-theoretically the reverse is true. The relevance of quality spacings to induction is that quality spacings would seem to largely determine what inductive projections we make. For example, if we know little of insects and are stung by a yellow jacket, we may subsequently judge a honey bee to be likely to sting on account of its similarity to the

yellow jacket but may not judge a butterfly as relevantly similar. Thus, since, in our quality space, we judge yellow jackets as more similar to honey bees than butterflies, we project propensity to sting for honey bees but not butterflies.

Let us now consider what it might mean for our quality space to match that of the cosmos. Our quality space, it might be suggested, matches that of the cosmos just when the things we judge to be similar are similar, e.g. as they are when we judge one lump of gold more similar to another lump of gold than to a lump of iron pyrite. But, unfortunately, this straightforward account of matching will not do. To see why, it must first be noticed that quality spaces concern generic similarity judgments, i.e. similarity judgments not made relative to respects. But, for a number of reasons, it is notoriously difficult to cash out such generic similarity judgments in objective terms (cf. Quine (1969) ch. 5). For example, in the last example, it seems to be taken for granted that chemical composition will determine which pair of metal lumps is most similar, but, clearly, other factors should also be considered, e.g. their shape, size, etc.--and it is unclear just how much weight these other factors should be assigned. It might be thought that each could be assigned equal weight and the question of similarity could be decided in terms of number of shared qualities, but the hopelessness of this proposal is easily seen once it is realized that the number of shared qualities (at least if we equate such sharing with the sharing of predicates) will be infinite for any two objects. I suspect that these difficulties in making sense of objective generic similarity (i.e. generic similarity in the cosmos) are insurmountable,

and, since the proposed account of matching assumes clear sense can be made of objective generic similarity, it would seem to be unacceptable.

To avoid these difficulties I suggest that we make the question of matching between our quality space and that of the cosmos hinge not on a correlation between our similarity judgments and the state of the cosmos but on a correlation between the projections we make using our quality space and the state of the cosmos. Thus, we can say that our quality space matches that of the cosmos to the extent that our actual and counterfactual projections using this quality space are/would be successful. Thus, if we predict that whales, being similar to fish, will have scales and be cold-blooded, our false prediction will indicate a divergence between our quality space and that of the cosmos.

It should now be clear why a match between our quality space and that of the cosmos should contribute to inductive success. Our quality space clearly plays a major role in determining what inductive projections we make. Indeed, if we are to make a projection based on past experience about any new situation, we will employ our quality space to determine the new situation's similarity to past situations. Thus, it stands to reason that a match between our quality space and that of the cosmos would contribute to inductive success.

However, it might be objected that this explanation of why a match of quality spaces would contribute to inductive success is vacuous. For we seem to have defined the degree of match between our quality space and that of the cosmos in terms of degree of inductive success; thus, it would seem to follow trivially that a

match would contribute to inductive success, for it is a logical consequence of having a high degree of match that one has a high degree of success.

For the most part, I must agree with this objection. We have failed to make sense of a match between our quality space and that of the cosmos other than by appeal to inductive success, so it is pointless to claim that a match somehow explains inductive success.<sup>50</sup> However, admitting this difficulty does not prevent us from putting our discussion so far to good use. Even if we dismiss talk of a "match" of quality spaces explaining inductive success as no more than a façon de parler, the truly important conclusion can still be drawn from the preceding discussion: different quality spaces can increase or decrease inductive success, so the shape of one's quality space (though not its matching the cosmos) can be instrumental in achieving inductive success.

Having clarified the ingredient notions in our argument that Gawanian inductive projections are realized despite their lack of logical or rational privilege, the argument itself can finally be presented.

The argument is that Gawanian inductive projections are successful because Gawanians, along with all humans, have a quality space that leads them, by and large, to project regularities that obtain in the world (i.e. a quality space that "matches" the cosmos), and that there are two factors that conspire to give humans such a quality

---

<sup>50</sup> A similar problem may infect the line of reasoning pursued by Quine in "Natural Kinds" (ch. 5 of Quine (1969)). He does not there present any detailed notion of what a match implies that would avoid this problem.



space. The first is natural. Since having a quality space that leads to inductive success has obvious survival value, biological evolution would select for it. Clearly, however, evolution alone is not sufficient, for we are naturally prone (by our innate quality space) to make predictive mistakes. For example, we mistake whales for fish and non-poisonous king snakes for the similarly colored poisonous coral snakes evolution has selected them to mimic. A second factor, a cultural one, enters in to correct the deficiencies in our innate quality space: our innate quality spaces are fine-tuned by second-order inductions (cf. Quine:1969, pp. 128-9). That is, when we see our inductions fail, we infer that such predictions will not work in the future and seek to alter our quality space accordingly. For example, we examine the coloration of the snakes more closely and, after noting two color patterns, bifurcate that section of our previously unified quality space. Such fine-tuning can be passed on from generation to generation and so become part of the cultural heritage human societies give their members. In these two ways, natural and cultural, humans, including Gawanans, come to have a quality space that enables them to project regularities that obtain in the world.

Reasons have now been given for believing that both elements requisite for the Gawanans' inductions to be an effective means of sanctioning actually reliable *bfp*'s--true premisses and the projection of existing regularities--obtain. Suppose they do obtain. Since beliefs acquired by sanctioned *bfp*'s are those to which knowledge is ascribed, this will mean that knowledge is ascribed to r-justified beliefs. To make plausible the claim that knowledge ascriptions are typically made of br-knowledge

(claim I), it only remains to be shown that the other elements of br-justification are present in beliefs to which knowledge is ascribed.

### 7.2.2 *Ascribing Knowledge to Br-justified Beliefs*

Br-justified beliefs are r-justified beliefs whose authority is r-recognized. Hence, since we have given good reasons to believe that the Gawanans' ascribe knowledge to r-justified beliefs, we can make plausible the claim that they ascribe knowledge to br-justified beliefs by showing that their knowledge attribution practice admits only beliefs whose authority is r-recognized, and that is the goal here.

Let us begin by recalling the two components of authority recognition for bilevel reliabilism:

*S* recognizes the authority of his belief in *p* iff:

- (a) *S* r-knows that his belief in *p* was acquired via *bfp* *b*.
- and (b) *S* r-knows that *b* is a reliable *bfp*.

There is an internalist and an externalist element to each of these clauses. The internalist element is the requirement that *S* have meta-beliefs of these sorts, the externalist element is that the meta-beliefs constitute r-knowledge.

Let us consider first whether the Gawanans ascribe knowledge only to cases in which the internalist element is present. To determine this, it will be useful to consider first- and third-person knowledge attributions (i.e. knowledge avowals and attributions of knowledge to others) separately. These are examined in turn.

For the Gawanans, part of the act of (correctly) avowing knowledge is to "ascertain" whether one's belief has been acquired by one of the socially approved

*bfp*'s, for if it has not the avowal is incorrect. Hence, in avowing knowledge, one does, as a matter of course, acquire the meta-belief required by clause (a)--a belief about which *bfp* led to the object-belief. Does one also acquire the meta-belief required by clause (b)--a belief about whether the *bfp* is reliable? If one is to attempt to avow knowledge correctly one must "ascertain" whether the particular *bfp* producing the object belief is one of those socially sanctioned. Making such an effort would not eo ipso produce beliefs about the reliability of the *bfp*, but if we consider the process of social sanctioning more closely, we will see that such beliefs will be produced. In Gawan society the sanctioning begins with inductions which have as their ultimate conclusions that certain *bfp*'s are trustworthy. Hence, the rationale for sanctioning is clearly that the *bfp*'s are trustworthy, i.e. reliable. In practical terms this means that parents will pass these sanctions on to their children giving trustworthiness as the rationale for the sanctioning, and that any *bfp* whose trustworthiness comes to be doubted may lose its sanction. Thus, if one comes to believe that his *bfp* is one of those socially sanctioned, one does, typically, also come to believe that it is a reliable *bfp*. Thus, it can be seen that, typically, one has the meta-belief pertaining to clause (b) as well; hence, the first-person internalist part of our present project is complete.

These internalist elements do not seem to be similarly required for third-person knowledge ascription. In such cases, since the ascription is done by another, the cognizer possessing the object-belief is not required by the act of ascription to have beliefs either about which *bfp* led to his belief or about the reliability/social sanction-

ing of that *bfp*. And I am not certain that there is any aspect of Gawanian epistemic practices that decisively requires internalist features in cases of third-person attribution. However, it is worth noting that typically in such a case the attributor would possess these internalist features, i.e. he must believe that the belief was acquired by his fellow Gawanian by a *bfp* *b* and believe that *b* is trustworthy; otherwise he would not be disposed to make the attribution. Thus, in such cases, the act of knowledge attribution would require at least that someone in the society, viz. the attributor, possess the internalist element of br-justification. In any event, this slight chink in our present argument is heavily mitigated by two factors. First, third-person knowledge ascription seems to occur relatively infrequently. Second, as seldom as third-person knowledge ascription occurs, third-person knowledge ascription to beliefs for which the cognizer lacks these meta-beliefs will occur even more seldom. Thus, though knowledge may sometimes be correctly ascribed to non-br-justified beliefs in such situations, it is not really a problem for the present explanation since it occurs so rarely that its overall effect is negligible. (Consequently, I will hereafter ignore problematic cases of third-person knowledge attribution.)

Next we examine whether the externalist element of authority recognition is typically present in the beliefs to which the Gawanians' ascribe knowledge, i.e. whether the meta-beliefs associated with these object beliefs constitute r-knowledge. Consider first the clause (a) meta-belief--the belief that one's belief has been acquired by *bfp* *b*. It seems reasonable to suppose that that meta-belief typically constitutes r-knowledge simply because it is acquired via introspective faculties that seem to be

quite reliable when it comes to determining just how we acquire a certain belief (or, at least, how we sustain that belief in cases in which we have forgotten just how we acquired it).<sup>51</sup> The clause (b) meta-belief--that the relevant *bf* is reliable--also seems typically to constitute r-knowledge. For, as was noted above, this meta-belief can be seen as being acquired via an inference from the belief that the relevant *bf* is socially sanctioned.<sup>52</sup> And acquiring beliefs via such an inference would seem to be a reliable *bf* in such cases, for the premiss will typically be true--determining whether a *bf* is socially sanctioned is not a difficult task--and, when it is, the conclusion will also typically be true, for, as was discussed in the previous section, the inductive practices employed by Gawanans are an effective means of sanctioning only reliable *bf*'s.

We can now see how all the elements of authority recognition will obtain, typically, for the beliefs to which the Gawanans' ascribe knowledge, and, thus, that it is plausible that such beliefs are mostly br-justified beliefs and, hence, mostly br-knowledge.

---

<sup>51</sup> It may be objected that there is a significant number of situations in which such faculties are not reliable. For example, emotions and desires often lead to the formation of belief via wishful thinking, and often in such cases we are unaware of their role in the formation of our belief. But epistemic practices (ours, and, we will assume the Gawanans' also) correct for such cases by the practice of challenging and defending a knowledge claim. Being challenged to defend one's belief or merely knowing that one may be challenged should lead a person to consider a knowledge ascription more carefully and not avow knowledge in such cases.

<sup>52</sup> Individualists may protest that if Gawanans' inferences typically proceed in this direction their judgments of reliability will be inordinately dependent on their society. However, relying on society for judgments of reliability can help avoid individual errors in making such judgments. For example, errors due to too limited experience with the *bf*, or the obtuseness of the particular individual, or his emotional biases.

### *7.3 That the Practice of Knowledge Attribution Promotes Br-knowledge and that such Promotion is Useful*

Having established the plausibility of the claim that Gawan knowledge ascriptions are typically made of br-knowledge (claim I), we are now ready to conclude the explanation of the function of the Gawan practice of knowledge attribution by arguing that this practice promotes br-knowledge (claim II) and is useful because it does (claim III).

Thus far we have been concerned with the Gawan practice of knowledge attribution primarily considered as a means for the classification of beliefs as knowings or not. However, to complete this final stage of our explanation we must construe the practice a bit more broadly, for clearly the mere ability to classify beliefs need not promote br-knowledge since the Gawans may never exercise the ability. If the practice is to promote br-knowledge, it must be more inclusive. To see what to include, let us consider the argument about to be presented. The general argument that the practice promotes br-knowledge is that Gawans desire, for a number of reasons, to possess "knowledge" (the quotation marks here indicate that the word is being used in the Gawan sense), and, since "knowledge" is applied mostly to br-knowledge, such desires will result in an increase in br-knowledge in the community. Thus, to give this argument a chance to succeed, we will include in the Gawan practice of knowledge attribution two demands for "knowledge" that are consistent with the epistemic institutions of the Gawan community and human societies generally.

The first such demand arises from the social demand that one's opinion be especially trustworthy in certain "high stakes" situations: e.g. when one is asked for information, when one is contradicting another with greater authority, or, at least in contemporary society, when one is determining the guilt or innocence of a defendant. Since knowledge, for Gawanans, is distinguished from other beliefs primarily in terms of its trustworthiness, a demand for trustworthy opinion is easily transformed into a demand for knowledge.

A second such demand arises not from social but from individual desires. Usually, ceteris paribus, we seek to be right about things. Since "knowledge" is held to be trustworthy, we will naturally seek it out as a way to be right.

Including these demands in the practice of knowledge attribution makes clear that it promotes br-knowledge in the community: members desire "knowledge" for at least these two reasons, and, since "knowledge" is mostly br-knowledge, their efforts to realize their desires should bring more br-knowledge into being. Such promotion is also clearly a useful thing, for as we saw in previous chapters, br-knowledge is a very useful kind of belief to have.

We have finally concluded the bilevel reliabilist explanation of the function of the Gawanans' knowledge attribution practice. The function of their practice is the useful promotion of br-knowledge. We must now attempt to apply what we have established so far to the conventionalist foundationalist version of the practice--the version that would seem to exist today--and then consider to what extent what we have established about Gawanans society applies to human societies generally.

#### 7.4 Explaining the Practice of Knowledge Attribution as it Exists Today

Continuing to use Gawan society as a paradigm, we now turn our attention to explaining the origin and function of the conventionalist foundationalist practice of knowledge attribution. We first attempt to explain its origin by considering how such a practice might evolve from the Gawanans' pedigree practice and then attempt to explain the function.

The conventionalist foundationalist practice was described in the previous chapter as follows:

The title of knowledge is conferred on just those beliefs which receive direct social sanction or which the believer can infer by socially sanctioned norms from claims which receive direct social sanction.

To see how such a practice might evolve from the Gawanans' pedigree practice recall that their pedigree practice sanctions *bfp*'s directly, and, by extension, sanctions beliefs in terms of their *bfp* ancestry. However, this unstructured and rather haphazard sanctioning process may become unwieldy. Since it sanctions *bfp* types in piecemeal fashion, there may come to be so many sanctioned *bfp*'s that individual Gawanans cannot effectively master the epistemological heritage of their community. An evidentialist practice such as conventionalist foundationalism might evolve as an attempt to introduce a structure to the sanctioning process to make it more manageable. Let us take a closer look at this practice to see how this evolution might occur.

The conventionalist foundationalist practice allows sanctioning of beliefs either non-inferentially or inferentially. A belief is sanctioned non-inferentially if it is determined to belong to one of a certain set of privileged kinds of beliefs, e.g. beliefs



in analytic propositions or beliefs about the way things appear. A belief is sanctioned inferentially if it is determined to be inferable by socially sanctioned inference norms from sanctioned beliefs. A switch to such a two-tiered sanctioning system could greatly streamline the sanctioning process for the community, but, of course, for such a switch to be acceptable the new process must sanction mostly the same beliefs as the old. This streamlining could occur while maintaining sanctioning equivalence as follows. The *bfp*'s sanctioned by the pedigree practice would be assigned to one of the two tiers--non-inferential or inferential--and equivalents to each introduced. *Bfp*'s involving an explicit psychological inference as well as many involving only implicit inferences would be assigned to the inferential tier, and the remainder would be assigned to the non-inferential tier. Translation at the non-inferential level would not involve streamlining because each *bfp* would be translated into a corresponding kind of belief, e.g. a sanctioned *bfp* of "direct observation" might be "translated" into a sanctioning of "beliefs resulting from direct observation". But at the inferential tier things could be much streamlined, for a variety of sanctioned *bfp*'s using one inference rule could be replaced by the sanctioning of the inference rule. For example, to use a somewhat contrived example, sanctioned *bfp*'s of believing-that-a-food-is-non-poisonous-after-directly-observing-people-eating-it-with-no-ill-effects-on-numerous-occasions and believing-that-tigers-will-attack-humans-after-directly-observing-attacks-on-numerous-occasions could be reduced to the sanctioning of the inference norm of enumerative induction (assuming the non-inferential sanctioning of direct observation is already in place). In this way, a switch to a conventionalist foundationalist practice

could both streamline the sanctioning process and maintain the same corpus of privileged beliefs.

It can be seen, then, that a conventionalist foundationalist practice would be a practical improvement over the Gawanans' pedigree practice and that such an advantage could provide the impetus for its evolution from the pedigree practice. To complete the explanation of the origin of a Gawanian conventionalist foundationalist practice we should explain the mechanics of this evolution, i.e. just how the transition from one practice to the other occurs, and why the conventionalist foundationalist rather than some other simplifying practice evolved. But I will not attempt to add these elements here since they would probably be tedious, highly speculative, and, as will become evident, are not required for the success of our overall project.

We now turn to the explanation of the function of the Gawanian conventionalist foundationalist practice described above. Above, it was noted that the evolution from a pedigree practice to a conventionalist foundationalist practice would tend to retain the same corpus of "knowledge", i.e. would count the same set of beliefs as knowledge. Given this extensional equivalence of the two practices, the task of explaining the function of the conventionalist foundationalist practice becomes one of simply transferring to it, mutatis mutandis, the explanation of the function of the pedigree practice. As with the pedigree practice, the explanation of the function of the conventionalist foundationalist practice is that it promotes the existence of br-knowledge via the combination of three features: (1) that its "knowledge" is mostly br-knowledge; (2) that there is a social demand for such "knowledge" in high stakes

situations; and (3) that individuals desire such knowledge because they desire to be right. Feature (1) is achieved because it was shown how it could be true for the pedigree practice above, and because the two practices are presumed to be extensionally equivalent. Features (2) and (3) are achieved as they were with the pedigree practice.

### *7.5 Summary of the Explanation and Consideration of the Generality of its Application*

We have now completed the presentation of the bilevel reliabilist explanation of the origin and function of the Gawanans' knowledge attribution practice. In sum, the explanation of the origin is that Gawanans were led to develop a knowledge attribution practice because, first of all, their inductive tendencies led them to have beliefs about the trustworthiness of various types of belief-forming processes. These beliefs, in turn, led to the community's sanctioning certain *bfp*'s as trustworthy. And a word for knowledge was finally introduced--and thus a knowledge attribution practice was fully realized--to abbreviate the otherwise tedious process of claiming that one's belief was acquired via one of the sanctioned *bfp*'s. From this pedigree knowledge attribution practice there then evolved, as a way of simplifying the sanctioning process, a conventionalist foundationalist practice.

The principal function of the Gawanans' pedigree knowledge attribution practice is promoting the existence of a highly useful type of belief, *br*-knowledge, within the community. Briefly, the explanation of how the practice achieves this function is as follows. Armed with a set of mostly true beliefs and an effective

quality space for making inductions that has been selected for by evolution and fine-tuned by second-order inductions, Gawanans sanction certain *bfp*'s as trustworthy. Since this belief set and quality space make Gawanans relatively efficient cognitive mechanisms, their sanctions succeed in picking out mostly reliable *bfp*'s. Beliefs resulting from these processes become known as knowledge, and, typically, the authority of these beliefs is r-recognized because of various requirements involved in the act of knowledge attribution pertaining to its place in the space of reasons and the space of communal interactions. Thus, knowledge comes to be attributed primarily to br-knowledge. Because of its presumed trustworthiness, there are strong social and individual demands for knowledge, and these demands cause the community to increase the production of knowledge. It is in this way that the pedigree knowledge attribution practice functions to promote br-knowledge. The conventionalist foundationalist practice promotes br-knowledge in just the same way, attributing, as it does, knowledge to just the same beliefs.

Let us now consider to what extent this explanation is applicable to human societies generally. Since, in the course of the explanation, no explanatorily important attributes were attributed to Gawanans society that could not be possessed by human societies generally (i.e. no special historical developments or religious beliefs), it seems clear that it is possible that this is the explanation, mutatis mutandis, for the practice of knowledge attribution in all human societies. But our initial goal, it will

be recalled, was to provide a plausible<sup>53</sup> explanation. And it seems to me that this too has been accomplished, for arguments have been provided for most parts of the explanation and these have relied upon factors, such as evolutionary and cultural pressure, that are universal to human societies. However, a greater plausibility has been established for some parts of the explanation, especially when it is considered as a general explanation, than others, and it will be worth noting these. Perhaps the least well supported parts of the explanation concern the explanation of the origin. Little reason was given to believe that inductions concerning *bfp*'s (rather than, e.g., inductions concerning reasons) should be the ones that lead to the development of a knowledge attribution practice. Moreover, since the plausibility of the claim that an evidentialist practice evolved from a pedigree practice depends on the plausibility of the claim that a pedigree practice developed first, the former claim would also seem to have relatively weak support. However, it should be noted that the plausibility of the explanation of the function is not directly dependent on the plausibility of the explanation of the origin. For the key to the explanation of the function was establishing that a conventionalist foundationalist practice promoted *br*-knowledge, and that could be true even if the practice evolved in some entirely different way.

---

<sup>53</sup> A "plausible" explanation, for present purposes, differs from a merely possible one primarily in that there will be some, though far from conclusive, good reasons for believing that it is actual, and that there will be no obvious reasons for believing that it is not actual.

### *7.6 Implications of the Bilevel Reliabilist Explanation of the Existence and Function of Epistemic Practices*

We turn now to consider the implications of the bilevel reliabilist explanation of epistemic practices. In the previous chapter a number of reasons for undertaking such a task were discussed: increasing our knowledge of an important part of human affairs; providing a defense for non-traditional epistemologies against charges of explanatory inadequacy; providing a weapon against epistemological anarchism, relativism, and nihilism; and providing support for bilevel reliabilism as a basis for constructive epistemology. In assessing the extent to which the explanation achieves these goals, we must bear in mind that it is only claimed to be a plausible, not definitive, explanation. Though it might be doubted that merely a plausible explanation could go very far in achieving such goals, it will become evident that it makes significant progress.

The first goal was to increase our knowledge of this important part of human affairs. I would argue that this goal has been achieved by the explanation. The explanation makes clear and, to my knowledge, for the first time, the details of how a practice of knowledge attribution could originate and be of great general utility despite there being no rational privileging of the sort the practitioners believe (at least in some of the practice's forms) to be the raison d'être of the practice. Thus, it offers researchers an appealing alternative to extant hypotheses, most of which have either the problem of relying upon the existence of rational privileging or, in denying the existence of rational privileging, make epistemic practices appear useless.

The second goal was to provide a defense for neo-Humean epistemologies against charges of explanatory inadequacy. Such charges, it will be recalled from the previous chapter, were that these epistemologies could not satisfactorily explain why such practices existed. For in denying the existence of rational privileging, they deny epistemic practices a general utility that could explain their near universal existence in human communities. The bilevel reliabilist explanation provides the tools for a defense against such charges since it shows how epistemic practices might have originated simply from universal human inductive tendencies.

The third goal was to provide a weapon against epistemological anarchism, relativism, and nihilism. The explanation succeeds here too (although, as will become evident, the weapon is primarily defensive in nature). But before examining how it succeeds, let us pause to clarify the meaning of these three terms in the present context. Our present concern is the current practice of knowledge attribution, which, for these purposes, can be thought of as a set of principles bestowing epistemic honor on some beliefs and theories and denying it to others. An epistemological anarchist<sup>54</sup> about this practice will hold that it is unnecessarily restrictive. The principles pick out no set of beliefs and theories more deserving of honor than any other set of beliefs and theories; thus, we would be better off without the practice since it denies, without justification, honor to certain beliefs and theories. An epistemological relativist holds that though there may be a point to honoring some beliefs/theories over others, there are many mutually incompatible and widely

---

<sup>54</sup> This term is from Feyerabend (1975).

divergent sets of principles that humans might adopt to distribute this honor, and the honor bestowed by one of these systems takes (and ought take) no precedence over that bestowed by another. An epistemological nihilist holds that the practice has no useful function and ought to be abandoned.

To see how the bilevel reliabilist explanation provides a weapon against such extremist views, consider the structure of the debate between epistemological extremists and moderates. Moderates wishing to oppose extremist views face a dialectical dilemma regarding the existence of rational privileging. Either they accept its existence or they do not. If they accept its existence, they should be able to marshal powerful attacks against extremist views. For, as was noted in the previous chapter, the current practice of knowledge attribution seems to be aimed at bestowing honor on rationally privileged beliefs, and, if there is such privileging and the practice is largely successful in picking out beliefs so privileged, then it is relatively easy to argue that the practice has a useful function and the honor bestowed is genuine and, hence, that extremist views are wrongheaded. But if moderates accept the existence of such privileging, then they must also accept the burden of defending their belief in its existence. However, the history of skeptical attacks on this view shows its defense to be very difficult--indeed, so difficult that for the purposes of this dissertation we have assumed it to be hopeless. Thus, the first horn of the dilemma is complete--there is no more dialectical room in which moderates can maneuver. Suppose, then, that moderates deny the existence of rational privileging. Extremists would seem to have good cause to rejoice at this move as well. For it is accepted by both sides (or



at least we will assume so for present purposes) that current practices are held by society at large to bestow honor on rationally privileged beliefs, and if such privilege does not exist, the conclusion seems clear that these practices fail at their appointed task. And, once this conclusion is drawn, the door seems wide open for nihilists to attempt to show that epistemic practices are pointless, for anarchists to show that the honor bestowed is bogus, and for relativists to argue for equal status for a variety of systems of bestowing honor since the current one has no special place from the standpoint of reason.

The bilevel reliabilist explanation provides a weapon for moderates by providing for them a way out of this dialectical dilemma. It provides a way out by allowing an opponent who opts for the horn denying the existence of rational privileging to block the advance of extremists by appealing to the plausible possibility presented by the explanation. To the epistemological nihilist, for example, he can say:

Yes, there is no rational privileging, and so we agree that the practice does not have the function it is generally supposed to have. However, the practice may have other useful functions. And I am not here attempting to block nihilistic moves merely by shifting to you the burden of proving that there are no such alternative functions, for the bilevel reliabilist explanation of epistemic practices gives a detailed and plausible account of how such practices might have just such a function.

And to the anarchist he can offer the bilevel reliabilist explanation as a plausible scenario according to which the honor bestowed is not bogus, for it is granted mostly to cases of br-knowledge. And to the relativist he can decry the equal status given various systems of distributing honor, for, if the bilevel reliabilist explanation is cor-

rect, our system bestows it largely to cases of br-knowledge--a sort of belief useful in most any community--and other systems might not. In this way, the bilevel reliabilist explanation can be of assistance to moderates who deny rational privileging and who--though tempted by the heady intellectual freedom of the nihilist, anarchist, and relativist positions--wish to avoid the epistemic chaos inherent in each.

As a way of further elucidating the manner in which the explanation serves as a weapon against extremist positions, I want to make clear that there is one means of using it as a weapon that I am definitely not advocating. This means could be employed only after the explanation had been extensively tested and confirmed so that there was widespread scientific agreement that it was the best explanation. Once this had been achieved, there would then be those who would wish to claim "our best science indicates the extremists are wrong--epistemic practices do have a useful function, and they do pick out beliefs that are deserving of honor, [etc.]". There are two difficulties with such a claim. First, the claim is appealing to the authority of science as a means of refuting epistemological extremism, but, of course, it is the very authority of science that the extremist questions, so such an appeal has no force with him. (Of course, if it is then claimed that the explanation establishes the authority of science, the appeal to the extremist will be blatantly circular.) A second difficulty is that the claim seems to assume the existence of rational privileging, for it seems to imply that we ought to heed the dictates of our best science because its results are rationally privileged.

It should now be clear that I do not intend to use the bilevel reliabilist explanation to provide any sort of science-based refutation of extremist positions. The dialectical move it provides the moderate position is not one of refutation, but of defense against attack. It can be used as a defensive, but not an offensive weapon. And the defense it provides does not involve any sort of claim of rational privilege for the moderate position, but only a defense of the tenability of the position.

In addition to providing moderates with an escape from the aforementioned dialectical dilemma, the explanation also can serve to defend a moderate position from a different, though closely related, sort of attack that the extremist might launch. This sort of attack claims that a position, in some respect, undermines itself. Such an attack would seem to be successful against a traditional epistemological position (one that affirms the existence of rational privileging). For such a view undermines itself in that if one employs the practices it recommends, i.e. the usual sorts of Wissenschaftliche methods<sup>55</sup>, to determine whether there are rationally privileged beliefs, the result is negative (or so I, and the extremist, would argue). The bilevel reliabilist explanation can be seen as helping to show that a moderate position denying rational privileging does not undermine itself. For the explanation shows that if one employs the epistemic practices recommended by moderate positions (which are the same as those recommended by the traditionalist) to determine whether epistemic practices serve a useful function, the answer is positive.

---

<sup>55</sup> By this expression I mean to refer to those methods generally regarded as sound scholarly/scientific methods.

(The preceding discussion may lead the reader to wonder, if I make no claim of rational privilege for the bilevel reliabilist explanation, just what, if any, sort of special status I do believe it has and why, in any event, I am advocating it. I suspect my answer will be unsatisfying to those inclined to ask such questions. My answer, in general terms, is just that I believe the explanation is that favored by the sort of methodology employed in generating it (i.e. Wissenschaftliche methodology). I do not believe this methodology has any sort of rational privilege, but it is consistent with the form of life<sup>56</sup> I favor, and that is why I advocate the views that I believe it favors.)

The fourth goal of undertaking an explanation of epistemic practices was to provide support for bilevel reliabilism as a basis for a constructive neo-Humean epistemology. Since bilevel reliabilism is a radical epistemology, to establish it as a basis for constructive epistemology requires that it be shown capable of accomplishing a variety of tasks. This has been accomplished, in part, by showing how it achieves the three goals discussed above. Thus, though it would probably be going too far to say that the radical nature of bilevel reliabilism has been completely justified, a substantial beginning for doing so has been achieved.

---

<sup>56</sup> Very roughly, by "form of life" I mean to refer to the fundamental assumptions one makes about the universe and one's place in it that do not derive from other beliefs one has.

## **BIBLIOGRAPHY**

- Alston, William: 1983, 'What's Wrong with Immediate Knowledge?', Synthese 55, 73-95.
- Alston, William: 1986, 'Internalism and Externalism in Epistemology', Philosophical Topics 14, 1, 179-221.
- Alston, William: 1988, 'An Internalist Externalism', Synthese 74, 265-83.
- Armstrong, David M.: 1973, Belief, Truth, and Knowledge, Cambridge U. Press, London.
- Austin, John L.: 1961, Philosophical Papers, Clarendon Press, Oxford.
- Bonjour, Laurence: 1985, The Structure of Empirical Knowledge, Harvard U. Press, Cambridge.
- Boyd, Richard: 1984, 'The Current Status of Scientific Realism' in Leplin (1984).
- Chisholm, Roderick M.: 1988, 'The Indispensability of Internal Justification', Synthese 74, 285-96.
- Cohen, Stewart: 1984, 'Justification and Truth', Philosophical Studies 46, 279-95.
- Dretske, Fred: 1971, "Conclusive Reasons", Australasian Journal of Philosophy 49, 1-22.
- Dretske, Fred: 1981, Knowledge and the Flow of Information, MIT Press, Cambridge.
- Feldman, Richard: 1985, 'Reliability and Justification', The Monist 68, 2, 159-74.
- Feldman, Richard and Conee, Earl: 1985, 'Evidentialism', Philosophical Studies 48, 15-34.
- Feyerabend, Paul: 1978, Against Method, Verso, London and New York.
- Foley, Richard: 1985, 'What's Wrong with Reliabilism?', The Monist 68, 2, 188-202.

- Fricker, Elizabeth: 'The Epistemology of Testimony', Aristotelean Society Supplementary Volume 61, 57-83.
- Goldman, Alvin I.: 1967, 'A Causal Theory of Knowing', The Journal of Philosophy 64, 12, 355-72.
- Goldman, Alvin I.: 1976, 'Discrimination and Perceptual Knowledge', The Journal of Philosophy 73, 20, 771-91.
- Goldman, Alvin I.: 1978, 'Epistemics: The Regulative Theory of Cognition', The Journal of Philosophy 75, 10, 509-23.
- Goldman, Alvin I.: 1979, 'What is Justified Belief?' in Justification and Knowledge, New Studies in Epistemology, G. Pappas ed., D. Reidel, Dordrecht, pp. 1-23.
- Goldman, Alvin I.: 1980, 'The Internalist Conception of Justification' in Midwest Studies in Philosophy, Volume V, Studies in Epistemology, P. French et. al., eds., U. of Minnesota Press, Minneapolis, pp. 27-52.
- Goldman, Alvin I.: 1986, Epistemology and Cognition, Harvard U. Press, Cambridge.
- Goodman, Nelson: 1979, Fact, Fiction, and Forecast, 3rd ed., Hackett, Indianapolis.
- Harrison, J.: 1962, 'Knowing and Promising', Mind 71, 284, 443-57.
- Kaplan, Mark: 1985, 'It's Not What You Know That Counts', Journal of Philosophy 82, 7, 350-63.
- Kornblith, Hilary: 1980, 'Beyond Foundationalism and the Coherence Theory', The Journal of Philosophy 72, 597-612.
- Leplin, Jarrett, ed.: 1984, Scientific Realism, U. of California Press, Berkeley.
- Levi, Isaac: 1980, The Enterprise of Knowledge, MIT Press, Cambridge.
- McDowell, John: forthcoming, 'Knowledge by Hearsay' in a collection of essays on the epistemology of testimony, B.K. Matilal and A. Chakrabarti, eds.
- Mackie, John L.: 1977, Ethics, Inventing Right and Wrong, Penguin, Middlesex.
- Papineau, David: 1985, 'Realism and Epistemology', Mind 94, 367-88.

- Peacocke, Christopher: 1986, Thoughts: An Essay on Content, Basil Blackwell, Oxford.
- Pollock, John: 1986, Contemporary Theories of Knowledge, Rowman & Littlefield, Totowa, New Jersey.
- Pollock, John: 1987, 'Epistemic Norms', Synthese 71, 61-95.
- Quine, W.V.O.: 1969, Ontological Relativity and Other Essays, Columbia U. Press, Cambridge.
- Ramsey, Frank P.: 1978, Foundations: Essays in Philosophy, Logic, Mathematics and Economics, D.H. Mellor, ed., Humanities Press, Atlantic Highlands, New Jersey.
- Rawls, John: 1971, A Theory of Justice, Belknap Press, Cambridge.
- Rorty, Richard: 1979, Philosophy and the Mirror of Nature, Princeton U. Press, Princeton.
- Salmon, Wesley: 1966, The Foundations of Scientific Inference, U. of Pittsburgh Press, Pittsburgh.
- Sellars, Wilfrid: 1963, Science, Perception, and Reality, Routledge and Kegan Paul, London.
- Smith, Norman Kemp: 1949, The Philosophy of David Hume, Macmillan, London.
- Sosa, Ernest: 1983, 'Nature Unmirrored, Epistemology Naturalized', Synthese 55, 49-72.
- Stich, Stephen: 1983, From Folk Psychology to Cognitive Science, MIT Press, Cambridge.
- Watling, John: 1954, 'Inference from the Known to the Unknown', Proceedings of the Aristotelean Society 55, 81-108.BIB