PYRRHONIAN
REFLECTIONS
ON
KNOWLEDGE
AND
JUSTIFICATION

ROBERT J. FOGELIN

Pyrrhonian Reflections on
Knowledge and Justification

This page intentionally left blank

PYRRHONIAN REFLECTIONS ON KNOWLEDGE AND JUSTIFICATION

Robert J. Fogelin

Oxford University Press

Copyright © 1994 by Robert J. Fogelin

For Florence

This page intentionally left blank

Preface

Some of the thoughts in this work have been with me a long time. Forerunners of my present ideas about knowledge appeared first in *Evidence and Meaning* (1967), and much that I said there still strikes me as broadly correct. What that earlier treatment lacked was an appreciation of the complexity and depth of the problems it was intended to solve.

At that time I was much under the influence of the writings of J. L. Austin, and I thought philosophical problems could be resolved in reasonably short order through paying careful attention to how words are actually used. I still think valuable insights can be gained in this way and have no desire to join those who now look back at Austin's work with something close to contempt. I moved away from Austin's views—or what I took to be Austin's views—for a number of reasons. Paul Grice's "Causal Theory of Perception" convinced me that Austin's treatment, and hence my treatment, of the problem of perception were wholly inadequate. Studying the works of ancient and modern skeptics convinced me that skeptical challenges are serious challenges. I came to see that these challenges could not be written off with the glib suggestion, common in the 1950s and 1960s, that they result from the skeptic's imposing arbitrary and impossibly high standards for what we can know. My views on these matters were given a considerable boost by hearing, then reading, Thompson Clarke's "Legacy of Skepticism." Immersing myself in the works of Wittgenstein cured me of any lingering idea that philosophical problems admit of easy solutions. Philosophy—including philosophy intended to abolish philosophy—is hard work.

I cannot thank all who have made important contributions to this work, for they include those who have been generous in commenting on my previous writings. I'll limit myself to those institutions and persons who contributed directly to this project.

Institutions first. Though it may not be obvious from its contents, this work was written at a series of pleasant—and almost invariably sunny—places. It was begun in Stanford, California, at the Center for Advanced Studies in the Behavioral Sciences, where I received partial support from the National Endowment for the Humanities. Most of Part I of this work was written in these ideal surroundings. Two years later I was a visiting professor at the University of California at Berkeley, where I offered a seminar to a group of very intelligent graduate students. This enabled me to test my ideas before they had hardened into unshakable commitments. I had hoped to bring the project to completion during a five-week stay at the Rockefeller Center in the Villa Serbelloni, in Bellagio, Italy, but instead I convinced myself that the work needed two additional chapters. The project was essentially finished in the spring of 1992 while I was on sabbatical, living in a Tuscan farmhouse south of Siena.

Of course, none of this would have been possible without the aid of Dartmouth College and its Department of Philosophy. Both have been understanding and flexible in allowing me to take advantage of the opportunities just noted. The college has also supported my research through sabbaticals, supplementary grants, general research funds, and funds associated with the Sherman Fairchild Professorship in the Humanities, which I hold. For more than a decade, Dartmouth College has provided an ideal setting for the dual occupation of teacher and scholar.

A great many individuals have made helpful criticisms of this work. I presented parts of it at a number of institutions, and profited from the vigorous Socratic questioning provided by such occasions. These institutions included Stanford University, Brooklyn College, the University of California at Berkeley, Dartmouth College, Edinburgh University, Glasgow University, St. Andrews University, and Sterling University. I also profited from the opportunity to discuss Donald Davidson's views with him at the Villa Serbelloni.

My chief debt, however, is to my colleagues in the philosophy department at Dartmouth. Without their energetic criticisms, this work would have been finished sooner with a number of flaws still in place. Specifically, Bernard Gert, John Konkle, James Moor, James Page, and Walter Sinnott-Armstrong participated in a study group that subjected the manuscript to high scrutiny.

I wish to thank Angela Blackburn who, as philosophy editor, shepherded this project through the Oxford University Press, and

also the anonymous reviewer who supported its publication while making important suggestions for its improvement. I also wish to thank Jane Taylor for her help in copyediting the text.

Finally, I must thank my wife, Florence Fogelin, for her intellectual support, her sharp eye for editorial improvement and, not insignificantly, her command of Italian—though it should be said that, in compensation, she exercised these virtues in splendid surroundings.

Hanover, New Hampshire   R.J.F.
September 1993

This page intentionally left blank

Contents

This page intentionally left blank

Pyrrhonian Reflections on
Knowledge and Justification

We have first raised a dust, and then complain, we cannot see.
GEORGE BERKELEY

Introduction: Philosophical Skepticism and Pyrrhonism

When speaking of a philosophical skeptic we might have in mind someone who doubts things—or calls things into question—on philosophical grounds. In this way of speaking, philosophical skepticism is philosophical because philosophy is the source of the skepticism. We could also speak about someone being a philosophical skeptic in order to indicate that he or she is skeptical *about* philosophy, just as we say that someone is a religious skeptic in order to indicate that he or she is skeptical about religion. In this second way of speaking, philosophical skepticism is philosophical because philosophy is its target. To avoid confusion, I will reserve the label "philosophical skepticism" for the first sort of skepticism, namely, skepticism that arises from philosophical reasoning. The second sort of skepticism—where philosophy is the target of the skepticism—I will refer to as skepticism *about* philosophy.

Is it possible to combine philosophical skepticism with skepticism about philosophy, that is, to have doubts about philosophy on the basis of philosophical arguments? There is, of course, something self-referentially peculiar about such a position, since the skepticism about philosophy would seem to undercut the philosophical arguments used to attack it. Whether or not this is a tenable position, it is clearly the position adopted by the Pyrrhonian skeptics—at least in its late form as represented in the writings of Sextus Empiricus.[1] Pyrrhonian skepticism, in its late form, uses self-refuting philosophical arguments, taking philosophy as its target. In what follows, when I speak of Pyrrhonian skepticism, I will have this position in mind.

Some of the paradoxical character of using self-refuting arguments may be removed when we recall that the Pyrrhonist sought suspension of belief to further a practical goal. When one attains suspension

of belief, one is supposed to find oneself in a state of ataraxia (qui-
etude), which, for the Pyrrhonist, is a form of blessedness.[2] Thus, the
Pyrrhonist philosophizes only as a temporary expedient, and, once the
anxieties produced by dogmatic philosophizing have been sur-
mounted, the Pyrrhonist's own skeptical arguments may be discarded
as a ladder no longer of use (*Sextus Empiricus: Against the Logicians*,
2: 481). But the Pyrrhonian arguments are not simply dispensable,
for, as specimens of philosophizing, they too stand in need of elimi-
nation. This is brought out in an image more apposite than that of the
ladder:

> For, in regard to all the Skeptic expressions, we must grasp first the
> fact that we make no positive assertion respecting their absolute
> truth, since we say that they may possibly be confuted by them-
> selves, seeing that they themselves are included in the things to
> which their doubt applies, just as aperient drugs do not merely
> eliminate the humours from the body, but also expel themselves
> along with the humours. (*PH*, 1: 206–7)

I will take this acceptance of self-refutation (peritrope) as a defining
character of Pyrrhonian skepticism as I understand it.

I think the historical movement known as Pyrrhonism as repre-
sented in the works of Sextus Empiricus is an example of Pyrrhon-
ian skepticism as I have explained it. Some would disagree. I don't
think there will be serious objections to the claim that Pyrrhonism
took philosophy as one of its chief targets; it was certainly a skepti-
cism about philosophy. Nor do I think there is any question that the
late Pyrrhonists embraced self-refutation in the sense in which I
described it. The texts on both these matters seem clear. A further
question is whether the Pyrrhonists limited their skeptical attacks
to philosophy. The answer to that is clearly no, for in *Against the
Professors* Sextus adopts a skeptical stance toward the claims made
by professors of grammar, rhetoric, geometry, arithmetic, astrology,
and music. Although each of these fields can have philosophical ele-
ments, Sextus explicitly did not treat them as areas of philosophical
investigation. All the same, as he tells us, the Pyrrhonists came to
suspend belief in these areas as well. He explains their reasons for
doing so this way:

> In respect of the Arts and Sciences, they [the Pyrrhonists] have met
> with the same experience as they did in respect of philosophy as a
> whole. For just as they approached philosophy with the desire of
> attaining truth, but when faced by the equipollent conflict and dis-
> cord of things, suspended judgment,—so also in the case of the Arts
> and Sciences, when they had set about mastering them with a view

to learning here also the truth, they found difficulties no less seri-
ous, which they did not conceal. (*Against the Professors,* 1: 6–9)

That is, finding no more agreement among the professors of the arts
and sciences than they found among the philosophers, the Pyrrhon-
ists again were led to a suspension of belief.

It is not clear, however, whether the teachings in all the branches
of the arts and sciences are fit targets for suspension of belief. Perhaps
medicine, pursued undogmatically, produces claims worthy of our
assent.[3] In fact, Sextus himself sometimes speaks with no apparent
reservation about things he says have been established in a particu-
lar science, as, for example, in the following remark: "For the body
is a kind of expression of the soul, as in fact is proved by the science
of Physiognomy" (*PH,* 1: 85).

It is not completely clear, then, which of the arts and sciences
are subject to the Pyrrhonist calls for suspension of belief and which
are not, but at least it is clear that these calls for the suspension of
belief encompass more than philosophy narrowly conceived and,
perhaps, less than the arts and sciences broadly conceived.

A more difficult question, and here commentators are sharply
divided, is whether Pyrrhonists called for, or attempted to induce, a
suspension of belief with respect to the common beliefs of everyday
life. In "The Beliefs of a Pyrrhonist,"[4] Jonathan Barnes raises this
question by distinguishing two sorts of Pyrrhonism that might be
found in the *Outlines of Pyrrhonism:*

> The first type I shall call, following Galen, *rustic Pyrrhonism.* The
> rustic Pyrrhonist has no beliefs whatsoever: he directs *epoche*
> towards every issue that may arise. The second type of Scepticism I
> shall call *urbane Pyrrhonism.* The urbane Pyrrhonist is happy to
> believe most of the things that ordinary people assent to in the ordi-
> nary course of events: he directs *epoche* towards a specific target—
> roughly speaking, towards philosophical and scientific matters. (2–3)

Which sort of Pyrrhonism is found in Sextus's *Outlines of
Pyrrhonism?* Barnes himself concludes that "the general tenor of *PH*
is, I think, indubitably rustic. But *PH* also contains important intru-
sions of urbanity" (18). Miles Burnyeat has also plumped for the rus-
tic interpretation.[5] Michael Frede has defended the urbane interpre-
tation—though he does not use this (somewhat tendentious) label.[6]

Barnes tells us that defenders of the rustic interpretation are
likely to defend it for two chief reasons:

> First, many of the arguments in *PH* appear to demolish *all* beliefs
> on a given topic if they demolish any beliefs: the attack on causation

or on time or on truth, say, do not appear to restrict their target to
scientific or philosophical positions in those areas; and the Five
Tropes of Agrippa, in terms of which much of the argumentation of
PH is conducted, seem wholly indifferent to any distinction
between scientific theory and everyday opinion. Secondly, *PH*
makes it plain that the opponents of Pyrrhonism regularly construed
Pyrrhonism in rustic fashion—the notorious argument that Sceptics
cannot act evidently presupposes that Pyrrhonists have no beliefs at
all. (3)

Barnes's second argument carries little force, given the common
practice among protagonists of attributing unacceptable views to
their opponents in order to refute them. Thus, for this part of Bar-
nes's argument to have force, a text must be found that states that
the skeptical tropes that were intended to induce a suspension of
belief with respect to philosophical and scientific beliefs were also
intended to be used, with the same effect, with respect to ordinary
beliefs. Barnes, however, candidly admits that no such text appears
in the *Outlines of Pyrrhonism*. "If we are concerned to discover the
scope of *epoche* in *PH*, it is precisely such humdrum sentences
[those expressing ordinary beliefs] which will most exercise us; yet
of them Sextus says nothing" (10).

So an argument intended to show that the Pyrrhonists called for
a suspension of everyday (humdrum) beliefs must be indirect, and, in
fact, Barnes rests his case primarily on his first point noted above,
namely, "that many of the arguments in *PH* appear to demolish *all*
beliefs on a given topic if they demolish any beliefs." That is, *as
stated*, these skeptical arguments seem to apply equally to common
beliefs and philosophical or scientific beliefs; if one goes, so must the
other.

Barnes uses the Pyrrhonian argument from the criterion to make
this point:

A Pyrrhonist will only believe that the water is tepid if he judges it
to be so; and he can only judge it to be so if he possesses a criterion
of truth by which to judge it. But the thesis that there is a criterion
of truth is itself a *dogma*—indeed it is a perfect specimen of those
philosophico-scientific tenets which the Greeks called *dogmata*.
Now the Pyrrhonist of *PH* rejects all *dogmata*. Hence he will not
have—or rather will not believe that he has—a criterion of truth.
Hence he will not be able to judge, or to believe, that the water is
tepid.

In general, the Pyrrhonist of *PH* will have no ordinary beliefs at
all. Ordinary beliefs are not *dogmata*. . . . Nonetheless, in rejecting

dogmata the Pyrrhonist must reject ordinary beliefs; for the possession of ordinary beliefs presupposes the possession of at least one *dogma*—the *dogma* that there is a criterion of truth. (11–12)

I think this argument is not only wrong, but deeply wrong, and wrong in a way that does not turn on subtle nuances or the technical meanings of certain Greek terms. The passage simply misrepresents the dialectical character of the Pyrrhonian attack on the dogmatists. The Pyrrhonist does *not* hold the view that judgments may not be made in the absence of a criterion of truth. That is a view held by the dogmatists, for example, the Stoic epistemologists who attempted to formulate such a criterion of truth. If the argument from the criterion is correct, it will have as a consequence that the *dogmatist* ought to suspend judgment on her dogmatic philosophical beliefs, and also on her ordinary beliefs, for, as Barnes rightly notes, the argument applies equally to both. But this leaves the Pyrrhonist untouched, for it is no part of his position to suppose that judgments may only be made on the basis of a criterion of truth. Not to see this is not to see what Pyrrhonian skepticism (whether it is right or wrong—persuasive or unpersuasive) is all about.

In "The Skeptic's Beliefs" Michael Frede has made this point with great force:[7]

> Since the skeptic wants to see whether his opponent at least by his own standards or canons has knowledge, he in his own arguments adheres to these standards. But this does not mean that he himself is committed to them. He is aware of the fact, e.g., that ordinarily we do not operate by these [the dogmatic] standards and that it is because his opponents want more than we ordinarily have that they try to subject themselves to these stricter canons; they want "real" knowledge, certain knowledge. (204)

I think Frede is exactly right, and this leads him to make a deeper point. In their writings, typically at the conclusion of a skeptical argument, the Pyrrhonists often state that "one ought to suspend judgment or withhold assent." Here, in calling for a suspension of judgment, the Pyrrhonist seems to be speaking in his own voice, and since the grounds he uses as the basis for this call for the suspension of judgment in the particular case equally apply to all judgments, the Pyrrhonist is *eo ipso* committed to the suspension of all judgment whatsoever. In response, Frede points out that these calls for the suspension of judgment are themselves made from within the dogmatist's framework, which the Pyrrhonist occupies only temporarily for dialectical purposes.

Their aim [in calling for the suspension of judgment] might just be
to point out to the opponent that by his own standards it would
seem that he ought to withhold assent.[8] But since the skeptic has
not committed himself to these standards there is also no reason to
think, just on the basis of these remarks, that he is committed to
the claim that one ought to withhold assent on a particular subject,
let alone to the generalization that one ought always to withhold
assent. (204–5)

We thus get the result that the generalization of the Pyrrhonist's
argument shows that the dogmatists ought to withhold assent on all
subjects including those concerning the affairs of everyday life. On
the other hand, the argument has no implications for what the
Pyrrhonist ought or ought not to believe, except for the subjunctive
claim that the Pyrrhonist ought to believe nothing were he a dog-
matist—which, of course, he is not.

If there are no texts, as Barnes concedes, that directly state that
the Pyrrhonist ought to suspend judgment concerning the affairs of
daily life, and if, as Frede has shown, there are no grounds for draw-
ing this conclusion from what *is* said in the text, there certainly are,
on the other side, a number of texts that seem to state unequivocally
that the Pyrrhonist was *not* calling for the suspension of all beliefs
including beliefs of everyday life. Thus, commenting explicitly on
the range of application of the Pyrrhonian expressions, Sextus
remarks: "We must . . . remember that we do not employ them uni-
versally about all things, but about those which are non-evident and
are objects of dogmatic inquiry" (*PH*, 1: 208). Then, more positively,
he tells us:

Adhering, then, to appearances we live in accordance with the nor-
mal rules of life, undogmatically, seeing that we cannot remain
wholly inactive. And it would seem that this regulation of life is
fourfold, and that one part of it lies in the guidance of Nature,
another in the constraint of the passions, another in the tradition of
laws and customs, another in the instruction of the arts. (*PH*, 1: 23)

It might be possible to go a little way—Barnes gives it a try—in
explaining how a person could pursue a life of this kind sans belief.
For example, what Sextus describes as the guidance of Nature might,
in some cases at least, be treated as an automatic or instinctive
response. The thirsty man, without thought or belief, simply gulps
down water much in the manner that a knee jerks when struck with
a hammer.[9] But, as Barnes himself sees, this automatic-response
account of the Pyrrhonist's ways of acting becomes less and less plau-

sible as we move down the list from the guidance of Nature to the tradition of laws and customs and finally to the instruction of the arts. In fact, it is already implausible when applied to passions and desires. Without possessing a complex set of beliefs, a person could not hanker after a seat on the New York Stock Exchange or want to see Lhasa before she dies. Faced with the prospect of attributing to the Pyrrhonists an implausible view concerning human thought and action while possessing no direct textual evidence to support this attribution, Barnes acknowledges that "*PH* . . . contains important intrusions of urbanity" (18). Still, he continues to maintain that "the general tenor of *PH* is . . . indubitably rustic" (18).

Where do we stand? There are no texts in the *Outlines of Pyrrhonism* that explicitly state a commitment to what Barnes calls rustic Pyrrhonism. There are, however, texts that run dead against this commitment. With this, the entire weight of the argument shifts to what we might call the generalization argument: Sextus is committed to rustic Pyrrhonism because he is committed to things whose natural generalization amounts to rustic Pyrrhonism. That argument, it seems to me, Frede has shown to be mistaken. The conclusion, then, is that the Pyrrhonism of Sextus was urbane. In any case, when I speak of Pyrrhonism I will understand it in this way— although I will henceforward drop the word "urbane."

With the support of Frede's arguments, I think m y understanding of late Pyrrhonism is historically sound. Turning now to an entirely different matter, I wish to ask what Pyrrhonism would look like if it were updated so that it would have direct application to contemporary philosophical debates. In particular, since contemporary philosophy has been deeply concerned with language, what would a Pyrrhonian philosophy look like when given a linguistic turn? My suggestion is this: Just as the traditional Pyrrhonist did not call for the suspension of common (nondogmatic) beliefs, our updated Pyrrhonist would have no complaints against common modes of *expressing* these beliefs. Furthermore, just as the traditional Pyrrhonist took as his target dogmatic beliefs that transcend common belief, our updated Pyrrhonist would take a like attitude toward forms of expression that attempt to transcend (for philosophical purposes) common modes of expression. In sum, our updated Pyrrhonism would have a strong family resemblance to the position developed by Wittgenstein in his later writings.

I have argued elsewhere that Pyrrhonian skepticism is the historical movement that Wittgenstein's late philosophy most resembles,[10]

and I'll return to a comparison between Pyrrhonism and Wittgen-
stein's late philosophy a number of times in this study. The point I
wish to make here is that the Pyrrhonists, as I understand them, are
not debarred from using such words as "know," "certain," "real," and
"true." As common terms used in a common way, they are no more
objectionable than other terms of common use. Thus, speaking in the
common way, the Pyrrhonist can avail herself of the standard con-
trasts: what is known rather than what is believed, what is certain
rather than what is very likely, what is real rather than what is coun-
terfeit, and what is true rather than what is false. Our Pyrrhonist (or
neo-Pyrrhonist) bridles only when these terms are pressed into the
service of dogmatists to do their bidding, for they are then converted
into what Wittgenstein calls superconcepts (*Überbegriffen*):

> We are under the illusion that what is peculiar, profound, essential,
> in our investigation, resides in its trying to grasp the incomparable
> essence of language. That is, the order existing between the con-
> cepts of proposition, word, proof, truth, experience, and so on. The
> order is a *super-order* between—so to speak—*super-concepts*.
> Whereas, of course, if the words "language," "experience,"
> "world," have a use, it must be as humble a one as that of the words
> "table," "lamp," "door." (*Philosophical Investigations*, 97)

Our neo-Pyrrhonists will not, of course, defend common language as
the privileged philosophical language, since they are not interested
in putting forward philosophical claims. In a parallel fashion, the tra-
ditional Pyrrhonists, though defenders of common beliefs against the
criticisms of dogmatic philosophy, were not proponents of a philos-
ophy of common sense.

    None of these remarks is offered as a defense of Pyrrhonism or
neo-Pyrrhonism. Though misunderstandings about Pyrrhonism can be
countered, it seems to me that Pyrrhonism admits of no direct justifi-
cation. Pyrrhonism seems to have this peculiar feature: If true, it can-
not be warrantedly asserted to be true. I'll return to this point. These
opening remarks, then, are intended only to describe the viewpoint
from which this work is written. They are intended to indicate what I
have in mind by calling these reflections Pyrrhonian reflections.

    This work falls into two parts. The first is entitled "Gettier and the
Problem of Knowledge." It is an attempt to rescue our ordinary con-
cept of knowledge from those philosophers who have assigned it
burdens it cannot bear, and, in the process, have made it seem deeply
problematic. Chapters 1 and 5 present my own response to Gettier

problems; chapters 2 through 4 present a critical survey of the main positions held by others. It is possible to go directly from chapter 1 to chapter 5, although doing so will leave unexplained why I have rejected a number of seemingly plausible alternatives to my position.

The second part of this study is called "Agrippa and the Problem of Justification." According to tradition, Agrippa was a leading figure in the Pyrrhonian revival that took place between 100 B.C. and A.D. 100. His particular contribution was to systematize Pyrrhonian arguments in what came to be known as the Five Modes Leading to the Suspension of Belief. These modes were intended to present a completely general procedure for refuting any claim that a dogmatist might make. Roughly, these modes were intended to show that the dogmatists, when forced to defend their assertions, must inevitably fall into circular reasoning, an infinite regress, or reliance on an unsupported assumption.

Even though this has gone largely unnoticed, there is an uncanny resemblance between problems posed by Agrippa's Five Modes and those that contemporary epistemologists address under the heading of the *theory of justification*. The similarity is so close that I will define the philosophical problem of justification as the attempt to take seriously and then avoid the consequences of Agrippa's Five Modes. The Pyrrhonian conclusion of this work is that recent philosophical writings on justification have made no significant progress in carrying out this program. Things are now largely as Sextus Empiricus left them almost two thousand years ago.

Since this work arrives at a strongly skeptical conclusion, I have tried to be scrupulous in giving the positions under consideration a fair and full explication. For those already familiar with the works I consider or for those who have little sympathy with the standpoint they represent (or both), this close tracking of the text may sometimes seem prolix or tedious. This feeling may be reinforced by my dismissal of many of these positions on rather simple grounds. If the positions are so obviously inadequate, why tire the reader by laying them out in detail? The answer is that it is not always obvious at first that a position is obviously inadequate. Often this becomes evident only after all the resources of the position have been displayed and their inadequacies made manifest. In any case, there is a strong presumption against the claim that philosophers of acknowledged talent have simply missed the boat in solving problems of knowledge they themselves have raised. The only way to defeat this presumption is to show in detail that these theories exhibit just the defects I attribute to them.

Since this work relies in many ways on Wittgenstein's later writings—and, in particular, on my reading of them as presenting a neo-Pyrrhonian standpoint—I have included an appendix examining both the neo-Pyrrhonian and non-Pyrrhonian themes in these writings. The strength of the arguments in the body of the text does not, of course, depend on the correctness of my reading of Wittgenstein. They must stand on their own whether they correctly represent Wittgenstein's views or not. As I read Wittgenstein's later writings—particularly *On Certainty*—it seems evident that there are non-Pyrrhonian strains in his later writings that suggest a replacement of a foundationalist theory of justification with an alternative holistic theory of justification. I am also aware that this aspect of Wittgenstein's later writings makes them attractive to many of his admirers precisely because it seems to supply a *refutation* of skepticism. Given my constant reference to Wittgenstein as a neo-Pyrrhonian, it seemed important to be clear about the relationship between the skeptical and antiskeptical themes in his later writings. Since the discussion is primarily a matter of interpretation, it did not sit well in the body of the text.[11] I therefore thought it better to present it as an appendix.

Notes

1. References to Sextus Empiricus are to *Sextus Empiricus with an English Translation* (Sextus Empiricus, 1961–71). *PH* is an abbreviation of the Greek title of the *Outlines of Pyrrhonism*, i.e., *Pyrr. Hypotyposis.*

2. For an excellent discussion of these matters, see David Sedly, "The Motivations of Greek Skepticism" (1983).

3. See, for example, the discussion of the Methodic School of medicine at *PH*, 1: 236ff.

4. Barnes, 1982.

5. Burnyeat, 1980; Burnyeat, 1984.

6. Frede, 1984; Frede, 1987.

7. Frede, 1987.

8. Frede's "might just be" should be read as a cautious Pyrrhonian "is." From the surrounding context, it is clear that he endorses the interpretation he is about to give.

9. The example is from Barnes, 1982, 27 n. 90.

10. Fogelin, 1981; Fogelin, 1987b.

11. This was pointed out by the anonymous reader for the Oxford University Press.

I

GETTIER AND THE PROBLEM OF KNOWLEDGE

This page intentionally left blank

1

Gettier Problems

In this chapter I will take a fresh look at a family of epistemological problems, known as Gettier problems, that were generated by Edmund Gettier's important article, "Is Justified True Belief Knowledge?"[1] I will argue that the source of the Gettier problems has been misunderstood, and because of this, responses to them have been at best out of focus and at worst beside the point. I will also argue that once their source is correctly identified, the Gettier problems can be solved in a straightforward way. With that solution in hand, it will then be possible to understand why epistemology has taken the form it has, and why its problems have resisted, and will in all likelihood continue to resist, solution.

Gettier's Formulation

Gettier's cases are intended to provide counterexamples to biconditionals of the following kind:

> S knows that P iff   (i) P is true,
> (ii) S believes that P, and
> (iii) S is justified in believing that P.

For easy reference, I will call this the doctrine that knowledge equals justified true belief.

Before presenting two cases intended to show that justified true belief is not a sufficient condition for knowledge, Gettier notes two points:

> First, in the sense of "justified" in which S's being justified in believing that P is a necessary condition of S's knowing that P,

it is possible for a person to be justified in believing a proposition that is in fact false.

Second, for any proposition P, if S is justified in believing that P and P entails Q and S deduces Q from P and accepts Q as the result of this deduction, then S is justified in believing that Q.

Many who have responded to Gettier's challenge have targeted his second point for attack.[2] I will not adopt this strategy but, instead, concentrate on his first point.

Gettier's Reasons for Accepting His First Point

Why does Gettier think a person can be justified in believing something that is false? Gettier offers no explicit reason for accepting this claim, but, given the state of discussion at the time he wrote his essay, he presumably had something of the following sort in mind. Suppose we were to insist that in order for S to be justified in believing that P is true, S's evidence for P must entail the truth of P.[3] It has been commonly held that the immediate upshot of this ruling would be that no belief could be justified on inductive grounds, and most of what we call empirical knowledge would thereby be ruled out from the start. In Roderick Chisholm's words, "such a move would have the consequence that S knows very little about the world around him."[4]

These reflections do not, however, establish Gettier's first point, for the following theses are not equivalent:

Anti-Deductivism: A person can be justified in believing something without possessing grounds that entail it.

Gettier's First Point: A person can be justified in believing something that is false.

For all that we have seen so far, the proposition

S is justified in believing that P

may itself entail P, even though S's grounds for believing P do not entail P. Notice that such a relationship holds for propositions of the following kind:

S correctly believes that P.

If S correctly believes that P, then P is true, regardless of whether S's grounds for believing that P entail P. Later we will see that the same

implication holds for claims of the form *S is justified in believing that P*, but it will take some preparatory work to get there.

Gettier's Understanding of Justification

Since the notion of justification plays a central role in Gettier's argument, it will be important to decide how he understands it. But Gettier gives no explicit account of the sense of "justified" he has in mind. Given his silence, we can only examine the cases he uses in his attempt to refute the claim that knowledge is justified true belief, and then try to extract from them the sense of "justified" relevant to his argument. The first of his two cases will serve our purposes.[5]

CASE 1

Smith has strong evidence for the following conjunctive proposition:

(d) Jones is the man who will get the job, and Jones has ten coins in his pocket.

Smith's evidence for (d) might be that the president of the company assured him that Jones would be selected, and that he, Smith, had counted the coins in Jones's pocket ten minutes ago. Proposition (d) entails:

(e) The man who will get the job has ten coins in his pocket.

Let us suppose that Smith sees the entailment from (d) to (e) and accepts (e) on the grounds of (d), for which he has strong evidence. In this case, Smith is clearly justified in believing that (e) is true.

But imagine, further, that unknown to Smith, he himself, not Jones, will get the job. And also unknown to Smith, he himself has ten coins in his pocket. . . . It is . . . clear that Smith does not *know* that (e) is true; for (e) is true in virtue of the number of coins in Smith's pocket, while Smith does not know how many coins are in Smith's pocket, and bases his belief in (e) on a count of the coins in Jones's pocket, whom he falsely believes to be the man who will get the job.

Schematically, Smith is justified in believing a false singular proposition of the form "φa," and from this correctly draws an inference to a true proposition of the form "At least one thing is φ." Gettier offers a second example paralleling his first: Schematically, Smith is justified in believing some false proposition *P*, and from this correctly draws an inference to a true proposition of the form "*P* or *Q*."

How, exactly, in the first case, is Smith justified in believing that (e) is true? The answer, I think, is that there was nothing wrong with

his *epistemic performance:* that is, there was nothing wrong with the way in which Smith arrived at this belief. He did not form this belief using no inductive procedure—for example, he did not simply guess; nor did he use a dubious inductive procedure—for example, he did not consult bird entrails. His inductive evidence for (d) was strong, and his conscious inference from (d) to (e) was flawless. Thus there was nothing epistemically irresponsible or otherwise defective about the way in which Smith came to believe that (e) was true.

In order to fix attention on the fact that Gettier's first case is most naturally read as evaluating the process by which Smith arrived at his belief, I will translate his understanding of the claim that Smith is justified in believing that (e) is true into the explicit *adverbial* remark

> Smith *justifiably came to believe* that the man who will get the job has ten coins in his pocket.

In line with this, the third clause of the traditional definition of knowledge takes the following form:

(iii$_p$) *S* justifiably came to believe that *P*.

In making such a claim, we are indicating that *S* has been *epistemically responsible* in forming his belief. I think this is what Gettier has in mind when he speaks of *S* "being justified in believing *P*." I think this is what most philosophers who have written on this subject have had in mind as well. Although this claim should be supported by an explicit argument—and I'll give one later—on this interpretation of justified belief, it seems that *S* can be justified in believing something that is false.

What, on this interpretation of justified belief, are we to say about Gettier's counterexamples? It seems to me that when justification is taken this way his counterexamples are completely decisive against the claim that knowledge equals justified true belief. If *S* can justifiably believe something that is false, then it is easy, though it took Gettier to see this, to imagine *S* drawing an inference to a *weaker* proposition that is true, justifiably arrived at, but not knowledge.

A Second Interpretation of Justification

There is, however, another way we can interpret the claim that *S* is justified in believing that *P*. The leading idea is that in saying *S* is justified in believing that *P*, we are not assessing the procedures *S* used in coming to his belief that *P*; we are, instead, evaluating the adequacy of *S*'s grounds for establishing the truth of *P*. In saying that

S is justified in believing that P is true, we are saying that the grounds on which S accepts P establish the truth of P. This suggests a second way of interpreting the justification clause in the traditional account of knowledge. I'll call the clause read this way the (iii_g) clause:

(iii_g) S's grounds establish the truth of P.

Note that in saying that S's grounds establish the truth of P, we are speaking about a relationship between a *proposition* that S accepts and the *grounds* on which he accepts it. The (iii_g) clause offers an assessment of S's reasons (or grounds), indicating that they are adequate to establish the truth of a certain proposition.

To appreciate the force of the idea that in making third-person epistemic judgments we are involved in an assessment of grounds and not simply commenting on a performance, it will help to imagine ourselves actually investigating Gettier's claims. That is, Gettier actually appears before us and tells us that there was a Smith, a Jones, and a president, and further tells us that they carried on in the way he describes. Treating the story as a real story makes a significant difference in how we will deal with it. In particular, none of the items in the story is *privileged* in the sense of being immune to epistemic evaluation. For example, as genuinely engaged, we have to decide whether to trust Gettier, for the story is, after all, rather strange. Perhaps he has left out important details. Perhaps he has the story wrong. Perhaps he made it up. These are matters that would naturally concern us in an actual situation. Noting them here helps bring us and our activity of assessing grounds into the picture.

To avoid the complexities of a Gettier case, we can first suppose that our investigations reveal that Gettier actually has the facts wrong. Contrary to what he has told us, things proceeded quite normally: the president was his usual reliable self, and Jones, who Smith somehow determined had ten coins in his pocket, received the promotion. (When we confront Gettier with these facts, he admits that he distorted the story as a hoax having something to do with tenure.) Once we have discovered all this, what shall we say about S's belief that someone with ten coins in his pocket would get the promotion? First, it remains true that Smith was justified in coming to believe what he did; his performance was epistemically responsible. Beyond this, in this non-Gettier version of the story, we will further judge that his *grounds* were adequate to establish the truth of what he believes. In saying that Smith was justified in believing what he did, we express our *agreement* with Smith.[6] In this sense of "justified," justification claims, as opposed to claims concerning whether

someone has reasoned justifiably, commit *us* to the truth of what is said to be justified.[7] Furthermore, with some free play, we commit ourselves to this belief on the *same* grounds that commit Smith to it.

These reflections on this non-Gettier example may suggest that we read the clause

(iii) S is justified in believing that *P*

as saying

(iii$_g$) S's grounds (or reasons) establish the truth of *P*.

But for reasons that will emerge later, we will not want to abandon what I have called the (iii$_p$) clause in favor of a (iii$_g$) clause, for both clauses are needed in a plausible account of knowledge claims. The result is that the justification clause in the traditional doctrine that knowledge equals justified true belief is now seen to have two components. The first concerns the manner in which S came to adopt a belief. This is the (iii$_p$) clause, which demands that he do this in an epistemically responsible manner. The second concerns a relationship between the proposition believed and the grounds on which it is believed. This is the (iii$_g$) clause, which demands that these grounds establish the truth of the proposition believed on their basis.

Having developed the notion of a (iii$_g$) clause through examining a non-Gettier example, we can now apply it to Gettier's first counterexample to the traditional account of knowledge. It should be obvious that its inclusion immediately disposes of this counterexample. To see this, we need only recall that Gettier's argument depended crucially on the truth of (d):

> Smith is justified in believing that Jones is the man who will get the job, and Jones has ten coins in his pocket.

As noted, if this is interpreted as a remark about the propriety of Smith's performance, then the claim is true, and if the traditional analysis of knowledge demands no more than this for a belief to be justified, then Gettier's counterexample is decisive against it. If, on the other hand, it is interpreted as demanding—at least in part—that

> Smith's grounds justify the belief that Jones is the man who will get the job, and Jones has ten coins in his pocket

then the claim is not true, and Gettier's counterexample fails. I claim that in every version of the Gettier problems we will find this same situation: S will be justified in his belief, having come to it in a

responsible manner; that is, his performance will satisfy the (iii$_p$) clause of the amended biconditional. At the same time, his grounds will not establish the truth of what he believes, so the (iii$_g$) clause will not be satisfied. The first fact inclines us to say that S is justified in his belief, the second leads us to deny him knowledge even though his belief was justifiably arrived at and true.

Isn't it, however, incoherent to grant that S has acted with epistemic responsibility in coming to believe P yet at the same time deny that his grounds establish the truth of that proposition? If S believes something on grounds that do not establish its truth, isn't that, by itself, enough to show that he has formed his belief irresponsibly? In general the answer to that question is yes, and it is for this reason that the grounds clause is easily obscured by the performance clause. The two clauses do, however, fall apart, that is, take different truth values, when we, who are trying to decide whether S knows or not, have access to a wider range of information than S does. In such a situation we may sometimes grant that S has formed his belief responsibly even though his grounds do not establish the truth of what he has come to believe. This situation characterizes Gettier problems in *all* their manifestations, for they provide cases where S has justifiably (i.e., responsibly) come to believe some truth on grounds that do not establish this truth. To see in detail how this is possible, we must examine some of the basic features of so-called inductive reasoning.

The Gettier Problems and Nonmonotonicity

Our question is this: How can a person accept a belief on grounds that do not establish its truth, yet not be epistemically irresponsible in doing so? I think we can answer this question, and get to the heart of the Gettier problems, by reflecting on a fundamental way in which inductive inferences differ from deductive inferences. It is commonly noted that in a valid deductive argument the truth of the premises necessitates the truth of the conclusion, whereas with inductive arguments this is not so. This is right, but insisting on this difference can place things in the wrong light. The Gettier problems do not, after all, depend on deductive chauvinism—the claim that the only good argument is a deductively sound argument. For that matter, they do not depend on the device, sometimes attributed to skeptics, of introducing artificially high inductive standards and then rejecting all knowledge claims that fail to meet them. In Gettier examples, we allow S to use ordinary inductive procedures in

normal ways, and in evaluating these cases we do the same. To suppose otherwise is to rob Gettier examples of their real significance.

For our purposes, it is useful to describe the difference between deductive and inductive reasoning from the following perspective: Adding further premises cannot convert a valid deductive argument into an invalid one, whereas adding further premises can degrade a strong inductive argument into a weak one. Acquiring additional information can lead us to revise our opinion about the truth of a premise of a deductive argument, and then declare that an argument that we previously considered sound is, in fact, unsound.[8] With an inductive inference, however, additional information can lead us to revise downward our original assessment of its strength without revising our opinion concerning the truth of *any* previously accepted premise. This difference is sometimes marked by saying that deductive inferences are *monotonic* whereas inductive inferences are *non-monotonic*.[9] Gettier-like situations arise, I suggest, because inductive inferences are nonmonotonic: it is this that drives them.

More specifically, it is because of the nonmonotonicity of inductive inferences that it is possible for someone to be epistemically responsible in coming to believe something on grounds that do not establish its truth. Given the information available to *S*, it might be completely responsible for him to suppose that his grounds establish the truth of what he believes on their basis. He uses a standard procedure in a standard way, and nothing in the context suggests that he should not. Indeed, it might be irresponsible for him to show any more care than he does. There is, after all, something called being too careful. We, however, given a broader range of information, can see that his justificatory grounds are, in one way or another, undercut, though not in a way that *S* could be expected to recognize or take provision against. My suggestion, then, is that the central feature of Gettier's original examples is this:

(a) Given a certain body of information, our subject *S*, using some standard procedure, justifiably comes to believe that a proposition, *h*, is true.

(b) We are given wider information than *S* possesses, and in virtue of this wider information see that *S*'s grounds, though responsibly invoked, do not justify *h*.

Information Possessed by Us ⟶ ◯ ⟵ Information Possessed by *S*
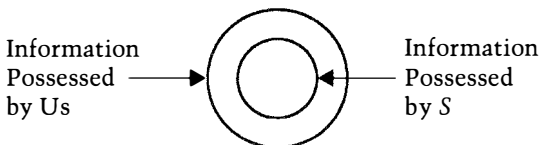
I think this double informational setting—this informational mismatch between the evidence S is given and the evidence we are given—lies at the heart of Gettier problems. It is this informational mismatch that inclines us to say, quite correctly, that S justifiably came to believe something true, yet at the same time deny him knowledge because, as we see, his grounds do not justify this claim.

If all this is correct, then it should be clear that the Gettier counterexamples have no force against the doctrine that knowledge is justified true belief when justification has at least the force of what I have called the (iii_g) clause. Furthermore, I think the same diagnosis holds for all the many intricate variations of the Gettier problems. They all depend on an informational mismatch that allows the (iii_p) and (iii_g) clauses to fall apart. I will sketch this claim in a general way next, and defend it in more detail in succeeding chapters.

Variations on the Gettier Problems

Gettier's original examples have had offspring,[10] many of them narrowly tailored to refute specific analyses of knowledge. Without worrying about fine details, I will divide Gettier-like counterexamples into two broad categories:

> I. Those that employ a normally sound justificatory procedure in a context where it is not, in fact, reliable, then arrive at something true by drawing a conclusion *weaker* than normally warranted by this procedure.

> II. Those that employ a normally sound justificatory procedure in a context where it is not, in fact, reliable, then arrive at a normal *strong* true conclusion by good fortune.

For ease of reference, I'll call examples that fall into the first category *weakening-inference* examples, and those that fall into the second category *epistemic-luck* examples. These labels are not entirely apt since, in a sense, all Gettier examples involve epistemic luck, but they should serve well enough to mark the salient features of these two sorts of Gettier examples.

Gettier's original cases fall paradigmatically into the first category, since they involve an inference from a justified false belief to a weaker belief that turns out to be true. In reply, critics responded, quite reasonably, that one cannot justify something on the basis of a belief that is false. To get around this, a new set of counterexamples appeared that were like Gettier's except that S does not rely on any

premise that is false. Thus the first category can be divided into two subcategories:

Ia. On the basis of a normally sound justificatory procedure, S arrives at a justified but false belief and then arrives at something true by explicitly drawing a conclusion *weaker* than normally warranted by this procedure.

Ib. On the basis of a normally sound justificatory procedure, S arrives at something true by drawing a conclusion *weaker* than normally warranted by this procedure, without, however, explicitly relying on a premise that is false.

As far as I know, Keith Lehrer was the first writer to present a clear example of this second kind. It will be helpful to develop his example in stages.

In "Knowledge, Truth, and Evidence,"[11] Lehrer introduced a Gettier-type example that has enjoyed a robust, if ever-changing, career.

> I see two men enter my office whom I know to be Mr. Nogot and Mr. Havit. I have just seen Mr. Nogot depart from a Ford, and he tells me that he has just purchased the car. Indeed, he shows me a certificate that states that he owns a Ford. Moreover, Mr. Nogot is a friend of mine whom I know to be honest and reliable. On the basis of this evidence, I would be completely justified in believing:
>   PI: Mr. Nogot, who is in my office, owns a Ford.
> I might deduce from this that
>   H: Someone in my office owns a Ford.

Since this is a Gettier story, not surprisingly, it turns out that Mr. Nogot does not own a Ford, whereas Mr. Havit (though I have no reason to suppose this) does. Thus I justifiably believe *H*, yet I do not know it. Since I have reached this result by drawing an inference from a false premise (*PI*), we have a standard, category Ia, Gettier counterexample.

Lehrer moved to a category Ib counterexample by having S (to return to the third person) draw the inference *directly* from the evidence to the conclusion *H* without using *PI* as an intermediate step.[12] In this way, no false premise plays a role in the argument. In *Knowledge*[13] Lehrer presented an even more elaborate example that has come to be known as the Clever Reasoner. All the facts are the same as before except for the following new wrinkle: Being clever and cautious, S reflects upon the fact that someone else in the office just might own a Ford, so to increase his chances of being right, he consciously retreats to the weaker claim, *H*, that *someone* in his office owns a Ford.[14]

Why, with this new case, do we feel so strongly, indeed just as strongly as we did with earlier category Ia cases, that *S* does not know that someone in his office owns a Ford? He certainly justifiably came to believe this, for he has just the sort of evidence that we normally use in justifying such a claim. (Indeed, from an everyday standpoint, his evidence is more than we expect. We don't usually wait for people to show us certificates of ownership before we believe they own a car.) I can see no other reason for our denying that Smith knows that *H* is true except that we will not grant that his grounds (in the context of other things we have been told) *established* the claim that someone in the office owns a Ford. In sum, we can deal with category Ib examples in precisely the way in which we dealt with the original Gettier example: We deny that Smith is justified in holding a certain belief because we see (even if he cannot) that his *grounds* do not adequately justify his belief.

These same points can be made, perhaps more forcefully, by turning to a counterexample that falls into category II—involving what I have called epistemic luck. The following comes from Ernest Sosa:

> Out for a drive in the country, I see a barn nearby standing out from its surroundings in sharp focus, and thus perceptually acquire the fully justified belief that it is a barn. In both directions along my road there are numerous barn facsimiles, however, mere shells presenting to the road a facade that would draw an attribution of barnhood no less justified than my own when by luck I happened to take notice of the one real barn in the area. Surely my justified true belief that I see a barn is not knowledge.[15]

I share Sosa's intuitions concerning this particular example: In these circumstances I do not know that there is a barn before me. I do not know this because the further information that the road is lined with facsimile barns shows that you cannot tell, as you usually can, that something is a barn just by looking. The acquisition of new information does not, however, lead to me say that I formed my original belief irresponsibly. The new information shows me that my original grounds, though responsibly invoked, did not establish that the object I was looking at was a barn.

This last Gettier variation differs from those that fall into category I, for here a robust (as opposed to a weak) conclusion is drawn on the basis of a procedure inadequate for the circumstances, and then, by the grace of nature, it turns out to be right. From these reflections we see that reliance on a false (though justified) premise is not essential to Gettier counterexamples. (Category Ib examples show this.)

Nor is the retreat to a weaker claim essential. (Category II examples show this.) For these reasons, narrow diagnoses of these particular cases will not serve our general purposes. As far as I can see, however, the underlying reason that we refuse to say that S's belief amounts to knowledge in *all* these cases is that we refuse to acknowledge that his grounds establish the truth of the claim he makes. In each case we have a violation of the (iii$_g$) clause in the revised traditional definition.

Epistemic Responsibility

If I am right in suggesting that the Gettier counterexamples depend on taking the notion of justification the wrong way, then it is tempting to think that taking it the right way will yield an acceptable account of knowledge. As it turns out, however, it seems that the *adequate-grounds* reading of the justification clause is not itself sufficient to yield an acceptable account of knowledge.

First, problems arise from what I think of as BonJour counterexamples.[16] They concern epistemic irresponsibility. The following is not from BonJour, but is in his style. Back to Gettier's original example:

> Suppose that Smith has grounds sufficient to justify the belief that
> (e), "A person with ten coins in his pocket will get the promotion,"
> is true. This time there is no Gettier funnybusiness: (e) is in fact
> true, and Smith commands grounds that establish the truth of (e).
> Furthermore, on the basis of these grounds he infers that (e) is true.
> Beyond this, however, Smith also sincerely holds the false belief
> that the president is a pathological liar and can never be trusted on
> any subject. We can suppose that he has very good reason to believe
> this. Considering this fact, Smith then reasons quite *badly* (or *stu-*
> *pidly* or *incorrectly*) that in this particular case the president is not
> lying. Thus, given his full set of beliefs, Smith ought not think him-
> self justified in believing (e), but he does so nonetheless. In this sit-
> uation, even though he possesses reasons that justify (e), I do not
> think we want to say that Smith knows that (e) is true. People do
> not know things that, given other things they believe, they ought
> not believe.

This situation is, in some ways, the reverse of Gettier situations. In the Gettier examples, we pictured Smith basing his knowledge claim on a batch of evidence that would normally justify the kind of claim Smith has made. We then supposed *ourselves* in the possession of some further evidence that degrades the evidential support on which Smith bases his belief. The wider set of beliefs we accept includes the evidential beliefs he accepts, but the evidential force of

his beliefs is degraded in our wider framework. In this new case, we imagine that Smith commands evidence that we accept as wholly adequate to justify the proposition (e), but beyond this we imagine that he also accepts as evidence propositions we do not accept and which, if true, would degrade the evidential support given by the propositions we do accept. In a case like this it seems wrong to say that Smith knows that (e) is true, for, turning things around, although Smith has grounds that justify the claim that (e) is true, he does not justifiably believe this. It seems, then, that a reference to justified performance, as well as a reference to justifying grounds, is (at least sometimes, or for certain kinds of knowledge) an essential ingredient in an account of knowledge as justified true belief.

A demand for what I have called a (iii$_p$) clause arises for other reasons as well. For example, we expect S's reasoning from his grounds to be non-fallacious. We will not grant that S knows that P if, though in possession of adequate grounds for P, he reasons badly to the conclusion that P.[17] Another reason for including a performance clause is that we do not want to say that a person knows everything entailed by the true things he believes. To cite a familiar example, a person who knows all of Euclid's axioms, definitions, and so forth, does not thereby know all the theorems of Euclidean geometry. In order to know one of these theorems to be true, a person must *base* her belief on the axioms, definitions, and previously proved theorems, and do this correctly.

These reflections suggest, as a first approximation, that we split the third clause of the original biconditional into two parts:

S knows that P iff   (i) P is true,
(ii) S believes that P,
(iii$_p$) S justifiably came to believe that P, and
(iii$_g$) S's grounds establish the truth of P.

But this is inelegant, first because the adverbial phrasing in (iii$_p$) makes the second clause redundant. Second, the fourth clause renders the first clause redundant. And this is important, because it reflects the fact that truth is not simply an independent feature of knowledge. I'll come back to this.

Finally, the formulation as it stands does not reflect a thesis central to this discussion: In saying that S knows that P, we are stating— perhaps with some free play—that the grounds that S takes as establishing the truth of P do establish its truth. That is, the (iii$_g$) clause has as its target S's grounds as referred to in the (iii$_p$) clause. In order

to keep the two justification clauses from drifting apart, we can com-
bine them and, dropping the two redundant clauses, arrive at the fol-
lowing biconditional:

 S knows that P iff   S justifiably came to believe that P on
                       grounds that establish the truth
                       of P.

If a slogan is needed: Knowledge is not simply justified true belief,
but, in the cadences of Woodrow Wilson, it is justified true belief,
justifiably arrived at.[18]

Conclusion

If this analysis is right, then I think we can understand why
responses to Gettier problems have taken the forms they have. My
general thesis is that all of these responses are misunderstood
attempts to find a surrogate for the grounds-evaluating (iii$_g$) clause in
the analysis I have presented. These positions seem persuasive to the
extent that they approximate this goal, and with sufficient ingenuity,
this approximation can be made quite close. More deeply, however,
in their attempts to find a surrogate for the grounds-evaluating aspect
of epistemic judgments, they are wrongheaded—wrongheaded in the
way that theories that commit the naturalistic fallacy are wrong-
headed. In the next three chapters I shall try to show this in detail.

Notes

1. Gettier, 1963. I speak in the plural about Gettier problems in order to
encompass the large and loose family of counterexamples spawned by the
Gettier originals. For a detailed survey of the various forms that the Gettier
problems have taken, together with an assessment of the various answers to
them, see Shope, 1983.

2. It can be noted in passing that the stated version of the principle of
closure for knowledge is wrong as it stands, since even though P entails Q,
S might deduce Q from P in some invalid way. In that case, I do not think
we would say S is justified in believing that Q. This difficulty is easily fixed
by insisting that S *correctly* deduces Q from P.

3. As far as I know, Robert Almeder is the only person involved in this
discussion to adopt this line (Almeder, 1974; Almeder, 1976). I cannot, how-
ever, tell from these essays how Almeder proposes to deal with inductively
based knowledge claims.

4. Chisholm, 1986, 248.

5. Many of the points I make concerning the interpretation of the

expression "*S* is justified in believing that *P*" were developed in detail by Catherine Lowy (1978). I repeat them here, first, because they are worth repeating, and second, because they will play a central role in an argument with a conclusion fundamentally opposed to the conclusion that Lowy herself drew from them.

6. Sometimes we can legitimately grant that Smith's grounds justify the belief that *P* is true without being able to produce these grounds ourselves. This happens, for example, with appeals to expert authority.

7. The position adopted here parallels one taken by Fred Dretske in his essay "Conclusive Reasons" (1971a). I advanced essentially the same position, though without application to Gettier problems, in *Evidence and Meaning* (1967). In *On Certainty*, Wittgenstein also comes close to adopting this line as well. For example:

> *OC*, 243. One says "I know" when one is ready to give compelling grounds. "I know" relates to the possibility of demonstrating the truth.

From the context it is clear that Wittgenstein is not thinking of *deductive* demonstrations.

8. Here I follow the now common convention of calling a deductive argument sound if it is valid and has true premises.

9. These terms were introduced on the basis of a not altogether apt analogy with (positive) monotonic functions that can be presented as constantly rising (or at least nonfalling) curves, as opposed to those nonmonotonic functions whose curves go both up and down.

10. Robert K. Shope (1983) has actually catalogued *ninety-eight* examples relevant to Gettier problems.

11. Lehrer, 1965.

12. Lehrer, 1970.

13. Lehrer, 1974.

14. Perhaps even more clearly, in "An Alleged Defect in Gettier Counterexamples," Richard Feldman produces the following specimen of this kind of example:

> We can alter the example slightly, however, so that what justifies *h* for Smith is true and he knows that it is. Suppose he deduces from [the original statement of the evidence] its existential generalization:
>
> > *(n)* There is someone in the office who told Smith that he owns a Ford and even shows him a certificate to that effect, and who up till now has always been reliable and honest in his dealings with Smith.
>
> *(n)*, we should note, is true and Smith knows that it is, since he has correctly deduced it from *m*, which he knows to be true. On the basis of *n* Smith believes *h*—someone in the office owns a Ford. (Feldman, 1974, 253)

Again we arrive at a category Ib Gettier result. Brian Skyrms has an example of much the same kind as Lehrer's Clever Reasoner (Skyrms, 1967).

15. Ernest Sosa, "Epistemic Presuppositions," in Pappas, 1979, 84.

16. BonJour's counterexamples are aimed primarily against so-called externalist or reliabilist accounts of knowledge. In chapter 3 I will argue that these counterexamples do not have the full force that BonJour attributes to them. They do, however, have important consequences for the analysis of certain kinds of knowledge. For a budget of BonJour counterexamples see BonJour, 1985, 37ff.

17. I owe this point to Michael Bratman.

18. Cf.: "Open covenants of peace openly arrived at."

2

Fourth-Clause Theories

Analyses of Knowledge

At the close of the previous chapter, I suggested that those who have tried to solve Gettier problems have been engaged in the misunderstood attempt to find a surrogate for the grounds-evaluating (iii_g) clause that must form part of a correct account of knowledge claims. This program can be developed in two ways. First, we might hold S to stricter standards of epistemic responsibility. We can demand that S behave as a suitably ideal gatherer of evidence. At the extreme, we might demand that to be epistemically responsible in believing that h, S must withhold judgment until he is apprised of *all* information relevant to the truth of h. This demand for total relevant evidence will eliminate the threat of Gettier situations by blocking the informational mismatches that generate them. The result of this approach, however, is that we must give up most of our ordinary claims to inductively based knowledge. A solution along these lines yields what might be called a *functional equivalent* of a strong version of skepticism. In the process of safeguarding empirical knowledge claims against the threats posed by Gettier situations, we concede to the skeptic that we do not really know a great many of the things that we commonly think we do know.[1] Because of this threat of skepticism, those who have attempted to solve Gettier problems have rarely attempted to achieve this through significantly raising standards of epistemic responsibility.

If, because of the threat of skepticism, demanding more of S's epistemic performance will not serve our purposes, then it seems that the only alternative open to those who are trying to preserve the core idea of the traditional account of knowledge is to place *side-constraints* on

the occasions on which a normally adequate procedure is employed such that the informational mismatches that yield Gettier problems cannot arise. That is, instead of asking anything more of S, a *fourth* clause is introduced specifying circumstances such that when S justifiably comes to believe something true in those circumstances, then S's belief cannot be denied the status of knowledge—at least for Gettier-like reasons. Since to know h, S need not know or even suppose that he is happily situated in this way, I'll call clauses of this kind *external* clauses. This, then, is the pattern for a so-called *fourth-clause* analysis of knowledge: S will be said to know that P if (i) P is true, (ii) S believes that P, (iii) S justifiably came to believe that P, and (iv) S forms his belief in, as we might say, a Gettier-free zone. The task of the fourth clause, then, is to lay down external side-constraints that situate S in a manner that protects him from Gettier counterexamples. To repeat, these constraints are external in the sense that S need know nothing of them in justifiably coming to believe something.

Given this distinction between internal clauses and external clauses, we now have a convenient way of sorting out and understanding various analyses of knowledge. I will not follow the common practice of simply dividing analyses of knowledge into internalist and externalist positions. This classification strikes me as being too crude, since some analyses—including those examined in this chapter—contain both kinds of clauses. Instead, if we begin with the base assumption that *knowledge is at least true belief*,[2] then analyses of knowledge can be classified in terms of the additional internal or external clauses they invoke:

	Internal Clause	No Internal Clause
External Clause	(1) Fourth-Clause Theories (Indefeasibility)	(2) Pure Causal and Subjunctive Theories (Goldman, Dretske)
No External Clause	(3) The Original Definition As Usually Interpreted	(4) Knowledge Equals True Belief

In this chapter I will examine the leading example of an analysis that includes both an internal clause and an external clause—so-called indefeasibility analysis. In the two succeeding chapters, I will examine analyses that drop the internalist (epistemic responsibility) clause altogether and rely completely on an external clause to block Gettier-like problems.

Indefeasibility Theories

In examining the traditional doctrine that knowledge is justified true belief, we have encountered cases where S, quite justifiably, bases his judgment that d is true on evidence $e_1, \ldots, e_n$. We, however, from our privileged position, are aware of some additional truth p (perhaps $\sim d$) that degrades or defeats this justification. We thus refuse to admit that S's grounds establish the truth of d, and for this reason will not say that S knows that d. If this correctly diagnoses the problem that Gettier examples raise for the traditional definition of knowledge, then it would seem that the most straightforward way to avoid this problem is to add a clause that specifically excludes this possibility. This has been attempted by a number of philosophers in a variety of ways, but the best-known and most fully elaborated example of this approach is found in analyses that supplement the traditional definition of knowledge with what have come to be known as *indefeasibility clauses.* In what follows I shall examine in some detail the ways in which indefeasibility theorists can be driven by counterexamples to add further and further qualifications to their initially plausible position. In examining this standpoint, I will interleave examples and arguments from various writers including Lehrer, Paxson, Swain, and Pappas, but I will use Marshall Swain's clear and subtle essay "Epistemic Defeasibility" to give structure to the discussion.[3]

Beginning in a standard way, Swain offers the following schema for the analysis of knowledge:

> S knows that h iff (i) h is true, (ii) S is justified in believing that h (that is, there is a true body of evidence e such that S is justified in believing e and e justifies h),[4] (iii) S believes that h on the basis of his justification and (iv) S's justification for h is indefeasible (that is, . . .). (163)

Swain's task is to give a specification of indefeasibility by filling in the *that-is* clause left blank in the fourth condition.

After a first attempt that is obviously too strong, Swain offers the following suggestion:

> (ivb) There is no true body of evidence e' such that the conjunction of e and e' fails to justify h. (164)

I think this is precisely what an indefeasibility clause should say, but Swain and others find reasons to move further and further away from such a simple formulation. I will follow this development through various stages.

Stage One

The first counterexample to the simple indefeasibility clause just
stated has come in various forms. Here's a version from Pappas and
Swain, who are restating an example from Lehrer and Paxson. I'll
label it the Mad Mrs. Grabit.[5]

> S sees a man named Tom Grabit steal a book from a library. We
> may suppose that S has whatever justification we like for believing
> that Grabit stole the book. However, we suppose that, entirely
> unknown to S, Tom Grabit's mother has said that Tom was not in
> the library, but that Tom's twin brother Tim was in the library. Let
> q be the statement that Tom Grabit's mother has said these things.
> This statement is such that if S came to have it as additional evi-
> dence, then S's justification would be lost. (29)

So far so good. Now the Gettier twist:

> But suppose that Tom Grabit's mother is demented, that Tom has
> no twin brother, and that the mother's remarks are thus completely
> wrong, all of which is unknown to S. (29)

Under these circumstances Pappas and Swain suggest that

> S can still be said to know that Tom Grabit stole the book, even
> though there is a true statement q that would sully the justifica-
> tion. (29)[6]

So it seems that our first indefeasibility clause is too strong and
some modification is needed.

   This task is given sharper focus if we ask, following Lehrer and
Paxson, how we are to distinguish this new case, where it seems that
a justification has not been defeated, from the apparently similar
Nogot case (discussed in the previous chapter), where the addition of
a further piece of information (that Nogot had no car) does under-
mine a justification. Lehrer and Paxson attempt to separate the two
cases in the following way: "In one case [i.e., the Nogot example] my
justification depends on my being completely justified in believing
the true statement to be false while in the other [i.e., the Grabit
example] it does not" (151).[7] In the Mad Mrs. Grabit example, S pos-
sesses no evidence that would justifiably lead him to believe that
Mrs. Grabit did not say the things she did. As the story is told, he has
no beliefs concerning a Mrs. Grabit at all, and thus no basis for
believing that something true of her is false. In the Nogot example,
on the other hand, S has evidence that justifies him in believing that
the true proposition—Nogot does not have a Ford—is false. Further-

more, if this true proposition were added to S's original evidence, this original evidence would be undercut.

Swain, agreeing with this diagnosis, adds a further qualification to (ivb) yielding: "(ivc) . . . there is no true body of evidence e' such that (a) e justifies S in believing that e' is false and (b) the conjunction of e and e' fails to justify h" (166).

Stage Two

This amended clause seems to allow us to distinguish the Nogot case from the Mad Mrs. Grabit case, allowing us to deny knowledge in the first case and, it seems, grant it in the second. But Swain thinks this new version of the indefeasibility clause has troubles of its own because of the following purported counterexample to (ivc). Roughly, a person sees a heavy rock heading toward a window and is, let's say, justified on this evidence e in believing that the rock will break the window.

Now the Gettier twist. On this same evidence, S is also justified in believing that he is about to hear and see the window break. Suppose, however, that, unknown to S, he will suffer a sudden physiological disorder and because of this will neither hear nor see the window break. That he will neither see nor hear the window break we will call e'. Then S's evidence seems to justify a falsehood (i.e., that he will see and hear the window break), and furthermore, the truth that he will not see and hear the window break, if added to his former evidence, will defeat his original justification. That is, his not hearing and not seeing the window break constitute evidence that will override his evidence in favor of the claim that it will break.

Skipping over technical details, we can see how Swain attempts to deal with this counterexample by introducing the notion of a person who is epistemically "ideally situated." "We can give an intuitive characterization of what is wrong when a man's justification is defeasible by saying that he is less than 'ideally situated' with respect to the evidence bearing upon h. The notion of being ideally situated is, of course, a pipe dream. In any full sense of the term, anyone who falls short of omniscience is thereby less than ideally situated" (168). But even if the notion of being ideally situated is a pipe dream, according to Swain it still makes sense to speak of a person's epistemic position as being more or less close to being ideal:

> From a given epistemic position a man might approach such an ideal state to any of a variety of degrees. He might pick up a bit of information here, and a bit there, in piecemeal fashion. Ironically,

a man will sometimes wind up in a position that is worse than the
one from which he began, even though he has moved closer to an
ideally situated position by acquiring some new information. . . .
The acquisition of such a limited portion of available new evidence
is, we might say, one extreme on the continuum of ways in which
a man could move toward a more ideally situated position. The
other extreme is the acquisition of all the additional information,
but this extreme is an epistemologist's pipe dream. We need some-
thing in between. (172–73)

After a number of intermediate steps, Swain arrives at the fol-
lowing suggestion for restricting the range of potentially defeating
evidence. It will be "a complete description of how things might
change for S if his body of evidence were purged of all false members,
each of these being replaced by its denial" (176). If we think that, to
the extent that a person has false beliefs, that person is benighted, we
can say that one who has had his false beliefs replaced by true ones
has become unbenighted. We thus arrive at a new version of clause
(iv) that captures, I believe, the spirit of Swain's proposal minus its
formal trappings:

(ivf) S would be completely justified in believing h even if he
becomes unbenighted.[8]

This clause seems to sort out the problematic cases in the way
that Swain wants: S does not know that someone in his office has a
Ford, for, unbenighted with respect to Nogot's misleading activities,
S's justification is defeated. However, S does know that Tom Grabit
stole the book since, having no views (true or false) concerning the
mad Mrs. Grabit, his justification is not defeated under unbenight-
edness. Finally, and this is Swain's nice point, S does know that the
window will break, for unbenighted he will understand why he nei-
ther sees nor hears it break and these facts will no longer defeat his
justification. In Swain's words, "the defeating effect of the available
counterevidence is in turn overridden" (180).

Stage Three

But, almost unbelievably, there are still difficulties. Swain notices
that S, when unbenighted, may come into possession of a new set of
true beliefs that may provide him with an entirely *new* justificatory
argument that turns out to be indefeasible.[9] To guard against this,
Swain tells us that "we need to require that a justification be inde-
feasible only if the [original] justification *itself* would continue to

hold under the process of becoming more ideally situated epistemically" (180–81).

I will not examine Swain's technically complex effort to capture this intuition in a formally correct statement of the indefeasibility clause. Informally, our final version of the indefeasibility clause, (ivg), might read as follows:

> (ivg) Except for some minor adjustments, the same evidence that justifies *S* in believing that *h* will continue to justify him in believing that *h* even after he becomes unbenighted.

As I read these passages, a vacillation between an adequate-grounds interpretation of justification and an epistemic-responsibility interpretation of justification strikes me as transparent. To see this, we can go back to the first version of an indefeasibility clause I cited from Swain:

> (ivb) There is no true body of evidence *e'* such that the conjunction of *e* and *e'* fails to justify *h*. (164)

Swain and others were led to abandon such a simple formulation of the indefeasibility clause because of concerns over Mad Mrs. Grabit.[10] I think he and other writers were misled by this example. Again, I think we see things more clearly if we abandon our position as privileged spectators and place ourselves in the context of the actual epistemic evaluation. As the story is told, *S* has a strong justification of a standard type for the belief that Tom Grabit stole the book. Thus, if we were placed as *S* was placed, we would judge the matter as *S* judged it. Concerning *S*, we would then say he possessed a justified true belief, justifiably arrived at. But suppose we heard, as *S* didn't, that Tom's mother, someone presumably in a position to know, had attributed the theft to a twin brother named Tim. Suppose that's all we hear. (And remember, we do not enjoy a privileged perspective that gives us the inside information about who really committed the theft.) Here, I think, we might continue to say that *S* was epistemically responsible in coming to believe that Tom stole the book, but we would no longer say that his grounds justify this claim. This is a practical matter, not a philosophical one. The reported existence of a twin presents a specific challenge to an eyewitness identification, which must be dealt with before this issue can be settled. To use a notion that will be developed more fully in chapter 5, the report of a possible twin triggers a higher level of scrutiny that must be met before we can have a justified belief concerning who stole the book.

Furthermore, with this change in epistemic standards, we will no longer *endorse* S's grounds as justifying his belief and no longer be willing to say that he knows that Tom Grabit stole the book.

Finally, suppose we follow our inquiries further and discover the sad story of Mad Mrs. Grabit's dementia. Now what are we to say? First, we might again feel content to say that *we* know who stole the book: Tom Grabit. Furthermore, we will have no reason to go back on our previous judgment that S justifiably came to believe that Tom stole the book. But having learned all we have learned, do we want to say, as Swain and others seem comfortable saying, that S *knew* that Tom stole the book? I confess that once I imagine myself actually engaged in the inquiry—getting information bit by bit—my intuitions tend to flip-flop. We, in our investigations, had to deal with a potentially defeating piece of evidence before we could think ourselves justified in claiming that Tom Grabit stole the book. Until we have completed our investigation, we certainly will not say that S knew that Tom Grabit stole the book, for as long as the Mrs. Grabit challenge remains unanswered, his grounds (and our grounds) are not adequate to establish that claim.

What do we say about S after this challenge has been met? If we are inclined to think that Mrs. Grabit's remark simply raised a difficulty for S's justification, then I think that once this difficulty is met we will be inclined to say that his grounds were justifying, and, indeed, he did know that Tom Grabit stole the book. This, I think, is the natural view of a spectator who is given all the relevant information at once. On the other hand, if we have actually been involved in the investigation, and have gone to some trouble—perhaps a great deal of trouble—to check the veracity of Mrs. Grabit's remark, then I think that our attitude toward S's grounds (and hence his knowledge) might be quite different. Instead of thinking of the facts vindicating S's justification, we may think of ourselves as replacing his justification with a better one of our own. In this case, I do not think we will be inclined to say that S knew that Tom Grabit stole the book, because we will not be inclined to say that S's grounds established the truth of this claim. I find that my intuitions alternate—duck/rabbit fashion—as I shift perspectives in this way.

To go back a bit, responses to the Gettier problems began by examining S's actual reasoning. It seemed right to put certain restraints on how S reasoned for this reasoning to count as justified. For example, it seemed right, as many insisted against Gettier's original examples, that S's reasoning not depend on the acceptance of falsehoods. But when the second wave of Gettier-like counterexamples appeared, it became clear that such a restriction, at least in a simple form, was

inadequate. It is at this point that indefeasibility theorists make a sub-
tle but profound change in tactics. Instead of placing further restric-
tions on S's *actual* performance, indefeasibility theorists attempt to
situate S in such a manner that any justification he accepts will be
acceptable to us. In particular, Swain's strategy of unbenighting S is an
ingenious attempt to guarantee that we will endorse S's reasoning, for
unbenighting him eliminates just those possibilities that previously
led us (at least in the examples under consideration) to withhold
endorsement. In a sense, the indefeasibility theorists bestow on S a
larger and larger share of those protections that *we* enjoy as fully
informed external spectators of Gettier situations. The critics of inde-
feasibility theories, in their turn, exploit this privileged spectatorial
position to find ever more subtle mismatches between the information
that we, as spectators, possess and that which S, even under idealiza-
tion, is assigned. The upshot is that ever more complex and subtle
indefeasibility clauses are produced to be met with ever more complex
and subtle Gettier-like counterexamples. Finally our intuitions flicker,
then go out.

   My suggestion, then, is that indefeasibility theorists, instead of
explicitly including an endorsing (iii$_g$) clause in their analyses,
attempt to capture the force of such a clause by idealizing S's justifi-
catory setting in a way that will win our endorsement. I have not
presented an argument showing that such an extensional match
between the (iii$_g$) clause and a defeasibility clause *cannot* be
achieved, though it seems clear that it cannot be achieved in any nat-
ural, intuitively clear manner. More deeply, if my diagnosis of the
driving force behind the indefeasibility theorists' program is cor-
rect—that they are trying to make an externalist clause do the work
of the (iii$_g$) clause—then it is hard to see what motivation remains for
pursuing it. Perhaps it is also clear why pursuing such a program
involves committing something akin to the naturalistic fallacy.

Notes

   1. This slide toward a functional-equivalent skepticism is a standard,
though often unacknowledged, feature of theories of epistemic justifica-
tion—both of the foundationalist and of the antifoundationalist kind. Such
theories are examined in detail in the second part of this study.
   2. Some have challenged the assumption that knowledge must involve
belief. There are, I think, examples that incline us in this direction. My own
reaction to these examples—which typically involve idiot savants and the
like—is that they present instances where we are inclined to extend our
concept of knowledge to cover a new and curious case. It is rather like the

decision to treat imaginary numbers as numbers even though they are not indicators of quantity.

3. Swain, 1974. Reprinted in Pappas and Swain, 1978. Page references are to the latter source.

4. Note that this parenthetical remark contains something like the (iii$_g$) clause.

5. The example appeared first, as far as I know, in Lehrer and Paxson, 1969. The version presented here is from the excellent survey introduction in Pappas and Swain, 1978.

6. In "Epistemic Defeasibility," Swain tells a parallel story involving a demented cardinal, and he draws a parallel conclusion from it. (See Pappas and Swain, 1978, 165–66.)

7. Lehrer and Paxson, 1969. Reprinted in Pappas and Swain, 1978. Page references are to the latter source.

8. Swain's version of the clause is more complex:

(ivf) There is an evidence-restricted alternative Fs* to S's epistemic Fs, such that "S is justified in believing that h" is epistemically derivable from the other members of the evidence component in Fs*. (177)

But I think I have caught its gist.

9. Swain's example goes back to Gettier's first case, which was discussed in the previous chapter. Actually, Swain's story is not Gettier's. In the Gettier original, Smith justifiably believes that someone else, Jones, will get the promotion, then Smith gets the promotion himself. In the Swain version, Jones justifiably believes that he himself will get the promotion, then someone else, Smith, gets it. Notice that this is more than a simple name reversal. However disorienting, these changes make no systematic difference. Anyway, Swain finds a remarkably ingenious way of modifying the Gettier original that produces a counterexample to (ivf). It goes as follows:

Suppose everything is as before, except that for some reason Jones happens to know that Smith (the man who is really going to get the job) has ten coins in his pocket. If so, then in the relevant evidence-restricted alternative to Jones's epistemic framework it will be evident for Jones that Smith will get the job (since it is *in fact* evident for him that Smith will not get the job) and *also* evident for him that Smith has ten coins in his pocket. But this renders it evident for him (in the alternative) that the man who will get the job has ten coins in his pocket. According to (ivf), Jones's justification for h is indefeasible, and he would have knowledge. But neither of these things is so. (180)

10. Actually, in his case, a mad cardinal.

3

Externalism

Responses to Gettier problems usually begin with a natural and almost always *correct* diagnosis of a flaw in a particular, relatively simple Gettier case. Defeasibility analyses are prompted by the reflection that a standard, quite reasonable, justification can sometimes be overridden or defeated by the discovery of an additional, usually out-of-the-way, fact. The strategy of the defeasibility theorists is to continue to hold the traditional view that to be knowledge something must be justified true belief, but then they go beyond this to include an external standard intended to situate *S* in a manner, perhaps quite unknown to *S*, that insulates him from Gettier counterexamples.

In this chapter I will consider various responses that depart more radically from the traditional definition by replacing the epistemic-responsibility clause with an externalist clause rather than simply supplementing it with an externalist clause. Following the classification given at the beginning of chapter 2, I will call such approaches externalist or purely externalist analyses, since they add an externalist, but no internalist, clause to the claim that knowledge is true belief.

Concerning externalist positions, I shall attempt to establish the following general theses:

I. Externalist theories that are *purely* externalist are refuted by counterexamples due to violations of the (iii_p) clause of the analysis offered in chapter 1.

This, I think, is a matter placed beyond doubt by arguments presented by Laurence BonJour. Next, dividing externalist theories into two categories—those that take the notion of causation as central

and those that appeal to the subjunctive—I shall argue for two fur-
ther theses:

II. Pure causal theories of knowledge fail because they do not
capture the force of the (iii$_g$) clause of the analysis I have pro-
posed.

III. Subjunctive theories gain their plausibility precisely because
the subjunctive constraints they introduce are guided by a tacit
acceptance of the (iii$_g$) clause.

The first two theses are defended in this chapter; the third thesis will
be defended in chapter 4.

The Attraction of Externalism

One way to motivate an externalist approach to knowledge is to note
that in Gettier's first case Smith came correctly to believe that
someone with ten coins in his pocket would get the promotion by a
fluke or by a happy *accident*. Smith, using a perfectly standard pro-
cedure that is usually reliable, applies it in a context in which it is
not reliable, and arrives at a belief that turns out to be correct any-
way. All Gettier problems have this property, but they appear in
their most striking form in category II—epistemic-luck—examples.
If we are struck by the accidental correctness of Smith's belief, it is
tempting to say that Smith's deficiency does not lie in the way his
belief stands to his *reasons* for holding it, but, instead, with the way
his belief stands to the *fact* believed. This diagnosis looks outward,
rather than inward. The most natural suggestion concerning the
nature of this relationship is to say that it is *causal*. To count as
knowledge, a belief must be causally related—in the right sort of
way—to the fact it concerns.[1]

A causal theory of knowledge faces a number of tasks. It must
specify what kinds of knowledge come under its purview. It must
tell us what sorts of causal relations are the *right* sort. Later in this
chapter I will examine in detail how Alvin Goldman, in his early
writings on this subject, attempted to deal with these topics. I will
show how, faced with counterexamples, his causal theory is driven
in the direction of a subjunctivist theory. But before going into
detail, I wish to examine the criticisms, briefly noted in chapter 1,
that BonJour has brought against reliabilist theories of knowledge,
and, indeed, against all purely external theories of knowledge. I think
there is something unquestionably right about these criticisms. I
also think BonJour overextends them and, as a result, obscures some

of the important insights of externalism. That externalism does contain important insights is, I think, worth insisting upon.

BonJour against the Reliabilists

In a variety of places BonJour has developed a series of counterexamples intended to show that a person's belief can stand in the kind of relationship that externalists take to be knowledge-conferring, and yet still not know it to be true.[2] To this end, he presents a series of vignettes with characters sharing the following features:

(I) They each, in fact, possess reliable clairvoyant powers.

(II) On the basis of these powers, they each come to believe that the president is in New York, which, in fact, he is.

Beyond this, BonJour attributes certain epistemic failings or shortcomings to each character, with the intention of showing that each, in his own particular way, is "highly irrational and irresponsible in accepting [the] belief when judged in the light of *his own* subjective conception of the situation" (1985, 3, emphasis added). Because of this irrationality and irresponsibility, BonJour denies that any of these characters *knows* that the president is in New York.

The string of examples unfolds as follows:

Samantha:
(a) believes she is clairvoyant, but without good grounds, and (b) further believes the president is in New York, contrary to good evidence she possesses.

Casper:
(a) believes he is clairvoyant without possessing good evidence for this belief and, indeed, possesses strong evidence to the contrary, and (b) further believes the president is in New York, but without good grounds.

Maud:
is like Casper but has good evidence that *no one* is clairvoyant.

BonJour thinks examples of this kind show that

> external or objective reliability is not enough to offset subjective irrationality. If the acceptance of a belief is seriously unreasonable or unwarranted from the believer's own standpoint, then the mere fact that unbeknownst to him its existence in those circumstances lawfully guarantees its truth will not suffice to render the belief epistemically justified and thereby an instance of knowledge. (41)[3]

With some reservations to be noted later, I think this is correct. My inclusion of the (iii_p) clause in the criteria for knowledge given in chapter 1 reflects this agreement.

BonJour next imagines the externalist patching up his position in a way that excludes just these counterexamples:

> In addition to the lawlike connection between belief and truth and the absence of reasons against the particular belief in question, it must also be the case that the believer in question has no cogent reasons, either relative to his own situation or in general, for thinking that such a lawlike connection does *not* exist, that is, that beliefs of that kind are not reliable. (40)

A modification of this kind already extorts a great deal from an externalist in the direction of a mixed internalist/externalist position, but BonJour wants more. Notice that in BonJour's first three examples, there are serious cognitive *conflicts* among the person's various beliefs. These conflicts form the basis of the charge that he or she is "highly irrational and irresponsible in accepting a belief." In BonJour's next example, his fourth, no such conflict exists, only unwarrantedness. Since it is crucial for our discussion, I'll quote this example in full:

> Norman, under certain conditions which usually obtain, is a completely reliable clairvoyant with respect to certain kinds of subject matter. He possesses no evidence or reasons of any kind for or against the general possibility of such a cognitive power or for or against the thesis that he possesses it. One day Norman comes to believe that the President is in New York City, though he has no evidence either for or against this belief. In fact the belief is true and results from his clairvoyant power under circumstances under which it is completely reliable. (41)

Does Norman know? Before trying to answer that question—and certainly before trying to draw any philosophical conclusions from our answer—it is worth noting that Norman seems a bit unhinged. As BonJour describes him, Norman has no idea that he is clairvoyant, so, from his point of view, he merely finds himself believing something that has just popped into his head. BonJour exploits just this point in order to charge Norman with irrationality.

> From his standpoint, there is apparently no way in which he *could* know the President's whereabouts. Why then does he continue to maintain the belief that the President is in New York City? Why isn't the mere fact that there is no way, as far as he knows, for him to have obtained this information a sufficient reason for classifying this belief as an unfounded hunch and ceasing to accept it? And if

Norman does not do this, isn't he thereby being epistemically irrational and irresponsible? (42)

On the basis of this example, BonJour draws a very strong conclusion
against reliabilists, and, by implication, against externalists in general:
"Part of one's epistemic duty is to reflect critically upon one's beliefs,
and such critical reflection precludes believing things to which one
has, to one's knowledge, no reliable means of epistemic access" (42).

I don't find this persuasive. I think BonJour's Norman example
gains its apparent strength from two sources: first, our well-founded
distrust of claims to clairvoyance, which is here irrelevant, and second, the curiously disconnected way that BonJour characters receive
their clairvoyant deliverances out of the blue, as it were. I think we
might have a different attitude toward clairvoyance if, in addition to
supposing that it exists, we also supposed that it enters smoothly
into the warp and woof of everyday life. To see this, consider the following example. Suppose that *S* exercises the following clairvoyant
power, though without realizing it. When reading a book, he now
and again clairvoyantly "sees" the opening line on the next page
before he turns to it. When he does this, he simply skips over to the
second line on the page, though he doesn't notice that he does this
either. (We can imagine a physiologist making this remarkable discovery while studying eye movements in reading.) Under these conditions, would we be willing to say that *S* knows that page 150 of
BonJour's *Structure of Empirical Knowledge* begins with the words
"can fulfill the implicit mandate to seek out possibly conflicting
observations" when, in fact, he "sees" these words clairvoyantly
before he turns the page and skips over them after he does so? Given
the story as told, I am inclined to say that he does know this. Furthermore, I do not change my mind in the slightest if I am further
told that *S* firmly believes, on very good grounds, that clairvoyance
does not exist.[4] A fortiori, BonJour's stronger claim that *S* does not
know this unless he has positive grounds warranting his reliance on
his clairvoyant powers strikes me as deeply counterintuitive.

This fictitious example can be supported by a real one. The orientation of our body, our posture, the disposition of our limbs, and
the like, are given to us by a faculty generally known as properception. There are particular receptors (in muscles, joints, and other
places) associated with properception. Most people know nothing
about properception, and, if they accept the folk theory of there
being only five senses, they actually hold views incompatible with
its existence. Certainly most people have done nothing like "reflect
critically" on the beliefs that come from this source. Most people

haven't the slightest idea how they can tell that their legs are slightly parted without looking at them or feeling any contact with surrounding objects. This, however, does not make a belief on this matter either irrational or irresponsible. If that is right, which I think it is, then BonJour's Norman example does not establish what it is intended to establish, namely, that we have a *positive* duty to reflect critically on the sources of our belief, and have behaved irresponsibly and irrationally if we do not do so.

Externalist Grounds

Here it might help to ask how these considerations bear on the two justification clauses that appear in the doctrine that knowledge is justified true belief justifiably arrived at. In the first place, these reflections suggest that causal theories in particular and, I think, externalist theories in general must invoke something like the (iii_p) clause in order to exclude cases of conscious cognitive conflict. For the reasons I have just presented, I do not think the BonJour examples show that the (iii_p) clause needs to be made stronger than this.

Turning to the (iii_g) clause, it captures the idea that in saying S knows something we endorse S's grounds[5] as establishing the truth of what S holds on those grounds. Previously, I have tended to use the words "grounds" and "reasons" interchangeably, and I have often spoken of evidence. This way of speaking suggests that S's justification must take the form of an *argument*, and when we endorse S's grounds, we are indicating that his argument, or something close to it, establishes the truth of the thing he claims to know. In these most recent examples, however, we are not dealing with arguments, but with *powers*. As a naive clairvoyant, S believes that page 150 of BonJour's book begins with certain words; it is true that it does; and he has come to this belief using a reliable power. Does his belief satisfy the (iii_p) clause in our definition? Does it settle the (iii_g) clause? I think the answer to both questions is yes.

In defense of these claims, we can first notice how a shift in perspective, from S's standpoint to ours, affects the matter. Unlike S, who is a naive clairvoyant, *we* can produce an *argument* showing that S was correct in believing that page 150 begins with certain words:

S is a reliable clairvoyant.

This was an occasion on which he exercised this power.

And so on.

It might be tempting to say that, since we can produce a justificatory argument and S cannot, we know but he doesn't. This, however, strikes me as simply an intellectualist prejudice that the example of unrecognized perceptual mechanisms of properception was supposed to cure.

Once this intellectualist prejudice is set aside, it becomes evident that justificatory performances come in various forms. Given the kind of creatures we are, there are certain claims that can be justified only by arguments. As a case in point, consider the assertion "Certain forms of schizophrenia are associated with location five on the human chromosome." To the question "How do you know?" we might receive a citation to some authority, but ultimately an assertion of this kind could only be known to be true on the basis of an elaborate piece of argumentation. Compare this with the following example of a different kind of justified empirical belief. S is told to close her eyes and relax. She is then asked to describe the position of her body, and she replies, "I am sitting down, slumped slightly forward, with both my arms and legs crossed." I'm not sure what someone would say if asked "How do you know this?" except, perhaps, "I paid attention to my body." Still, these remarks can be made either responsibly or irresponsibly. It would, for example, be irresponsible for S confidently to report the position of her limbs if, *as she knew*, she was being treated with a mind-altering drug. In this case, she could responsibly report where her limbs seemed to be, but not where they were. So the (iii_p) clause is applicable even to simple and immediate reports of perception. Furthermore, in the standard case, the sincere report of the subject that her limbs are in a certain position is sufficient to establish the truth of this claim both for her and for us. Additional facts can undercut the force of such a sincere avowal. We might know, as S doesn't, that she has been given a mind-altering drug. In this circumstance it is not epistemically irresponsible for her to think her legs are crossed, though we, given our additional information, realize that her sincere avowal does not establish the truth of this claim. So the (iii_g) clause applies here as well. Furthermore, since it is possible for the (iii_p) and (iii_g) clauses to take different truth values, it is easy to construct a Gettier example even for this simple level of perceptual judgment. We need only suppose that S has been given a drug that will make her feel (and hence think) that her legs are crossed no matter what position they are in. The experimenter asks her the position of her legs when, as it turns out, they happen to be crossed, and she reports

that they are crossed. Does she, in this case, know that her legs are crossed? Of course not. Though, as the story is told, she is epistemically responsible in believing her legs to be crossed, the feeling and willingness to believe that they are crossed do not establish that they are.

It seems, then, that justificatory performances come in a variety of forms ranging at the extremes from those that involve complex ratiocination to those that rely on the unreflective use of a perceptual power or capacity. Both can be carried out in a responsible or an irresponsible manner; both can establish or fail to establish the truth of some belief. Both are sources of knowledge. Until relatively recently, philosophers have often tended to think of knowledge as solely the product of intellectual activity. The externalists have made an important contribution to epistemology by breaking the spell of this intellectualist prejudice. As we shall see, externalism, in its various forms, has serious shortcomings, but its rejection of intellectualist prejudices is not among them.

To backtrack, my earlier remarks were intended to create a presumption in favor of the first thesis announced at the beginning of this chapter:

I. Externalist theories that are *purely* externalist are refuted by counterexamples due to violations of the (iii_p) clause.

I think BonJour-style counterexamples establish this point. In the end, however, this is not a strong criticism, since externalist positions are easily modified to meet it without compromising their central nonintellectualist (non-Cartesian) core. Nothing more is needed than the provision that S not believe that he *lacks* the knowledge-giving power in question. And perhaps even that provision is not needed if the beliefs formed on the basis of this power fit in with and support beliefs derived from other sources.

In conclusion, BonJour's criticisms will cut deeply against an externalist position only if he is right in saying that: "Part of one's epistemic duty is to reflect critically upon one's beliefs, and such critical reflection precludes believing things to which one has, to one's knowledge, no reliable means of epistemic access" (42). BonJour's fourth example—the one concerning Norman—was intended to establish this thesis. I believe I have shown that it does not, and, as a result, a central insight of externalism—that knowledge can arise through the unreflective use of a perceptual capacity—has been preserved.

Goldman and Causal Theories of Knowing

At the head of this chapter I announced three theses. The second read:

II. Pure causal theories of knowledge fail because they do not capture the force of the (iii_g) clause of the analysis I have proposed.

By a *pure* causal theory I mean one that restricts itself to *actual* causal relations and does not traffic in *subjunctive* alternatives. A theory of this kind was developed by Alvin Goldman in his early essay "A Causal Theory of Knowing."[6] I will spend some time on this essay because its central idea carries immediate intuitive force; it contains important insights that should be preserved; and, beyond this, seeing (as Goldman himself came to see) what's wrong with it will move us along in the direction of a subjunctivist account of knowledge.

Goldman targets Gettier's second example for criticism. In this example, Smith justifiably believes the falsehood that Jones owns a Ford and from this draws an inference to the disjunctive proposition that Jones owns a Ford or Brown is in Barcelona. This disjunctive proposition turns out to be true in virtue of Brown's being in Barcelona, something that Smith has no reason to believe. Concerning this example, Goldman notes, as others had before him, that

> what *makes* p [i.e., "Jones owns a Ford or Brown is in Barcelona."] true is the fact that Brown is in Barcelona, but . . . this fact has nothing to do with Smith's believing p. That is, there is no causal connection between the fact that Brown is in Barcelona and Smith's believing p. (68)

The first sentence in this passage suggests a principle that strikes me as obviously sound:

P1. If the fact that makes p true has nothing to do with S's believing it, then S's believing that p does not count as his knowing that p.

This principle is not particularly constraining, and not constraining at all if we think that any two things will always have *something* (however remote) to do with one another.

Next, moving very quickly, Goldman cashes this principle in with respect to a *causal* relation. The reason, according to Goldman, that Smith does not know that p is because his belief that p is not

causally related to the fact that makes it true, that is, Brown's being in Barcelona. But, as Goldman seems to realize, some kind of causal relation can be found between any two facts (or events, if you prefer),[7] so for Goldman's causal rendering of P1 to be plausible, it must be stated something like this:

> P2. If the fact that makes *p* true is not causally related in the *right sort of way* to *S*'s believing it, then *S*'s believing that *p* does not count as his knowing that *p*.[8]

This principle will be probative only if we are given an adequate specification of sorts of causal relations that count as the *right* sorts. The bulk of Goldman's paper is an attempt to provide—in outline at least—just this specification. But I do not want to go into this just yet.

The point I do want to make is that Goldman's analysis of this *particular* case along causal lines strikes me as persuasive. To see this, we can consider a simpler case, not involving a Gettier problem, where Smith, who lives in Detroit, believes that Brown is in Barcelona. We will further assume that Smith is correct in this belief. How might he come by it? Perhaps it was communicated to him, say, by a phone call from Brown. Perhaps he infers this because at some earlier time Brown told Smith that he (Brown) was going to Barcelona.[9] We can, of course, think of other ways that Smith might come to have this belief, but in all these cases we are dealing with causal connections of a standard kind. Furthermore, given the kind of belief in question, one concerning a particular remote event or fact, and given the kind of creatures we are, for example, not preprogrammed Leibnizian monads, it is hard to see how Smith's belief that Brown is in Barcelona could be warranted in the absence of certain standard causal relations connecting this remote fact with Smith's belief in it. I think, then, that the following principle is true:

> P3. For a certain class of beliefs (yet to be specified), if the fact that makes a belief in that class true is not causally related in the *right sort of way* to *S*'s believing it, then *S*'s believing it does not count as his knowing it.

This may sound more like a promissory note than a principle, but it is, I think, on the right track.

Perhaps in the flush of youth, Goldman then went on to make a claim that I find more problematic, even with respect to his particular example: "Alternatively, if Jones did own a Ford, and his owning the Ford was manifested by his offer of a ride to Smith, and this in turn resulted in Smith's believing *p*, then we would say that Smith

knew that *p"* (68). Here the underlying principle seems to be something like this:

> P4. For a certain class of beliefs (yet to be specified), if the fact that makes a belief in this class true is causally related in the *right sort of way* to *S*'s believing it, then *S*'s believing it counts as knowledge.[10]

What sorts of beliefs might satisfy P4? It seems to apply most naturally to simple judgments of perception and memory.[11] Since they are more compact than his remarks about perception, I'll cite what Goldman has to say about memory. First he lays down a necessary condition for someone's remembering that *p:*

> *S* remembers *p* at time *t* only if *S*'s believing *p* at an earlier time is a cause of his believing *p* at time *t*. (70)

I'm not sure this is correct. Perhaps in the past I had a veridical experience that did not engender a belief because I had what I thought to be good reasons for thinking that the experience was not veridical. Later this experience comes back to me, but with the earlier scruple forgotten, and then, for the first time, I believe. This seems to be a case of remembering something not previously believed. The point, however, even if correct, is not damaging to Goldman's general program.

This supposed necessary condition for remembering cannot serve as a sufficient condition, for, as Goldman notes, "not every causal connection between an earlier belief and a later one is a case of remembering" (70). It does not take exotic examples to show this. Perhaps years ago I recorded one of my beliefs in a diary, then completely forgot it. I read this entry now and believe what it says. Here an earlier belief that *p* stands in a causal relationship to a present belief that *p*, but not in a relationship that counts as remembering. Exactly what sort of causal relation, then, is sufficient for remembering? Goldman candidly admits that he cannot answer this question in detail:

> As in the case of perception, however, I shall not try to describe this process in detail. This is a job mainly for the scientist. Instead, the kind of causal process in question is to be identified simply by example, by "pointing" to paradigm cases of remembering. Whenever causal processes are of that kind—whatever that kind is, precisely—they are cases of remembering. (70–71)[12]

Paradigm-case arguments are now out of style—I suppose because they were overused—but here, at least, their application strikes me

as sound. Simple perceptual beliefs and simple beliefs derived from
memory exemplify the sort of causal relationship that Goldman has
in mind even if it defies adequate representation. In any case, causal
theories of knowing like Goldman's have their most natural appli-
cation, and gain their initial intuitive plausibility, with respect to
these simple cases of perception and memory.

A great deal of our knowledge is not of this (relatively) immedi-
ate kind. We think we know many things that we do not now per-
ceive or remember. Adopting traditional terminology, Goldman
marks this as the difference between inferential and noninferential
knowing.[13] Having told a causal story about noninferential (empiri-
cal) knowledge, Goldman attempts to extend the tale by giving a
causal account of inferential knowledge as well. He introduces his
leading ideas using the following example: "S perceives that there is
solidified lava in various parts of the countryside. On the basis of
this belief, plus various 'background' beliefs about the production of
lava, S concludes that a nearby mountain erupted many centuries
ago" (72). Using solid lines to indicate causal relations (e.g., a vol-
cano spewing lava across the countryside) and dotted lines to indi-
cate inferential relations (e.g., applying established geological princi-
ples to the observed facts), Goldman diagrams the epistemic
situation as shown:

$$\{p\} \longrightarrow \{q\} \longrightarrow B_s(q) \underset{\diagdown}{\diagdown}$$
$$B_s(p)$$
$$B_s(r)$$
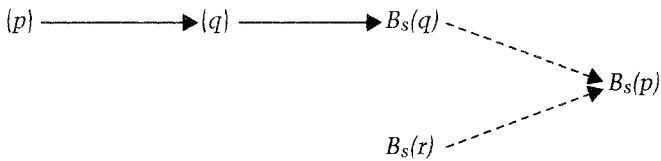
Here it seems natural to speak of a causal/inferential account of
empirical knowledge, but Goldman wants to call the entire chain a
causal chain. He achieves this by stipulation: "if a chain of inference
is 'added' to a causal chain, the entire chain is causal" (73). Given
this stipulation, Goldman can then say: "Thus, in figure 1, p is a
cause of S's belief of p, whether or not we regard S's belief of q a
cause of his belief of p" (73).

At the very least, this way of speaking seems highly artificial, for
I do not think anyone would say that the mountain's erupting was a
cause (or one of the causes) of S's believing that it erupted. The cause
of S's coming to hold this belief was, instead, a rather complex piece
of causal reasoning. Yet Goldman attempts to preserve his initial
causal picture even when the inferential component becomes the

dominant factor. This comes out in his treatment of the following example. Suppose the mountain does spew lava over the landscape, but some time later someone removes it. Later still, another person covers the landscape with lava so that it looks as if the mountain had spewed lava over the landscape. In this case, according to Goldman, S does not know that a volcano had erupted centuries ago, because "the fact that the mountain did erupt is not a cause of S's *believing that it erupted*" (72, emphasis added). Surely this is a peculiar way to describe the situation, for it would seem much more natural to forget about S's doxastic states and say that S doesn't know that the mountain erupted centuries ago because his reasoning concerning the source of the lava has turned out to be no good. Everything is handled quite naturally by an application of what I have called the (iii_g) clause.

In fact, a few pages later, Goldman's analysis begins to move in the direction of introducing a (iii_g) clause. Discussing another example whose details need not concern us, he remarks: "As in the lava case, S knows p because he has *correctly reconstructed* the causal chain leading from p to the evidence for p that S perceives" (74, emphasis added). Notice that here a reference to a causal account of S's *belief* has simply dropped out. S is said to know because he has produced an accurate reconstruction of the causal sequence leading from the eruption of the mountain centuries ago to the present dispersal of lava across the landscape. At this point the emphasis switches to the legitimacy of the causal reconstruction:

> If [S] is to know p, his reconstruction must contain no mistakes. Though he need not reconstruct every detail of the causal chain, he must reconstruct all the important links. An additional requirement of knowledge based on inference is that the knower's inferences be warranted. That is, the propositions on which he bases his belief of p must genuinely confirm p very highly, whether deductively or inductively. Reconstructing a causal chain merely by lucky guesses does not yield knowledge. (75)

It now seems that Goldman's position has undergone a fundamental change. Where he first told us

(i) S knows that p only if S stands in the right sort of causal relationship to the fact that p,

he now seems to be telling us that, for certain cases at least,

(ii) S knows that p only if S can support his belief that p with the right sort of causal reconstruction.

That these are very different positions is clear because the first con-
dition would not disbar animals from possessing knowledge, whereas
the second condition presumably would.

Let me stop here and draw a number of tentative conclusions
concerning a causal criterion of knowledge:

I. Setting aside the minor problem noted above, the causal crite-
rion of knowledge does seem to lay down a *necessary* condition
for certain kinds of knowledge, and in virtue of this provides a
way of solving Gettier's original problems.

II. But even in those cases where the causal criterion finds its
most natural application (i.e., with memory and so-called imme-
diate perception), BonJour examples show that it does not pro-
vide a sufficient condition for knowledge.

III. Goldman's causal criterion for knowing becomes more and
more attenuated as the knowledge in question becomes com-
plexly inferential. In fact, for complex inferential knowledge,
Goldman abandons his purely externalist standpoint in favor of
a mixed externalist/internalist position. Under pressure, the
theory is slowly turned outside in.

Yet for all these difficulties, there still might be something
importantly right about a causal account of knowing, at least in the
region of its paradigmatic application: noninferential judgments of
perception and memory. This, as Goldman came to see, is wrong, for
even in the area of its paradigmatic application, a causal theory will
not provide us with a sufficient condition for knowing.

The Move to the Subjunctive

At the beginning of "Discrimination and Perceptual Knowledge,"[14]
Goldman's second major effort to give an account of perceptual
knowledge, he tells us: "Like an earlier theory I proposed, the envis-
aged theory would seek to explicate the concept of knowledge by ref-
erence to the causal processes that produce (or sustain) belief. Unlike
the earlier theory, however, it would abandon the requirement that
a knower's belief that p be causally connected with the fact, or state
of affairs, that p" (120). Perhaps I am wrong, but I do not think this
passage accurately states the relationship between Goldman's origi-
nal account, as he presented it in "A Causal Theory of Knowing,"
and the new, reliable causal process theory he develops in this essay.
In particular, in the very *broad* sense in which Goldman uses the
notion of a causal connection (or causal chain) in the original article,

I do not think he "abandon[s] the requirement that a knower's belief that *p* be causally connected with the fact, or state of affairs, that *p*." I think he still, tacitly at least, accepts this as a necessary condition for perceptual knowledge and then goes on to strengthen it in response to problems he had not noticed before.

Along the way, Goldman offers a number of examples intended to show that the earlier, simpler, causal theory will not do, but the Category II case of facsimile barns serves as his central example. In broad outline, Goldman proposes to solve this problem along the following lines: "What, then, is my proposed treatment of the barn example? A person knows that *p*, I suggest, only if the actual state of affairs in which *p* is true is *distinguishable* or *discriminable* by him from a relevant state of affairs in which *p* is false. . . . So, once we are apprised of the facsimiles in the district, we are inclined to deny that Henry knows" (124). Put crudely, according to Goldman, we are not inclined to say that *S* knows that there is a barn before him, for he would have formed precisely the same belief, though now falsely, had he been looking at one of the facsimile barns.

For those who find this barn example problematic, let me cite another example that strikes me as being completely decisive. It comes from Fred Dretske and concerns Goldman's lava example:

> Not far from *M* is another mountain *N*. The geology of the area is such that at the point in time at which *M* erupted something, so to speak, was bound to give. If *M* had not erupted, *N* would have. Furthermore, the location of *N* is such that if it, rather than *M*, had erupted, the present distribution of lava would have been, in all respects relevant to *S*'s taking it as a reason for believing that *M* erupted, the same. (In Pappas and Swain, 46)

Here, even though the geologist's belief that *M* erupted stands in what Goldman earlier deemed the right sort of causal relation to the eruption of *M*, the geologist cannot be said to know that *M* erupted because she would have believed the very same thing if it had been *N* that erupted instead.

Reflections of this kind led Goldman to a new primitive insight intended to replace (or at least supplement) his earlier primitive insight that *S* does not know something when there is no causal connection between the fact he claims to know and his knowing it. Setting aside subtleties, the contrast between the two positions comes to this:

> Old Insight: If the fact that *p* did not cause *S* to believe that *p*, then in believing that *p*, *S* does not know that *p*.

New Insight: If *S* would have believed that *p*, even if *p* were false, then in believing that *p*, *S* did not know that *p*.

The old insight provided a way of avoiding what I have called Category I Gettier counterexamples, for example, the Gettier originals. It did not provide a way of eliminating Category II counterexamples, for example, the fake barns. The new insight is specifically intended to overcome this problem.

The key difference here is that the new criterion is explicitly cast in the subjunctive mode, a fact given prominence by Goldman when he remarks: "The theory of knowledge I envisage, then, would contain an important counterfactual element" (120).

As far as I can see, what I have labeled Goldman's New Insight, if correct, would hold quite generally for inferential as well as noninferential knowledge—indeed, for nonempirical as well as empirical knowledge,[15] but Goldman restricts his discussion to noninferential perceptual knowledge. Informally, he states his basic position in these words: "What our analysis says is that *S* has [noninferential] perceptual knowledge if and only if not only does his perceptual mechanism produce true belief, but there are no relevant counterfactual situations in which the same belief would be produced via an equivalent percept and in which the belief will be false" (139). As Goldman sees it, two things here demand careful explanation: the notion of an *equivalent* percept and the notion of a *relevant* counterfactual situation. But even though Goldman himself spends a great deal of time explaining the notion of perceptual equivalence, I propose to concentrate on what he says about relevant counterfactual situations.

Going back to the barn facsimile case, Goldman remarks: "The presence of the facsimiles does not 'create' the possibility that the object Harry sees is a facsimile. Even if there were no facsimiles in the district, it would be possible that the object on that site is a facsimile. What the presence of the facsimiles does is make the possibility *relevant;* or it makes us *consider* it relevant" (124). In common parlance, we might say that the presence of facsimiles makes it a *real* possibility that the object Harry sees is a facsimile.

We want to know, then, what makes a counterfactual situation a relevant or real possibility, and in trying to answer this question, we encounter a characteristic philosophical problem: Specifications at the extremes are unsatisfactory, and there seems to be no obvious way of producing a satisfactory specification between these extremes. Goldman notes one side of this problem: "If knowledge required the elimination of all *logically possible* alternatives, there

would be no knowledge (at least of contingent truths)" (124–25). On the other side, if the range of alternative subjunctive possibilities is narrowly restricted, say to those cases where the alternative situation is exactly alike in every respect to the original situation, then Goldman's new requirement becomes empty, placing no further counterfactual restrictions on what will count as perceptual knowledge than his original causal account did.

How does Goldman deal with this problem in specifying relevant alternatives? The short answer is he doesn't. After reflecting at large on recent literature in the field, he says, "I shall not try to settle the question of whether the semantic content of 'know' contains rules that map the putative knower's situation into a unique set of relevant alternatives. I leave open the question of whether there is a 'correct' set of relevant alternatives, and if so, what it is" (128).

Why doesn't this admission, so candidly made, simply take the stuffing out of Goldman's enterprise? Why isn't he more worried about this? Here's what Goldman says:

> In defending my analyses of "perceptually knows," however, I shall have to discuss particular examples. In treating these examples I shall assume some (psychological) regularities concerning the selection of alternatives. Among these regularities is the fact that speakers do not *ordinarily* think of "radical" alternatives, and take them seriously, if the putative knower's circumstances do not call attention to them. (128)

By "radical" alternatives, Goldman has in mind such things as the Cartesian hypothesis of a deceiving demon, or Russellian speculation that the world came into existence five minutes ago with memories, geological strata, and so forth, all in place.

In fact, however, in discussing particular examples, Goldman's assumptions go beyond bracketing "radical" alternatives: he also brackets remote (and not so remote) alternatives. Normally it does not cross our minds to wonder whether a barn before us might be a facsimile barn, yet this possibility is not radical in the sense in which the Cartesian demon and the Russellian creation present radical alternatives. First, the possibility of a facsimile barn is straightforwardly realizable. We could actually build such barns, perhaps with the aid of a grant from the National Endowment for the Arts. More importantly, the facsimile-barns hypothesis is not globally dislocating in the way in which the Cartesian and Russellian hypotheses are. There are straightforward ways of deciding whether a barn is a facsimile barn: you look around the back or try to get inside. The distinctive feature of the Cartesian and Russellian hypotheses—

whether this counts as a strength or a weakness—is that all such standard procedures for checking are themselves called into question.

On Goldman's account, S cannot be said to know that p (e.g., that there is a barn in front of him) because he *would* have believed this *had* a relevant counterfactual situation obtained where p is false (i.e., if there were a facsimile barn before him). Goldman's puzzle is why we can count this as a reason for saying that S in a land containing facsimile barns does not know, whereas in normal circumstances the procedure he used (i.e., looking out his car window) would give him knowledge. Goldman's answer is that S overlooked a relevant counterfactual alternative (i.e., that there is a Potemkin barn in front of him).

But, to ask the most obvious question, why is the facsimile barn possibility relevant? The answer is that *we* know that there are such barns in the vicinity since we have been told this in the setting-up of the example. Furthermore, we realize that their existence undercuts S's normally reasonable grounds by raising the level of scrutiny needed to form warranted perceptual beliefs. Thus we explain why a certain counterfactual alternative is relevant by pointing out how certain facts bear on the warrant a person's grounds provide for a certain proposition. If this is right—and I shall argue for this in detail in the next chapter—then in saying that pure causal theories can be refuted by an appeal to counterfactual considerations, we are implicitly acknowledging that they fail through violations of the (iii$_g$)—or epistemic-appraisal—clause for the analysis of knowledge. This was the second thesis stated at the beginning of this chapter:

II. Pure causal theories of knowledge fail because they do not capture the force of the (iii$_g$) clause of the analysis I have proposed.

However, as the argument has developed, this thesis was made to depend on another:

III. Subjunctive theories gain their plausibility precisely because the subjunctive constraints they introduce are guided by a tacit acceptance of the (iii$_g$) clause and are thus approximations of it.

I will defend this thesis in the next chapter.

Notes

1. D. M. Armstrong captures the initial intuitive plausibility of this approach using the *causal* metaphor of a reliable instrument, in particular,

a reliable thermometer. A reliable thermometer indicates that the temperature is, say, 98.6° because that is what the temperature is. The temperature's being 98.6° causes the thermometer to indicate this. (See Armstrong, 1961.)

2. Actually, BonJour is more narrowly concerned with accounts of empirical *justification*, but, as he himself notes at various points, his counterexamples apply equally against externalist analyses of *knowledge*.

3. BonJour, 1985.

4. Nor would it make any difference if S's semiclairvoyant reading of BonJour's arguments totally convinced him that Norman didn't know.

5. More carefully we should add the qualification "or something very close to S's grounds" because of cases where we accept S's grounds as establishing some conclusion, but do so only after vindicating his argument in the face of some objection.

6. Goldman, 1967. Reprinted in Pappas and Swain, 1978. Page references are to the latter source.

7. In extremis, we can trace (or claim to trace) a causal connection back to the Big Bang, or we can construct conceptually bizarre causal relations in the style of Nelson Goodman.

8. It should be noted that Goldman, from the start, restricts his discussion to empirical knowledge, remarking "I think that the traditional analysis is adequate for knowledge of nonempirical truths" (67).

9. The two cases exemplify the two patterns of causal relations that Goldman explores. Roughly, in the first pattern, S's believing that p has the fact that p as a causal ancestor; in the second pattern, S's believing that p and the fact that p share a common causal ancestor.

10. P4, unlike the previous principles, lays down a *sufficient* condition for knowledge and therefore is subject to BonJour counterexamples. Perhaps S, along with his sound belief that Jones owns a Ford, further believes (justifiably or not) other things incompatible with it. In cases of this kind, BonJour would argue that Smith does not know and, as indicated before, I think he is right in this. As a nod in the direction of this problem, we can add a dummy clause to P4:

> P5. For a certain class of beliefs (yet to be specified), if the fact that makes a belief in this class true is causally related in the right sort of way to S's believing it, then S's believing it does count as his knowing it, provided that S does not believe other things incompatible with his believing it.

(With this said, BonJour problems are quarantined until further notice.)

11. Beyond this, P5 also applies to beliefs formed from information communicated to us.

12. Here Goldman cites H. P. Grice's important article "The Causal Theory of Perception" (Grice, 1961).

13. Many have argued that even the simplest acts of perception are inferential. I'll return to this topic in Part II of this study.

14. Goldman, 1976. Page references are to Pappas and Swain, 1978.

15. In this latter case I am thinking, for example, of mathematical knowledge. Suppose that S uses a proof procedure that determines theorem-hood on the basis of the number of words in the theorem. Let's suppose that, using this method, the Pythagorean theorem is declared true. Here we will not say that S has shown that the Pythagorean theorem is true, for he would have declared it true even if it happened to be false.

Some think there is a puzzle here since the Pythagorean theorem, being a necessary truth, cannot possibly be false. I've never understood the force of this puzzle since I have no difficulty in imagining that a sentence (or a statement used in making a sentence) might express something false even though, in fact, it expresses something necessarily true. Looked at this way, we are not confronted with the task of trying to conceive of something necessarily true being possibly false. But I digress.

4

Subjunctivism and Subjunctivitis

Conclusive Reasons

In the previous chapter I examined the way in which Goldman's pure causal theory evolved into a causal/subjunctive theory. I further stated, without argument, that the subjunctive component gains its plausibility through being guided by what I have called the (iii$_g$) clause. The burden of this chapter is to provide an argument for this claim. But instead of following the development of Goldman's thought further, I will begin by examining a statement of the subjunctivist position in a pure and relatively simple form as it is found in Fred Dretske's remarkably fine article, "Conclusive Reasons."[1]

Dretske states the broad thesis of this essay in these words:

> In those cases where knowledge that P rests on evidence, grounds, or reasons, and when the question "How does S know?" can sensibly be asked and answered, the evidence, grounds, or reasons must be conclusive. Anything short of conclusive reasons, though it may provide one with justified true beliefs, fails to give the kind of support requisite to knowledge. I shall also urge that the possession of conclusive reasons to believe, properly qualified, is also a sufficient condition for knowledge. (1–2)

Dretske then gives two characterizations of a conclusive reason. At the head of his essay he uses explicit modal terminology:

> Letting the symbol "◊" represent the appropriate modality (a yet-to-be-clarified sense of *possibility*) I shall say, then, that R is a conclusive reason for P, if and only if, given R, ~◊~ P (or alternatively ~◊(R . ~ P)). (1)

More important for our purposes, he also uses a subjunctive con-
struction to characterize conclusive reasons:

> R [is] a conclusive reason for P if and only if R would not be the case
> unless P were the case. (12)

This said, we can summarize the fundamental features of Dret-
ske's position in terms of the interrelationships among the three fol-
lowing formulas:

(1) S knows that P and he knows this on the basis (simply) of R.

(a2) R is a conclusive reason for P.

(2) R would not be the case unless P were the case. (1)[2]

Although Dretske does not explicitly present his position in this
way, with respect to these formulas, he is clearly committed to the
following three theses:

> I. One cannot consistently affirm an instantiation of (1) and
> deny the corresponding instantiation of (a2). [That is, if S knows
> that P and he knows this on the basis (simply) of R, then R is a
> conclusive reason for P.]

> II. One cannot consistently affirm an instantiation of (1) and
> deny the corresponding instantiation of (2). [That is, if S knows
> that P and he knows this on the basis (simply) of R, then R
> would not be the case unless P were the case.]

For Dretske, thesis I and thesis II are related by a strong thesis
already mentioned:

> III. R [is] a conclusive reason for P if and only if R would not be
> the case unless P were the case.

By spreading things out this way, I can summarize my evaluation of
Dretske's position as follows: (i) Thesis I is a variant of my (iii$_g$)
clause and is true. (ii) Thesis II follows from thesis I, and so is also
true. (iii) The appeal to the subjunctive for the analysis or elucidation
of conclusive reasons fails because thesis III is clearly false.

Dretske attempts to establish the intuitive plausibility of thesis
II using a series of examples. The first and leading example has the
following form. Suppose the following instantiation of (1) is true:

> (1a) S knows that his child's temperature is normal and he
> knows this on the basis of the (normal) reading on the ther-
> mometer (which he has placed in the child's mouth, etc.).

He then asks whether one can "consistently affirm (1a) and deny the corresponding instantiation of (2)?"

(2a) The thermometer would not have read 98.6° unless the child's temperature was normal. (2)

Dretske expects a negative answer to this question.

To see the force of this much of the argument, we can consider a mistaken objection to it. Someone might complain that the thermometer might have read 98.6° even if the child's temperature was not normal—perhaps it was a defective thermometer, one that always reads 98.6°. Dretske, I think, says some questionable things concerning this example, but this much strikes me as clear. If in an actual case we have good reason for suspecting the reliability of the thermometer and for that reason refuse to accept (2a), then *for that very same reason* we will refuse to accept (1a). That's the point of thesis II, which strikes me as being perfectly correct.

Before examining the difficulties hinted at in the previous paragraph, we can examine the straightforward way in which this approach handles Gettier problems. In order to get this application, the notion of conclusive reasons must be relativized to a person's *having* them, or perhaps better, *commanding* them. Dretske attempts this using the following formulation of the necessary and sufficient conditions for *S* having conclusive reasons:

> Of course, *R* may *be* a conclusive reason for believing *P* without anyone believing *P*, much less having *R* as their [sic] reason for believing. I shall say, therefore, that *S has a conclusive reason, R,* for believing *P* if and only if:
>
>   (A) *R* is a conclusive reason for *P*,
>   (B) *S* believes, without doubt, reservation, or question, that *P* is the case and he believes this on the basis of *R*.
>   (C) (i) *S* knows that *R* is the case or
>   (ii) *R* is some experimental state of *S* (about which it may not make sense to suppose that *S knows* that *R* is the case; at least it no longer makes much sense to ask *how* he knows). (12–13)[3]

Given this specification of *S*'s having a conclusive reason, *R*, for *P*, a solution to the category I Gettier problems falls out at once. It was characteristic of those examples that *S* could be justified in believing that *P* on the basis of *R*, though *P* was *false*. This is not possible on Dretske's analysis, since *S*'s possession of conclusive reasons for *P* precludes the possibility that *P* is false. Though he tucks

it away in a footnote, Dretske explicitly relates his analysis to the original Gettier problems:

> It is this stronger connection which blocks the sort of counterexample which can be generated to justified-true-belief analyses of knowledge. Gettier's (and Lehrer's) examples, for instance, are directed at those analyses which construe knowledge in terms of a *degree* of justification which is compatible with being justified in believing something false. . . . On the present analysis, of course, the required connection between *S*'s evidence and *P* is strong enough to *preclude P*'s being false. One cannot have *conclusive* reasons for believing something which is false. (13n)

I think the final sentence in this passage is correct. In fact, though I drew no conclusions regarding the Gettier problems, I made a similar claim a number of years before Dretske's article appeared. In *Evidence and Meaning*,[4] I offered the following schema as a first approximation for the analysis of "knowing that":

> *X* possesses adequate grounds for *p*. (60)[5]

Contrasting this with the position developed by Chisholm, I remarked: "When Chisholm uses the phrase 'adequate evidence' he does not intend it in the strong sense that a proposition [must be] true if there is adequate evidence for it—there can be adequate evidence for a false proposition" (62). Over against this, I held: "If someone possesses adequate grounds available for some assertion, then there *are* adequate grounds available for that assertion, and from that it follows that the assertion is true" (62). Thus I used my notion of adequate grounds in the same strong sense that Dretske used his notion of conclusive reasons. For this reason both positions are immune to Gettier's original counterexamples, though unlike Dretske, I did not have the wit to note this—or footnote this.

The area of agreement between Dretske (then) and me (then) consists in accepting what I earlier called thesis I:

> I. If *S* knows that *P* and he knows this on the basis (simply) of *R*, then *R* is a conclusive reason for *P*.

But Dretske thinks it is important to defend the subjunctive counterpart of thesis I. I have labeled it thesis II:

> II. If *S* knows that *P* and he knows this on the basis (simply) of *R*, then *R* would not be the case unless *P* were the case.

I think Dretske prefers this subjunctive version of his position because it seems to give an account or explanation of what it is for

something to be a conclusive reason. I think this is wrong. But before pressing this criticism, we can examine some of the strengths of the conclusive-reason approach, for they will be worth preserving even if the subjunctive analysis of conclusive reasons must be rejected.

We can note first that Dretske's analysis allows us to attribute a conclusive reason, *R*, to *S* without committing ourselves to saying that *S* knows or even believes that *R* is a conclusive reason. As a result, we can attribute knowledge to animals (as we often do) without implausibly attributing to them sophisticated knowledge or beliefs *about* conclusive reasons. A similar point can be made concerning *human* perceptual knowledge.

Beyond this, Dretske's analysis blocks a *certain* form of skepticism. Skeptics are often pictured as reasoning in the following way:

It is possible that I am now dreaming.

If I am dreaming, I cannot know that there is a table before me on the basis of my present experience, since, when dreaming, I might have the same experience without there being a table before me.

But (let's say) I do not have conclusive reasons for believing that I am not dreaming.

Therefore, I do not know that there is a table before me.

To his credit, Dretske does not adopt any of the facile responses common at the time he wrote "Conclusive Reasons." He does not place the blame on the skeptics, accusing them of "setting artificially high standards for knowledge" (19). He does not appeal to some version of verificationism in an effort to declare the skeptic's doubt meaningless. He does not employ blunt paradigm-case arguments. Instead, he simply points out that it is possible, on his analysis, to know that *P* (there is a table before me) on the basis of *R* (my present visual experience) without knowing *A* (that *R* is a conclusive reason for *P*). As Dretske puts it: "One qualifies for knowledge when one has conclusive reasons for believing; one need not, in addition, know that one has conclusive grounds" (17). So if the skeptic claims that certain radical possibilities show that we can not know, say, that there is a table in front of us, Dretske can reply, even granting that these radical possibilities cannot be excluded, that *for all we know, we do know such things.*

I think this is correct, and brilliant. But it is important to note that this argument has no force against most forms of traditional philosophical skepticism. Sextus Empiricus, for example, is willing

to admit that for all we know we may know a great deal, say, about the external world. He thinks that this admission distinguishes Pyrrhonian skepticism from Academic skepticism, which he rejects as a form of negative dogmatism. The Pyrrhonian is content to point out that certain radical possibilities show that for all we know we do *not* know anything. Having apprised others of these radical possibilities and the consequences that flow from them, he then challenges others to justify their continued use of knowledge claims. The response "Well, for all we know we do know" does not meet this challenge. In fact, Dretske never claims that his analysis of knowledge has force against traditional skeptical arguments, but still, it is important to see that it does not. Here, however, I am treading on matters that will be discussed in more detail later in this chapter.[6]

Subjunctivism

Earlier I asked why Dretske preferred to express his position in terms of thesis II,

II. If S knows that P and he knows this on the basis (simply) of R, then R would not be the case unless P were the case,

rather than in terms of thesis I:

I. If S knows that P and he knows this on the basis (simply) of R, then R is a conclusive reason for P.

What is the point of this shift into the subjunctive? The subjunctive formulation is not needed for the derivation of the antiskeptical consequences of Dretske's analysis, for they follow directly from the fact that S can have R as a conclusive reason for P without knowing this or even believing it. As far as I can see, the appeals to the subjunctive do no work that could not have been done directly by the concept of a conclusive reason. If this is right, subjunctive conditionals will do honest work in this context only if they help elucidate the notion of a conclusive reason.

My own view is that subjunctive conditionals do not help us understand conclusive reasons (and, derivatively, knowledge) for, *if anything*, the illumination experience occurs in the other direction: our notion of conclusive reasons is needed to place constraints on truth conditions for subjunctive conditionals. First I will try to make this point with respect to Dretske's informal discussion in "Conclusive Reasons"; later I will make it again with respect to possible-worlds accounts of subjunctive conditionals.

As already noted, we avoid Gettier's original counterexamples (and some of their near kin) if we drop the standard weak justification clause in the definition of knowledge, and replace it with the following:

(iii$_{cr}$) S has conclusive reasons, R, for P.

or, alternatively, for Dretske,

(iii$_{sub}$) S has a reason, R, for P such that R would not be true unless P were true.

It seems that an analysis of knowing generated by using either of these clauses will be subject to BonJour counterexamples.[7] Moreover, as Dretske himself realized, an analysis relying on the (iii$_{sub}$) clause seems to be subject to what I have called category II counterexamples, namely those that involve epistemic luck.[8] As a variation on Dretske's original example, suppose that S took the thermometer from a box of thermometers that were being discarded because the manufacturer had written saying they were defective. In fact, all the thermometers were defective except for the one chosen by S; it was in perfect working order. In this case it is true that the thermometer would not have registered 98.6° unless the child's temperature were normal, yet we seem unwilling to say that S, on the basis of reading the thermometer, had a conclusive reason for believing that P. Since S lacks a conclusive reason for his belief, we will also deny that he knows that the child's temperature was normal.[9] This example seems to drive a wedge between Dretske's subjunctive clause and his conclusive-reason clause, which, on his approach, should come to the same thing. More to the point, the conclusive-reason clause survives this counterexample, whereas its subjunctive counterpart does not.

One response to this difficulty—and in a way it is perfectly correct—is to say that the counterexample depends on an inappropriate description of what S did in coming to the conclusion that his child's temperature is normal. What he did, as the story tells us, is to reach into a box containing thermometers, all but one of which were faulty. Then it is not true that a thermometer selected in this way would not have read 98.6° unless the child's temperature were normal. Thus if we get the description right, the proposed counterexample seemingly fails. Dretske attempts to capture the force of these reflections in the following statement: "When one has conclusive reasons, then this is sufficient for knowing that P is the case when those reasons are *properly specific*, both with regard to *what it is*

that displays the particular features on which one relies *and* the particular features themselves" (22).

But this modification to the *conclusive-reason* clause is not responsive to the problem at hand, for no objection has been found to the conclusive-reason analysis, and therefore it stands in need of no modification. It seems that in the thermometer example just given, *S* did *not* have conclusive reasons for believing that his child's temperature was normal even though it was true that the thermometer would not have read 98.6° unless the child's temperature was normal. If this is right, we do not need a qualification on the conclusive-reason clause; what is needed instead is a qualification on the subjunctive clause to get it back into alignment with the conclusive-reason clause. Targeting the subjunctive clause rather than the conclusive-reason clause would yield the following result:

If *S* has a reason, *R*, for *P* such that *R* would not be true unless *P* were true, then this is sufficient for knowing that *P* is the case when those reasons are *properly specific*, both with regard to *what it is* that displays the particular features on which one relies *and* the particular features themselves.

The problem is to find some well-motivated basis for determining when reasons are properly specific, and this amounts to finding a well-motivated basis for selecting the descriptions to be used in the formulation of the appropriate subjunctive conditional. (The specific thermometer used by *S* would not have read 98.6° unless the child's temperature were normal, but a thermometer taken from that particular box might have read 98.6° even though the child's temperature was not normal.) I do not think Dretske's remarks about subjunctive conditionals, as they appear in "Conclusive Reasons," provide an adequate way of determining which description is relevant or appropriate. Furthermore, as I shall argue later, I do not think so-called possible-worlds semantics will provide any help in this matter either. In general, it does not seem possible to explicate the notion of a relevant description in terms of the truth of subjunctive conditionals, since their truth seems always to depend on an antecedent specification of what considerations are *allowed to be* relevant.

As the story is told, it does seem clear that *S* does not know that his child's temperature is normal. Where does this intuition come from? To return to a theme stressed in previous chapters, our clear decisions on these questions are derived from our privileged position relative to the story told. We know (by fiat) that all the thermome-

ters in the box, save one, are faulty. We see at once that this is a relevant consideration because we see, using commonsense considerations, how this bears on the legitimacy (or warrant) of measurements made using thermometers so chosen. Furthermore, we do not have to loop through a subjunctive conditional in order to gain this insight, although now that we see that S's grounds for believing that his child's temperature was normal were not warranted, the context has been sufficiently fixed to allow us to deny the subjunctive conditional:

> A thermometer selected in the way that S selected it need not have read 98.6° unless the child's temperature were normal.

But to suggest that S's lack of conclusive reasons (and hence his lack of knowledge) is somehow explicated in terms of the falsehood of this subjunctive conditional gets things backward—or at least assigns a priority where none has been shown to exist.

A deeper problem with Dretske's subjunctive analysis was pointed out to me by Dretske himself. The counterexample comes from Raymond Martin.

> Suppose that S buys a ticket at a local horse track. The ticket is such that if S picks the winner in the first race or the winner in the second race, or both, S wins; otherwise, S loses. S picks Gumshoe to win the first race and Tagalong to win the second. Before the first race begins S is called away from the track and does not return until the end of the second race. While S is away Gumshoe wins the first race . . . and Tagalong finishes last in the second. . . . On returning to the track and still not knowing the results of either race S goes to the cashier's window and presents his ticket. Without speaking a word, the cashier gives twenty dollars in exchange for S's ticket. (215–16)[10]

In this case, given that Tagalong had lost, S would not have received the twenty dollars at the time he collected it unless Gumshoe had won, but still, at the time he received it, his receiving twenty dollars was not a conclusive reason for his holding that Gumshoe had won. This argument strikes me as completely decisive.[11] Since an appeal to proper specificity does not seem to help here, the counterexample seems fatal to the subjunctivist position Dretske developed in "Conclusive Reasons."

It is easy enough to see what happens in this example. Given that we know that Tagalong lost the second race, we can see that S would not have won anything unless Gumshoe had won the first. From our perspective, we can also see that winning twenty dollars

on his bet did not provide *S* with a conclusive reason for believing that Gumshoe won, even though it is true that if Gumshoe had not won, then *S* would not have won twenty dollars. One of the leading ideas of an externalist account of knowledge is that a person may have a conclusive reason for believing something without having the background information in virtue of which it can be seen to be a conclusive reason. This may be a reasonable position with respect to some forms of knowledge, perhaps knowledge based on immediate perception or memory, but the position loses its plausibility when it is extended to include knowledge based on reasoning. Since in Martin's example *S* was absent from the track and was not described as being a clairvoyant, he could only have known that Gumshoe had won through a piece of reasoning based on the winnings he received. Since this information did not provide adequate grounds for believing that Gumshoe had won, and, beyond this, since *S* reasoned quite badly given the information he did possess, *S* did not know. From this we may conclude that thesis III,

> *R* is a conclusive reason for *P* if and only if *R* would not be the case unless *P* were the case,

though it holds from left to right, does not hold from right to left. The claim that something is a conclusive reason will guarantee the truth of a counterpart subjunctive conditional, but not conversely.

Subjunctivitis

Under this tendentious title I will examine the notion that so-called possible-worlds semantics can play a crucial role in the explication of knowledge claims. In particular, I will examine accounts of knowledge that begin, as Dretske's did, by introducing subjunctive conditionals into the analysis of epistemic statements and next use possible-worlds semantics to establish truth conditions for these subjunctive conditionals. This, I think, is a subterfuge. To show this, I'll concentrate on some things Robert Nozick has said.[12]

Nozick's account of knowledge is initiated by a specific diagnosis of Gettier problems. It is then driven and shaped by a specific philosophic goal: the refutation of what he takes to be a brand of skepticism. He begins in a standard way by announcing:

> Our task is to formulate further conditions to go alongside
> (1) *p* is true
> (2) *S* believes that *p*. (172)

For him, these conditions should be individually necessary for knowledge, and jointly sufficient.

The first move is to add a clause that will deal with Gettier problems, and to this end, Nozick recognizes, as others had before him, that in Gettier's second example, the following is true of *S:* If no one in the office owns a Ford, *S* would still believe someone did.

Because this is true of *S*, it strikes Nozick as intuitively obvious that *S* does not know that someone in his office owns a Ford (173). So it seems that a correct account of knowledge ought to contain a clause that denies (or implies the denial of) the occurrence of a situation of this kind. Nozick first suggests that the following clause meets this demand:

> (3) If *p* weren't true, *S* wouldn't believe that *p*. (172)

Nozick next notes that the conjunction of this additional clause with the original two clauses does not "rule out every problem case." In particular,

> There remains . . . the case of the person in the tank who is brought to believe, by direct electrical and chemical stimulation of his brain, that he is in the tank and being brought to believe things in this way; he does not know this is true. However, the subjunctive condition is satisfied: if he weren't floating in the tank, he wouldn't believe he was. (175)

Nozick diagnoses this example in the following way:

> The person in the tank does not know he is there, because his belief is not sensitive to the truth. Although it is caused by the fact that is its content, it is not sensitive to that fact. The operators could have produced any belief, including the false belief that he wasn't in the tank; if they had, he would have believed that. Perfect sensitivity would involve beliefs and facts varying together. . . .
> [Thus] the subjunctive condition
>   (3) not-*p* ⇒ not-(*S* believes that *p*)
> tells us only half the story about how his belief is sensitive to the truth-value of *p*. It tells us how his belief state is sensitive to *p*'s falsity, but not how it is sensitive to *p*'s truth; . . .
>   This additional sensitivity is given to us by a further subjunctive: if *p* were true, then *S* would believe it.[13]
>   (4) *p* ⇒ *S* believes that *p*. (175–76)

Roughly, Nozick's fourth clause is the subjunctive counterpart of the demand found in causal theories of knowledge that the causal connection between the belief and the fact believed be of the right kind.

Nozick concludes this exposition by stipulating that he will "say of a person who believes that *p,* which is true, that when 3 and 4 hold, his belief *tracks* the truth" (178).

The notion of tracking the truth, so specified, is Nozick's primitive insight concerning knowledge: Knowledge is true belief that tracks the truth. But, as Nozick notes at once, the conditions laid down in this simple theory are too strong. To see why, it is sufficient to cite just one of Nozick's examples. "A grandmother sees her grandson is well when he comes to visit; but if he were sick or dead [and did not come to visit], others would tell her he was well to spare her upset. Yet this does not mean she doesn't know he is well (or at least ambulatory) when she sees him" (179). Quite clearly, this example shows a defect in Nozick's third clause, for the grandmother's method of coming to know that the child is well was, let's say, perfectly reliable, and it seems quite beside the point that, if she had not had the opportunity to employ this reliable method, she would have fallen back on another method (listening to the reports of solicitous friends) that would, in fact, be unreliable. So Nozick concludes that "we must restate our conditions to take into account the ways and methods of arriving at beliefs" (179).

Relativizing his position to methods in the right sort of way is an essential aspect of Nozick's position, for, as we have just seen in the example of the sharp-eyed, but gullible, grandmother, without such a relativization, the analysis can easily be shown to fail. In fact, when we get into the details of this discussion, and examine competing methods and available but unused methods, the issues become extremely complicated. But for our present purposes, we do not have to examine these refinements—or "epicycles," as Nozick calls them—since our main concern will be with Nozick's appeal to subjunctive conditionals and possible-worlds semantics, and the central points can be made without reaching the more refined, method-relativized version of his analysis.

Subjunctive Conditionals and Possible Worlds

On Nozick's account, the truth of a knowledge claim depends on the truth of two subjunctive conditionals. How, then, do we determine the truth of a subjunctive conditional? Nozick says little on this, and the little he does say is altogether damaging to his cause.

Stated simply, the difficulty with subjunctive conditionals— especially contrary-to-fact conditionals—is that there seems to be no

nonarbitrary way of deciding whether they are true or false, or, per-
haps, indeterminate in their truth values. Examples illustrating this
have been in the literature for a long time. A vivid one from Quine
will serve our purposes: "If Bizet and Verdi had been compatriots,
would Bizet have been Italian, Verdi French, or what?"[14] It seems
plain, on its face, that this question cannot be answered without an
arbitrary stipulation concerning which of these composers' birth-
places is to remain fixed. That just means that this subjunctive con-
ditional has no nonarbitrary truth value. Furthermore, even after this
first stipulation has been made, others must follow in its train. Sup-
pose that we hold fast to the claim that Verdi was Italian and let
Bizet's nationality float. It then seems that the subjunctive condi-
tional "If Bizet and Verdi were compatriots, then Bizet would have
been Italian" perhaps comes out true. But how are we to deal with
Bizet having been born in Paris? Again we have a choice: we can give
up that claim or we can hang on to it. Suppose we hang on to it and
also continue to suppose that Verdi was Italian. We then seem to get
the following result: "If Bizet and Verdi had been compatriots, Paris
would have been in Italy." And so forth.

Turning now to Nozick's remarks on determining the truth
value of subjunctive conditionals, what does he have to say? The
answer is not much. He notes first that his subjunctive clauses place
constraints on knowledge claims that are "not easy to satisfy, yet
not so powerful as to rule out everything as an instance of knowl-
edge." In particular: "A subjunctive conditional 'if p were true, q
would be true' does not say that p entails q or that it is logically
impossible that p yet not-q. It says that in the situation that would
obtain if p were true, q also would be true" (173). Speaking impres-
sionistically, subjunctive conditionals seem to be just strong enough
to do the job that epistemologists need done.

However, not any subjunctive conditional will do, for, as noted
above, some subjunctive conditionals seem indeterminate in their
truth values. How, then, are the appropriate subjunctive condition-
als to be specified? If Nozick has anything to say on this crucial sub-
ject, it appears in the following passage, which is a continuation of
the passage just cited:

> This point is brought out especially clearly in recent "possible
> worlds" accounts of subjunctives; the subjunctive is true when
> (roughly) in all those worlds in which p holds true that are closest
> to the actual world, q also is true. (Examine those worlds in which
> p holds true closest to the actual world, and see if q holds true in all

of these.) Whether or not q is true in p worlds that are still farther away from the actual world is irrelevant to the truth of the subjunctive. (173–74)

Thus, in order to characterize the truth conditions for subjunctive conditionals, further structures must be imposed on the set of possible worlds. To do this, Nozick, following Robert Stalnaker and David Lewis,[15] speaks of possible worlds that are *close* to the actual world. Intuitively, in evaluating a subjunctive conditional, we consider only those possible worlds where p is true (p-worlds) that *resemble* the actual world most closely. This brings us to Nozick's recommendation:

> Examine those worlds in which p holds true closest to the actual world, and see if q holds true in all of these. (173)

If this condition is met, then p subjunctively implies q.

The first, and perfectly correct, response to this recommendation is to say that it is idle, since nothing counts as *examining* possible worlds to see what is true in them. Talk of examining possible worlds (closest or otherwise) is a metaphor that is deeply misleading. In order for possible-worlds semantics to provide truth conditions for subjunctive conditionals, an independent specification must be given concerning what resemblances count (or can be made to count). Such a specification (even if it can be carried out in a nonarbitrary way) is not a matter for formal semantics.

Nozick actually seems to be making—or at least hinting at—this same point in a footnote appended to his comments on possible-worlds semantics:

> If the possible-worlds semantics is used to represent counterfactuals and subjunctives, the relevant worlds are not those that are closest or most similar to the actual world, unless the measure of closeness or similarity is: what would obtain if p were true. *Clearly, this cannot be used to explain when subjunctives hold true, but it can be used to represent them.* (174n, emphasis added)

This is correct, but it is an extraordinary admission on Nozick's part. In the body of the text, he seems to suggest that in some way or other the *examination* of close possible worlds will make sense out of our calling certain subjunctive conditionals true rather than false. It now seems that it is our instincts concerning the truth of subjunctives that tutor us concerning what resemblances or similarities count and, thus, which possible worlds possess the requisite measure of closeness. Things have been turned around, as Nozick

implicitly acknowledges in the following technical flourish: "It would be of interest to know what properties hold of distance metrics which serve to represent subjunctives, and to know how subjunctives must be structured and interrelated so that they can be given a possible worlds representation" (174n).

At least this much is clear: The appeal to possible worlds does no substantive work in establishing the truth conditions for subjunctive conditionals—a point that Nozick, in his more candid moments, acknowledges. Derivatively, then, it does no substantive work in establishing the truth conditions for knowledge claims. Why, then, has Nozick appealed to possible-worlds semantics at all? At one point he suggests that his use is purely heuristic: "I am not committed to this type of account. I sometimes use it, though, when it illustrates points in an especially clear way" (174). Then in a footnote appended to this passage he tells us: "Our purposes require, for the most part, no more than an intuitive understanding of subjunctives" (680n).[16] In this same footnote, he also makes it clear that he is not a realist with regard to possible worlds. Why, then, the appeal to possible worlds at all? Why, more importantly, does the qualifying phrase "for the most part" appear in the disclaimer just cited? The answer is that possible-worlds semantics—and not just an intuitive understanding of subjunctives—will play a central role in his refutation of a position he calls skepticism. I'll turn to this next.

Nozick against the "Skeptics"

Since it is hard to improve on it, let me begin with Peter Klein's concise statement of Nozick's characterization of a skeptic's argument:

> Suppose that some skeptical scenario (call it "SK") were true, for example, that I am in the vat on Alpha Centauri and that I am having the same experiences that I am presently having. Since being in a room in New Brunswick entails not being on Alpha Centauri, if I know that I am in New Brunswick, then I know that I am not on Alpha Centauri. But I don't know that I am not on Alpha Centauri. Thus I don't know that I am in New Brunswick. (Luper-Foy, 1987, 267)

For our present purposes, it will be useful to unpack this argument as an explicit reductio ad absurdum:

1. *S* knows that he is in New Brunswick. (hypothesis)
2. If *S* is in New Brunswick, then he is not on Alpha Centauri.

3. If *S* knows that he is in New Brunswick, and being in New Brunswick entails his not being on Alpha Centauri, then *S* knows that he is not on Alpha Centauri.
4. But *S* does not know that he is not on Alpha Centauri.
5. So (contrary to assumption 1) *S* does not know that he is in New Brunswick.

The third premise in this argument is an instance of the principle of closure for knowledge, namely, that if *S* knows that *p*, and *p* entails *q*, then *S* knows that *q*. Obviously that principle, so stated, is false, since, for example, a person who knows all of the axioms (definitions, postulates, etc.) of Euclidean geometry does not thereby know all its theorems. For our present purposes, however, we can avoid this difficulty by assuming that in the cases we examine *S* actually draws the required inference. Even with this qualification the principle of closure for knowledge is false, since *S* may draw the required inference on irrational grounds, for example, in the belief that any two propositions entail each other. Here we will suppose that *S* doesn't draw his inferences in such an irrational way. In general, we will ignore finicky objections to the principle of closure for knowledge so that the rejection of this principle, if we reject it, is important.

Returning to the argument itself, it is certainly valid. But, noting this gives us absolutely no reason to suppose that the conclusion is true. A valid argument equally invites us to deny the conclusion as the basis for denying the conjunction of the premises or to accept the premises as a basis for affirming the conclusion. *Modus tollens* and *modus ponens* are on a par. Given this, various responses to the argument can simply be laid out.

(I) The Skeptic's Move. The skeptic, as Nozick pictures him, accepts premises (2) through (4) and then, by an argument reductio ad absurdum, denies premise (1). To be persuasive, the skeptic must present good reasons for accepting all three premises (2) through (4), but historically he has usually concentrated on establishing the truth of the fourth premise.

Those who reject this argument have a variety of options:

(IIa) Moore's Response. With breathtaking simplicity, G. E. Moore has argued that one can know the sort of thing that *S* is said to know in premise (1), and, indeed, know this with certainty.[17] Then, taking premises (2) and (3) for granted, he *directly* concludes that premise (4) is false.

(IIb) The Justificationalist Response. Like Moore, justification-alists—in the broad sense in which I will use the term—reject premise (4). Unlike Moore, they attempt to produce a systematic argument *showing* that (4) is false. In the second part of this study, I will argue that all such justificationalist attempts fail.

(III) The Geographical Response. It's at least possible to deny the second premise, holding that New Brunswick is (or may be) on Alpha Centauri. (Alternatively, a crazed Whiteheadean might suggest that the inference involves the fallacy of simple loca-tion.) I will ignore this response.

(IV) Nozick's Response. Since on his account of knowledge, *S* could know that he is in New Brunswick, he must reject one of the premises that leads to the conclusion that he could not. He picks on premise (3), the closure principle. It is important to note, however, that in order to reach the conclusion that the clo-sure principle is false, he must join the skeptic in maintaining that the skeptical premise (4) is true.

All these responses except III are worth taking seriously and they will be considered in their proper place. Nozick's response is our pre-sent concern. His goal is to reach the following result: Although (i) *S* cannot know that he (or at least his brain) is not floating in a jug on Alpha Centauri, (ii) he can know that he is in New Brunswick. With respect to (i), Nozick invokes the standard *indiscriminability* argu-ment: Since *S* could have the same beliefs and experiences whether he is jugged or unjugged, he, from his point of view, cannot know which situation he is in. Here we might challenge Nozick's inter-nalist assumptions concerning the content of a belief, but I'll let that pass.[18] Our present concern is his defense of the claim that *S* can know that he is in New Brunswick even though he cannot know that he is not on Alpha Centauri.

In order to defend this position, Nozick attacks the third premise in the skeptic's argument: the principle of closure for knowledge. Again, we are not going to concern ourselves with minor problems concerning this principle, nor will we pause to correct them. To his credit, Nozick is not raising finicky objections to this principle. He is willing to suppose that someone can know that *p* is true, know that *p* entails *q*, on the basis of this knowledge infer *q* from *p*, but still, for all this honest labor, *not* know that *q*. The supposition to the contrary Nozick calls the skeptic's "short step" from a skeptical scenario to skepticism concerning ordinary beliefs. Let me quote the key part of Nozick's argument intended to establish this point:

Let us begin with condition

(3) If p were false, S wouldn't believe that p.

When S knows that p, his belief that p is contingent on the truth of p, contingent in the way that subjunctive condition 3 describes. Now it might be that p entails q (and S knows this), that S's belief that p is subjunctively contingent on the truth of p, that S believes that q, yet his belief that q is not subjunctively dependent on the truth of q, in that it (or he) does not satisfy:

(3') If q were false, S wouldn't believe that q.

For 3' talks of what S would believe if q were false, and this may be a very different situation than the one that would hold if p were false, even though p entails q. That you were born in a certain city entails that you were born on earth. Yet contemplating what (actually) would be the situation if you were not born in that city is very different from contemplating what situation would hold if you weren't born on earth. (206–7)

The argument is capped with an appeal to possible worlds:

There is no reason to assume the (closest) not-p world and the (closest) not-q world are doxastically identical for you, and no reason to assume, even though p entails q, that your beliefs in one of these worlds would be a (proper) subset of your beliefs in the other. (206)

There are various ways of responding to this argument. One is to take Nozick's argument on his own terms and ask whether the possible-worlds representations where the principle of closure for known implication fails are intuitively plausible on other counts. When the principle of closure for known implication fails, what else fails with it, and are these failures tolerable? Though the material has not reached print at the time of this writing, I understand that Saul Kripke has argued, in a number of venues, that Nozick's representation of possible worlds where the principle of closure for known implication fails carries with it an avalanche of further inferential failures that are deeply counterintuitive. If that is correct, then the rejection of the principle of closure for known implication is bought at an unacceptably high price.[19] More generally, I have argued that it doesn't make sense to invoke possible-worlds semantics to specify truth conditions for substantive subjunctive conditionals. Although it is possible to impose structures on possible worlds such that the principle of closure for knowledge fails—and in this sense *represents* failures of this principle—it does not make sense to speak of examining close possible worlds and discovering that this principle actually does fail.

The Failure of Epistemic Closure

Over against both these claims it might be argued that Nozick, and Dretske before him, provide clear examples where the principle of closure for known implication fails. If that is correct, then the question of how such a representation should be implemented becomes a matter of technical detail. I will argue that the supposed examples of failure of closure under known implication, properly understood, exhibit no such failure.

Instead of examining skeptical scenarios, I will pitch the discussion in terms of remote possibilities.[20] As we shall see, the same patterns will emerge that Nozick provides in his own examples. Suppose that I am walking through a remote wilderness and spot what I take to be a great horned owl. The light is good; great horned owls are easy to identify; and I am an expert birder. I have no hesitation in thinking that I know that the bird is a great horned owl. I then reflect as follows: If that object is a great horned owl, then it is not a dummy made of wax. I know that it is a great horned owl, but do I thereby know that it is not a dummy made of wax? Well, from where I stand, there is no way of telling whether it is a wax dummy or not. It seems, then, that I do not know that it is not a wax dummy, and using *modus tollens* on the closure principle, I seem driven to the conclusion that I was wrong, previously, to think I knew that the bird was a great horned owl. This example parallels Nozick's brain-in-the-vat skeptical scenario. If we are inclined to hold on to the claim that I did, for all that, know that it was a great horned owl in the tree, then I seem driven to abandon the principle of closure under known implication. I'll call this case I.

To continue the story, and get case II, suppose that I check and discover that the object is a wax dummy. I now reason (using *modus ponens* on the closure principle): I know that this is a wax dummy; I also know that if it's a wax dummy then it is not a great horned owl, so I know that it is not a great horned owl. In this reasoning I have again used the principle of closure for knowledge, and I have used it, it seems, in a perfectly legitimate way. Why do we tend to have different reactions to these two cases? Why, that is, are we inclined to reject the principle of closure in the first case, yet accept it without hesitation in the second?

The answer, I think, is that in our common use of knowledge claims—including *serious* uses of knowledge claims—our assertions are not responsive to every possibility that might refute them. We are willing to make knowledge claims while recognizing that sometimes

odd things come up that lead us to withdraw them. This is precisely
what happened in the first owl example given above. To use an anal-
ogy drawn from computer chess programs, our epistemic activities
are governed by parameters that control width and depth of search.
These parameters are determined by a number of factors: the com-
plexity of the situation, the supposed risk of error, the importance of
getting things right, the transaction costs of gaining new information,
and so on. The important fact for the present discussion is that these
parameters of search can change as circumstances change; in particu-
lar, the exigencies of inquiry can spread them wider or drive them
deeper.

A third variation on the owl example will illustrate these points.
As before, I am walking in a remote wilderness and spot what I take
to be a great horned owl. The identification strikes me as wholly
unproblematic. I then come upon a cleverly fabricated pileated
woodpecker tacked to the trunk of a pinon pine. My suspicions
aroused, I search the area and find the landscape cluttered with wax
imitations of birds. What now should I say about my identification
of the great horned owl? In case II, I withdrew my claim that I knew
that it was a great horned owl because I came to know something
(i.e., that it was a wax dummy) that implied that it was not an owl
at all. In this third case, I apply *modus tollens* to the principle of clo-
sure, and come to the conclusion that I do not know that the object
is a great horned owl because I do not know that it is not a wax
dummy.[21]

Thus the three cases sort out in the following way:

In case I we have a seemingly problematic application of the
principle of closure in drawing the conclusion that I do not
know that it is a great horned owl because I do not know that it
is not a wax dummy, since nothing in the context prompts me
to worry about the possibility of wax dummies.

In case II we have an unproblematic application of the principle
of closure because I have discovered something (i.e., that the
object is a wax dummy) that implies that it is not a great horned
owl and thus cannot be known to be one.

In case III, we have an unproblematic application of the princi-
ple of closure because I have discovered something (i.e., a batch
of wax dummies) that does prompt me to increase the level of
search in a way that will bring this principle into play.

Here, then, is my general diagnosis of the apparent failures of the
principle of closure for knowledge to which Nozick has pointed. In

each of his narratives, Nozick has misrepresented a fact about the level of inquiry as a fact about the logic of epistemic terms. In each of Nozick's cases, the depth of search must shift as we pass from considering the antecedent p to considering the consequent q. The level of inquiry shifts—in fact deepens—when I pass from the question "Is that a great horned owl?" to the question "Is that a dummy great horned owl?" When the depth of search is held constant, no failures—at least of the kind that Nozick suggests—of the closure principle arise.[22]

Nozick against the Skeptics

Finally, and briefly, we can note that Nozick's argument, even if it were decisive against skepticism as he represents it, would have no tendency to refute those skeptical arguments that philosophers have felt threatening. I have argued this point in my review of Nozick's *Philosophical Explanations*,[23] and this same point has been developed in detail by David Shatz.[24] Since it will help to clarify what the skeptical challenge amounts to, I will repeat these criticisms here.

We can say that a person Nozick-knows that p if her belief that p satisfies Nozick's conditions for knowledge. It seems clear that, for all I know, I may Nozick-know a great many things. Certain of my beliefs may track the truth without my having any good reason to suppose they do. So if knowing is identified with Nozick-knowing, then it should be clear that S can know (i.e., Nozick-know) that he is in New Brunswick without knowing (i.e., Nozick-knowing) that he is not in a jug on Alpha Centauri. With all the paraphernalia stripped away, this is precisely what Nozick's antiskeptical argument comes to. Concerning this I remarked:

> Now if this is the argument against the skeptic, I do not think it will overly embarrass him. Skeptics deal in arguments. Descartes, for example, maintains that there are no arguments, or at least no nontheological arguments, that can establish that we are not dreaming. Similarly Hume produces a no-argument argument to the effect that nothing can justify our reliance on inductive inferences because no argument can justify the principle of the uniformity of nature on which they rely. Perhaps something can be said in response to these skeptical arguments, but Nozick's reflections on the verb "to know" do not seem to bear on them. If our world is orderly and our projective mechanisms are well tuned to this orderliness, then our projections will often track the truth. Such projections will constitute Nozick-knowledge. But a person can Nozick-know that p without having the justified belief that p. (This is a

central feature of an externalist account of knowledge.) Therefore
the existence of Nozick-knowledge is compatible with the sound-
ness of Hume's argument for inductive skepticism and cannot
refute it. At most Nozick can issue only a small complaint against
the skeptic. If the skeptic goes on to say that his argument shows
that we do not (or cannot) know certain things, and if Nozick's
account of knowledge is essentially correct, then the skeptic can be
accused of *speaking* incorrectly. The skeptic should be content
with saying that certain fundamental principles are incapable of
justification, but [if Nozick's account of knowledge is correct]
should not go on to say that therefore we cannot know them to be
true. (822–23)

I do not think this criticism is stated pointedly enough. The cru-
cial fact is that Hume applied his skeptical doubts to *particular* judg-
ments, for example, that the sun will rise tomorrow or that bread
will continue to nourish. Nozick's response is that even if we can-
not know, for example, that the future will resemble the past, we
might still know that particular judgments of this kind are true
because, in fact, these beliefs track the truth. Hume, however, is
dealing with justification, and according to him we have no justifi-
cation for these particular beliefs, because we have no justification
for a principle we rely on in coming to them. We can then say, more
pointedly, that Hume's inductive skepticism with regard to particu-
lar beliefs is compatible with possessing Nozick-knowledge of them.
Thus the existence of Nozick-knowledge with respect to particular
inductively based beliefs does not refute Hume's skepticism con-
cerning them.

Dretske against Epistemic Closure

The attack on the principle of closure for knowledge did not begin
with Nozick. As Nozick grudgingly admits, Fred Dretske, in his arti-
cle "Epistemic Operators,"[25] "had all this and had it first" (689 n.
53).[26] I will close by looking at Dretske's attack on this principle for
two reasons. First, it attempts to make his rejection of this principle
seem intuitively plausible instead of basing it on arcane appeals to
possible-worlds semantics. Second, although, as it seems to me,
Dretske's position is wrong, it contains some things that are impor-
tantly right.

Dretske's central example for exhibiting a failure of the principle
of closure for knowledge is similar to my owl example.[27] It concerns
zebras:

> You take your son to the zoo, see several zebras and, when ques-
> tioned by your son, tell him they are zebras. Do you know they are
> zebras? Well, most of us would have little hesitation in saying that
> we did know this. . . . Yet something's being a zebra implies that it
> is not a mule and, in particular, not a mule cleverly disguised . . . to
> look like a zebra. (1015–16)

Dretske acknowledges that in this situation he did not know that
the creatures were not cleverly disguised mules. He continues:

> I part company with the skeptic only when he concludes from this
> that, therefore, you do not know that the animals in the pen are
> zebras. I part from them because I reject the principle he uses in
> reaching this conclusion—the principle that if you do not know
> that *Q* is true, when it is known that *P* entails *Q*, then you do not
> know that *P* is true. (1016)

I part company from Dretske because in the example as he describes
it I do not think we do know that the animals are zebras.

At least this much seems clear. If I observe the supposed zebras
while standing next to a cage with lions shaved and painted to look
like tigers, then, noting this, my epistemic attitude will change. In
particular, the level of epistemic scrutiny will rise, and I will not
accept the claim that these animals are zebras, at least until I have
ruled out the possibility that they are not some other kind of animal
painted to look like zebras. In fact, assurance on that point might not
be good enough, for people who go around painting mules to look
like zebras are presumably capable of all kinds of chicanery. Less
obviously, it seems to me that as soon as we are led to entertain the
possibility that the animals are painted mules, not zebras, *this itself*
is enough to lead us to withdraw the unqualified claim that we know
that the animals before us are zebras. Here I am suggesting that
merely dwelling on such possibilities is sufficient to raise our level
of scrutiny.

I expect resistance on this point. But consider the following:
Does *S*—to reintroduce him—know that it is a zebra *rather than* a
painted mule that he sees? The answer to this, as Dretske would
acknowledge, is surely no. But if *S* does not know that it is a zebra
rather than a painted mule, how can *S* know that it is a zebra? The
answer might be that *S* is asserting that the animals before him are
zebras rather than, say, deer, gazelles, or wildebeests. *S* is not assert-
ing that they are zebras rather than painted mules (or wax dummies
or holograms). Dretske seems to be saying something like this when
he remarks:

To know that x is A is to know that x is A within a framework of relevant alternatives, B, C, and D. This set of contrasts, together with the fact that x is A, serves *to define what it is that is known* when one knows that x is A. One cannot change this set of contrasts without changing *what a person is said to know* when he is said to know that x is A. (1022, emphasis added)

There is something right about this. When S identified the animals as zebras, his only concern was not to confuse this kind of animal with another—he was not worried, for example, about painted mules.[28] That is correct, but contrary to what Dretske seems to be suggesting, these considerations do not affect the *content* of what S claims to know. Suppose we discover that they are painted mules; then S is just wrong and doesn't know what he claims to know. And we might reject his knowledge claim even if we are not in a position to see that what he claims to know is actually false. We might see (as he cannot) that the cage next to the supposed zebras is filled with shaved lions painted to look like tigers. That strange discovery by itself might be sufficient for us to deny that S knows that the animals he is looking at are zebras, even if in fact they are. This shows that it doesn't help to suggest that S, in his knowledge claim, only *meant* to exclude a certain range of relevant possibilities, that is, that the animals were not zebras but, instead, some other similarly striped animals. What he said was that they were zebras, and if the additional information we have raises important concerns about the *strength* of S's evidence, then we will not grant that S knows, quite independently of the question whether what concerns us was something that ought to have concerned him.

This brings us back to a familiar point: In judging whether S's grounds are adequate to establish the truth of what he claims to know, we are in no way constrained by the practical considerations that guided S in forming his beliefs. He may have been entirely reasonable in not worrying about some remote defeating possibility. But, if, from our perspective, we see that this possibility is, or may well be, realized, then S's grounds (though perhaps responsibly invoked) do not establish the truth of what S claims to know, and then S does not know.

Notes

1. Dretske, 1971a.

2. (1) and (2) are Dretske's numbers which it will be convenient to preserve. (a2) is my own invention. It is oddly numbered to fit between Dretske's formulas (1) and (2).

3. Apropos of the (Ci) clause, Dretske remarks: "The appearance of the word 'know' in this characterization (in Ci) does not render it circular . . . since it can be eliminated by a recursive application of the three conditions until (Ciii) is reached" (13). I am skeptical concerning the possibility of any such recursive process of elimination, but this is a topic considered in the second half of this study.

4. Fogelin, 1967.

5. This schema was elaborated in various ways, for example by replacing "possesses" with "commands," but these details do not matter here.

6. I have made this same point in commenting on the third chapter of Robert Nozick's *Philosophical Explanations* (Nozick, 1981), where he restates Dretske's views at some extraordinary length.

7. In a footnote, Dretske explicitly states that "not only should we not require that S believe he has conclusive reasons in order to know, but also we should not require that he not believe that he doesn't have conclusive reasons" (17n). I am convinced that the first three BonJour counterexamples discussed earlier show that this is wrong, but, again, I don't think the BonJour counterexamples cut very deep, and analyses that are subject to them are easily corrected by introducing a (iii_p) clause.

8. Dretske's own example involves chemical indicators, but his remarks about this case are easily adapted to the somewhat simpler thermometer example I shall use.

9. Alvin Goldman uses this case as a counterexample to Nozick's analysis of knowledge. (See Goldman, 1986, 45).

10. Martin, 1975.

11. In "Tracking Nozick's Sceptic: A Better Method," Martin has applied essentially this same counterexample in a similarly decisive way against Nozick's subjunctivist account of knowledge (Martin, 1983).

12. I should make it clear that I have no general objection to possible-worlds semantics. Its use, for example, by Kripke for establishing completeness proofs in modal logic strikes me as money in the logical bank. Kripke does not, however, employ possible-worlds semantics in order to lay down *truth conditions* for subjunctive conditionals. With him I object to this practice. In *Naming and Necessity,* Kripke puts the matter this way: "The apparatus of possible worlds has (I hope) been very useful as far as the set-theoretic model-theory of quantified modal logic is concerned, but has encouraged philosophical pseudo-problems and misleading pictures" (Kripke, 1980, 48 n. 15). "A possible world," he tells us, "isn't a distant country that we are coming across, or viewing through a telescope" (44).

13. This second construction seems contrary to English usage. Stephen F. Barker makes the point nicely:

Nozick's handling of this grammar becomes especially awkward when, in regard to the provisos (1), (2), and (4), he writes, "Not only is *p* true and *S* believes it, but if it were true he would believe it" [Nozick, 1981, 176]. Here in a single remark he combines together

the asserting of a condition, worded as though he assumed its antecedent to be false, with the asserting of its antecedent. (Luper-Foy, 1987, 288–89)

Barker attempts to help Nozick by reviving the somewhat archaic construction "If p should be true, then S would believe it," but I confess that this doesn't sound right to me either. Barker's conclusion is that "in general the type of statement that is called for here is simply the conditional" (289). (Presumably, a subjunctive conditional for [3] and an indicative conditional for [4].)

Over against both Nozick and Barker, I don't think (4) should be expressed by a conditional construction at all. Given the situation as Nozick describes it, the most natural way to express (4) is as follows:

(4$_n$) S believes p because p is true.

Perhaps Nozick has a good reason for avoiding this natural construction, but he gives no indication what it is.

14. Quine, 1982, 23.

15. See Stalnaker, 1984, and Lewis, 1973.

16. In the remainder of this long footnote, Nozick speculates on the form a correct possible-worlds semantics for subjunctives might take, provided, that is, that some methods for determining distance metrics have been antecedently established.

17. See Moore, 1925, and Moore, 1939.

18. In chapter 8, I examine internalist accounts of meaning and belief content.

19. On these matters, I have profited from conversations with Michael Della Rocca concerning a yet-to-be-published paper he has written on this subject.

20. My preference for examples involving remote possibilities rather than standard skeptical scenarios is explained in the next chapter.

21. In general, in what I have called cases of epistemic luck, we refuse to grant that S knows something (e.g., that the object before her is a barn) because her grounds are not of a sufficient depth relative to the context as we recognize it.

22. Robert Audi has offered counterexamples to the principle of deductive closure for knowledge—which he refers to as the principle of deductive transmission of knowledge—that suffer from the same defect. Suppose, for example, that I hear a backfire, and, now quoting Audi, further suppose that

I am sufficiently acquainted with the sound to *know* that it is a backfire. Then, from what I know, it follows that it is not the sound of a firecracker with a similar muffled sound. Do I know that it is not? What if I have no evidence that there is no one around setting off such firecrackers? At most, I can say from general experience that this is improbable. It is not clear that, simply though validly

concluding it from my inferential ground, I know that the sound is not that of a firecracker with a similar muffled quality. One might now say that this just shows that I did not know in the first place that a vehicle backfired. But I do not see that we must say that. It may be equally plausible to say that because one *now* realizes that one's basis for believing this might not have been decisive, one *no longer* knows it, yet did know it in the first place. (1988, 77)

As with other examples we have examined, this one involves a shift in the level of scrutiny. Is it plausible, as Audi suggests, that once this shift in the level of scrutiny has occurred, I will now look back and say that previously I did know? That strikes me as completely implausible. The temptation to think otherwise is due, I believe, to being misled by those contexts where, for dramatic or rhetorical effect, we speak from the perspective of an assumed context, saying such things as "I knew exactly where it was, but couldn't find it." By saying "I knew" rather than "I thought I knew," I speak as I would have spoken then and thereby dramatically indicate the degree of my commitment at that time. I discuss such statements in *Evidence and Meaning* under the heading of amphibious—not amphibolous—constructions (Fogelin, 1967, 41–43).

23. Fogelin, 1983.
24. Shatz, 1987.
25. Dretske, 1970.
26. Nozick, 1981.
27. This is not surprising since my example was modeled on his.
28. Dretske's example would have been better if he had chosen an animal a bit less distinctive than a zebra. That way he could have contrasted the problem of confusing this animal with a somewhat similar other animal with confusing this animal with a fake version.

5

Epistemic Grace

*It is always by the grace of nature
that one knows something.*

In the Introduction, I pictured the Pyrrhonian skeptic going through
the world claiming to know certain things, and sometimes claiming
to be sure or even absolutely dead certain of them. The Pyrrhonian
skeptic freely participates in common epistemic practices, drawing
on all the practical distinctions embodied in them. These practices
are often fallible. Often this fallibility doesn't matter, since the price
of being wrong is not high. When the cost of error becomes exces-
sive, the skeptic, like others, may seek ways to improve these prac-
tices so that the chances of error are reduced. Pictured this way, the
skeptic is rather like Hume's moderate skeptic (whom he improperly
contrasted with the Pyrrhonian skeptic): cautious, agreeable, and
sane.

Historically, the Pyrrhonian skeptics have targeted the philoso-
pher as the object of their skeptical attack. Here the philosopher is
understood as someone who either (1) attempts to *replace* our com-
mon fallible modes of thinking about the world with new modes
that transcend them, or (2) accepts these common modes of think-
ing, but attempts to *ground* them in modes that transcend them. The
first is the revisionist strategy, the second the justificationalist strat-
egy. They are often combined in various proportions. Suggestions
have been made concerning the source of this philosophical drive: a
flight from peril (Dewey), a quest for unconditional comprehension
(Kant), a feeling of rancor against time (Nietzsche), language going on
a holiday (Wittgenstein), and so forth, but whatever it is that drives
people to engage in this kind of philosophy, the Pyrrhonian skeptic

is against it. The point of Pyrrhonian skepticism is to reject all such moves that attempt to transcend—rather than to improve or perfect—our common justificatory procedures.

But doesn't the philosophical critic of our common ways of making epistemic judgments have a legitimate complaint? How can we say that grounds *establish* the truth of a proposition while at the same time admitting that these grounds do not completely exclude the possibility that the proposition in question is false? If we recognize that a proposition might be false, don't we have *grounds for doubting* that proposition, and isn't having grounds for doubting some proposition incompatible with *knowing* it to be true? More pointedly, with inductive justification, it seems that grounds never completely establish the truth of a belief; thus we seem to be driven by the analysis presented in chapter 1 to the conclusion that there is no such thing as inductive knowledge. This is the cry of the deductive chauvinist, but perhaps it is time to stop calling names and try to say what is wrong with deductive chauvinism.

Justificatory Procedures

As we have seen, it is an important feature of the Gettier problems that they do not turn on imposing uncommonly high standards of justification on our epistemic performances. It is just this feature that separates the Gettier problems from many traditional skeptical ploys. As noted, the genius of Gettier's original argument is that it seems to generate profound problems while at the same time allowing us to use perfectly ordinary justificatory procedures. By an ordinary justificatory procedure—I am not trying to produce a definition—I mean a process that we go through when we are seeking justified true belief. Examples of justificatory procedures include such things as looking through a peephole to see who is at the door, consulting an expert to find out whether parking tickets are tax deductible, using a standard solution to check the accuracy of a chemical analysis, and so on. These justificatory procedures come in a wide variety of forms having, at most, a family resemblance to one another. Simple or complex, theoretical or nontheoretical, these are procedures we actually use; as a result of using them, we come to believe that we have grounds for accepting certain propositions as true.

Such justificatory procedures are the basis for both the judgment that *S* justifiably came to believe something and the judgment that *S*'s grounds establish the truth of something. In the former case we

are commenting on *S*'s performance—whether he carried out the procedure correctly, or perhaps whether he carried out the correct procedure. Even when positive, such an assessment does not commit us to the truth of what *S* claims to know. In the latter case we accept *S*'s justificatory framework as legitimate and, concurring in *S*'s evidential appraisal, join him in accepting a set of grounds as establishing the truth of what he claims to know. We have seen that it is possible for these two assessments to diverge. *S*'s performance may be adequate—that is, not irresponsible—but his grounds inadequate. This typically arises when we are privy to defeating information that *S* cannot be expected to know. (Gettier examples, in all their forms, build upon such situations.) We have also seen cases where *S*'s grounds, though adequate, are not invoked correctly by him. Such bad reasoning—if bad enough—from good grounds to a correct conclusion does not count as knowledge.

It is important to see that even in doing epistemology, we largely take these background justificatory procedures for granted; that is, our intuitions are guided by these common justificatory procedures with little explicit reflection on them. Even philosophers who emphasize epistemic responsibility operate in this way, for in presenting counterexamples, say against the externalist, they are typically trying to show us that externalism runs counter to our *common* ways of judging.

Doubts

Complementing the claim that reasons or grounds are given within justificatory practice is the further claim that *doubts* are expressed within justificatory frameworks as well. To express a doubt is to make a move—or, if you like, a countermove—within what I am calling a justificatory practice. To the question "What makes you doubt that?" the typical response is to point to some failure or inadequacy in meeting the appropriate justificatory standards. We are not always able to articulate the grounds for our doubts. Sometimes we can say little more than that something seems suspicious, peculiar, fishy, out of whack. Things can feel wrong or smell of paradox, though we will take such unarticulated doubts seriously only when the person who expresses them has experience in such matters.

For reasons that will emerge shortly, it will be helpful to distinguish various kinds (or levels) of doubt encountered in philosophical writing concerned with the theory of knowledge:

(A) Hyperbolic Doubts: those that rest on systematically une-liminable possibilities, as generated, for example, by so-called skeptical scenarios. (As I view an owl, I may wonder whether I am awake or asleep, or, perhaps, whether my experiences are caused by extraterrestrials who control my brain.)

(Bi) Eliminable but Impractical Doubts: those that rest on extremely unlikely possibilities that could be eliminated, but for which it would be a mark of obsessiveness to do so. (What I take to be an owl may be an ingeniously projected hologram.)

(Bii) Eliminable Legitimate Doubts: those that rest on possibil-ities whose elimination is demanded by the justificatory proce-dure being employed. In common parlance, these are *real possi-bilities*. (I recall that a young great horned owl can be somewhat similar to a mature long-eared owl, and I haven't really checked this out.)

Doubts of type (A), unlike doubts of types (Bi) and (Bii), have sometimes been declared illegitimate *because* they raise challenges that are systematically uneliminable. Wittgenstein adopted this line in both his early and his late writings. Thus, in the *Tractatus*, we find passages of the following kind:

> 6.5. If a question can be framed at all, it is also *possible* to answer it.

> 6.51. Skepticism is *not* irrefutable, but obviously nonsensical, when it raises doubts where no question can be asked.

Similar passages occur in *On Certainty*, written forty years later:

> OC, 3. If, e.g., someone says, "I don't know if there's a hand here" he might be told "Look closer." This possibility of satisfying one-self is part of the language-game. Is one of its essential features.

Such verificationist (or transcendental) arguments have enjoyed considerable vogue in this century. Here I will concentrate on the eliminable doubts that fall into the (Bi) and (Bii) categories.[1]

It is, I think, a plain fact that we often make serious knowledge claims in the face of nonexcluded remote (and sometimes not so remote) undercutting possibilities. Wittgenstein notes this in the fol-lowing passage from *On Certainty:*

> OC, 375. One must realize [that] the complete absence of doubt at some point, even where we would say that 'legitimate' doubt can exist, need not falsify a language-game. For there is also something like *another* arithmetic.

I believe that this admission must underlie any understanding
of logic.[2]

What interests me here is Wittgenstein's reference to *legitimate*
doubt, which, I take it, stands opposed to philosophical doubts of the
kind that fall into the (A) category given above. I take this to be a
deeper claim—a more far-reaching claim—than his attack on skepti-
cal scenarios just noted. Wittgenstein is saying that we accept
things, believe them, and act on them, in the face of identifiable
risks that we could eliminate but do not. He further claims that this
is simply how we employ our language, and accepting this is essen-
tial for a proper understanding of how our language works.[3]

Later in *On Certainty*, Wittgenstein applies these reflections to
our use of the verb "to know," but the passages are obscure and, in
fact, easy to get backward. Some lead-in will help:

> OC, 503. I look at an object and say "That is a tree," or "I know that
> that's a tree."—Now if I go nearer and it turns out that it isn't, I may
> say "It wasn't a tree after all" or alternatively I say "It *was* a tree but
> now it isn't any longer." But if all the others contradicted me, and
> said it never had been a tree, and if all the other evidence spoke
> against me—what *good* would it do me to stick to my "I know"?

Commenting on this passage, Wittgenstein continues:

> OC, 504. Whether I *know* something depends on whether the evi-
> dence backs me up or contradicts me.[4]

From the context, it is clear that by evidence Wittgenstein simply
means the facts of the matter. This brings us to this passage:

> OC, 505. It is always by the grace of nature [*von Gnaden der Natur*]
> that one knows something.

In making knowledge claims, or at least claims to empirical knowl-
edge, we rely on the grace of nature not to defeat us—at least when
we have behaved reasonably well. When so graced, we are said to
know. The philosopher, we might say, wants to replace this cove-
nant of grace with a covenant of work.

Although our common justificatory procedures do not demand
that we eliminate all potential defeaters, it is part of these procedures
to have built-in mechanisms for dealing with epistemically risky cir-
cumstances. For the most part we believe what people tell us, but some
circumstances—for example, dealing with the proverbial used-car
salesman—trigger the cautionary flag, "Don't believe everything you
hear." We usually trust our eyes in forming beliefs about someone's

actions, but at times it is worth remembering that the hand is quicker than the eye. And so on. In the language of the law, certain circumstances trigger strict scrutiny. What those circumstances are and how strict the scrutiny becomes is built into our justificatory procedures.

Levels of Scrutiny

Going back to Wittgenstein's remark that "complete absence of doubt . . . even where a 'legitimate' doubt can exist, need not falsify a language-game," we now see that, in a way, it is misleading. Properly understood, *legitimate* doubts are defined from within an epistemic practice. A legitimate doubt is one we *should* consider and remove before we make our knowledge claim. Failing this, we will not be said to know even if, by the grace of nature, our belief turns out to be true. In fact, since this is a typical Wittgensteinian point, I do not think he would reject it, and that, I think, is why he put the word "legitimate" in scare quotes. I think he uses this term in order to make vivid the point that we employ justificatory practices that do not demand that we exclude all everyday—as opposed to hyperbolic—grounds for doubt. In certain circumstances where we might doubt, we simply don't, and our language games rest on this. In *On Certainty*, Wittgenstein makes the point this way:

> OC, 509. I really want to say that a language-game is only possible if one trusts something (I did not say "can trust something").

The parenthetical remark is crucial, for it indicates that it is the fact that I trust something, not the inherent trustworthiness of what I trust, that lies at the base of my language games.

We can see the force of these last remarks by reflecting on what happens if we are relentless in attempting to eliminate remote defeating possibilities. Normally we ignore these possibilities, but if we dwell on them, our level of scrutiny will rise, and we will find ourselves unwilling to claim to know many things that we usually accept as items of knowledge. Do I, for example, know my own name? This seems to me to be as sure a piece of knowledge as I possess. But perhaps, through a mix-up in the hospital, I am a changeling. I'm really Herbert Ortcutt, and the person who is called "Ortcutt" is actually RJF. These things, after all, do happen. Given this possibility, do I know my own name? I'm inclined to say that I do not. Not only that, philosophical naïfs, namely those who do not see that such an admission may lead to forlorn skepticism, tend to agree. When pressed in this way, people are likely to become impatient, even

angry, yet, under pressure, most will acknowledge that strictly speaking—if you are going to be picky—given that they do not know they are not changelings, they do not know their own names.[5]

Seeing how reflection on remote possibilities can raise the level of scrutiny and thus lead us to withdraw epistemic commitment in a wholesale way shows that we do not have to introduce skeptical scenarios in order to generate skeptical problems. Reflection on unexcluded remote or *not so remote* possibilities can lead us to think we almost never know the things we claim to know. As long as we maintain this "intense view of things" we will be disinclined to think we know things or are justified in believing things that we normally accept without hesitation. When we return to practical affairs of life, our standards will return to their normal moderate level and this disinclination will fade. This is all very Humean, for it suggests that the application of certain concepts depends on the nonlinguistic fact that human beings lack the motivation, inclination, ability, or imagination to employ them in certain ways.

To philosophers imbued with the Cartesian ideal, or Cliffordian scruples, this line of thought may seem completely outrageous. Admittedly, some language games may rest on the practice of not doubting things that we might doubt, but this can hardly be true of the practice of making knowledge claims whose very point is to exclude all doubt. This, after all, seems to be a direct consequence of the analysis of knowledge presented in this work, namely:

> Knowledge that *p* is a belief that *p* justifiably arrived at on *grounds that establish the truth of p.*

But if that is correct, then it seems entirely natural to ask how grounds can establish the truth of something when at the same time there are undercutting possibilities that have not been eliminated. The answer to this—and this, I think, is Wittgenstein's central point—is that this is how we employ epistemic terms. We *assert* something, thus committing ourselves to it without reservation, while at the same time leaving eliminable refuting possibilities uneliminated. This is a fact—a fact about how we employ knowledge claims. Although this is misleading in a way to be noted in a moment, we might say that in making knowledge claims, we always (or almost always) assert more than we have a right to assert.

We have thus arrived at two factual claims. The first concerns what we mean when we say that *S* knows that *p*.

> I. "*S* knows that *p*" means "*S* justifiably believes that *p* on grounds that establish the truth of *p.*"

The second describes how people actually employ and actually understand knowledge claims:

II. When people claim to know things, they do not do so in the belief that they have eliminated all eliminable refuting possibilities, nor do their auditors suppose that they believe this.

Is There a Fact of the Matter in Knowing?

If a claim is made that *S* knows that *p*, is there a fact of the matter the presence of which would make this claim true, and the absence of which would make it false? That question is easily trivialized, for it is obvious that "*S* knows that *p*" is true, if and only if *S* knows that *p*. To try again, do epistemic claims obey the law of bivalence—that is, is it either definitely true or definitely false that *S* knows that *p*, or does this always depend on some further decision such that no simple yes or no answer is possible? Taking this question in the sense in which I think it is intended, my answer is that the position developed thus far implies *nothing* one way or another on that matter.

There is, it must be said, a fact or set of facts in virtue of which an epistemic claim will be correctly *deemed* true from within a justificatory framework functioning at a particular level of scrutiny. If a person claims to know someone's telephone number on the basis of looking it up in a telephone directory years out of date, then that person has acted irresponsibly relative to the standards for finding such things out and thus, given this fact, is not justified in his belief and does not know. (Of course, this person might have independent reasons for believing the directory, though out of date, is reliable, and considering that fact could change our epistemic assessment.) A person who finds the number in a current directory will be said to know—provided he turns out to be right. (Of course, reasons could arise for distrusting the accuracy of the current directory, perhaps because it is riddled with errors. We might also distrust it with respect to a particular person's number because he often moves.) So, from within a justificatory framework, setting aside questions of vagueness and indeterminacy, epistemic claims will be said to be determinately true or false, depending on the facts of the matter. From within the justificatory framework we will say that it is a fact of the matter that the person knows—or perhaps doesn't.

Notice that in the previous paragraph I have not said that *S* *knows* things from within a given justificatory framework. I have only described what he will be *said* to know—*deemed* to know— from within that framework. I have not relativized knowledge to

justificatory frameworks. I have simply described how epistemic claims operate within such frameworks. It is precisely here that I part company with so-called contextualist or perspectivist accounts of knowledge. I do not hold any of the following theses:

I. S is justified in believing that p if p is justified within the framework in which S is operating.

I may deny this for various reasons:

(a) I may reject S's justificatory framework. (S may be using astrological tables.)

(b) I may accept S's justificatory framework but think S has not used it correctly.

(c) I may grant that S has been epistemically responsible, but think his grounds have been defeated.

Nor am I committed to the following thesis:

II. S is justified in believing that p if p is justified within the framework in which I am operating.

That's wrong too, for the (b) and (c) shortcomings are still possibilities: S may have used an acceptable justificatory framework incorrectly or S may have been epistemically responsible, but, nonetheless, defeated in his justification by facts he could not be expected to know.

Finally, and this is the crucial point, I am not even committed to this thesis:

III. S is justified in believing that p if S bases his belief on the justificatory grounds that I accept using a justificatory procedure I accept as adequate.[6]

If S's epistemic performance satisfies these conditions, I will think it reasonable for me to hold that S is justified in his belief, but that, of course, does not entail that S *is* justified in his belief. There is nothing special about my justificatory framework.

To see the force of this last remark, consider what happens when we state matters in the first person:

IV. I am justified in believing that p if I base my belief on grounds I accept that are appropriate for the justificatory procedure I am using.[7]

This conditional fails because in claiming to know something (and thus to be justified in believing it) I am not relativizing my claim to

the grounds I possess nor to the justificatory framework I am employing. That I am doing neither of these things is clear because, if information appears that either defeats some of my reasons or shows that my justificatory procedure was inappropriate, I will then, looking back, say that my grounds did not justify me in believing what I did, though, perhaps, I may still think I was not irresponsible in thinking they did. Furthermore, I will not say I did know *then* within that framework, but do not know *now*, within the present framework. Of course, the perspective from which I make this new remark is no more insulated or unconditionally privileged than the original perspective.

In particular uses of knowledge claims, the character of the justificatory procedure invoked will be determined by the context in which the knowledge claim is made. But even if we invoke justificatory procedures in making knowledge claims, our knowledge claims are not claims *about* these justificatory procedures. The following is correct:

S knows that *p* if and only if S justifiably came to believe that *p* on grounds that establish the truth of *p*.

The following is incorrect:

S knows that *p* if and only if S came to believe that *p* on grounds that establish the truth of *p* within (i.e., relative to) some justificatory procedure.

The second formulation, but not the first, carries the contextualist or perspectivist implications that my position, which is expressed in the first formulation, does not.

I can illustrate the point I am trying to make through examining an ingenious counterexample raised by Fred Michaels to the position I am defending. The example is his, the wording mine.

> Suppose that Smith is worried whether her brakes are safe. To test them, she drives at a moderate speed and then hits the brakes quite sharply. When the car comes to a smooth stop, she concludes that they are safe. Let's suppose two things: (i) this is a good rough-and-ready way of testing brakes, and (ii) Smith's brakes are, in fact, safe. Thus by applying a standard procedure and by getting the correct result, Smith can, on the present approach, be said to know that her brakes are safe.
>
> Next suppose that Jones, an auto mechanic, tests the brakes in the same way, and then concludes, as Smith did, that the brakes are safe. Jones, however, knows that this test, thought reliable in a rough-and-ready way, sometimes does not detect faulty brakes.

Given the epistemic irresponsibility of his performance, I think we are inclined to say that Jones, even though it is true that the brakes are safe, does not know this. This leads to the seemingly paradoxical result that, given the same evidence, an amateur can know something that an expert cannot.

If asked what *I* think about this situation, I would say that neither Smith nor Jones knew that the brakes were safe, for even though they are safe, thanks to Michaels, I am privy to special information about the proper way of testing brakes.

Another related objection comes from Storrs McCall who suggested that, on my account, whether *S* knows something depends on who is making the epistemic assessment. From *A*'s perspective, *S* may know that *p*, but not from *B*'s. That's wrong. What is right is that *A*, from his perspective, may grant that *S* has justifiably come to believe that *p* on grounds that establish the truth of *p*, whereas *B*, from his perspective, will not grant this. That *A* grants something does not make it so; and, equally, that *B* does not grant this does not show that it isn't so.

This brings us back to the question whether there is a fact of the matter that determines whether an epistemic claim is determinately true or determinately false. My answer is that my analysis implies nothing one way or another on this matter, and this is precisely because nothing in this analysis either privileges or refuses to privilege particular justificatory frameworks. For all that has been said so far, there may be one ultimate justificatory framework that grounds all others. There may be a plurality of justificatory frameworks that ground various domains of knowledge. There may be no justificatory framework that stands up under the unlimited heightening of scrutiny. In fact, this third possibility strikes me as being correct, but that is not something that follows from the analysis of knowledge claims I have presented. It is something that has to be shown in detail by examining various attempts to produce philosophical theories of justification. That's the task of the second part of this study.

It's Hard to Say

Finally, it is worth noting that it is difficult to express the view I am presenting in a fashion that is not misleading. I claim to be doing nothing more than describing what we mean when we claim to know something and (following Wittgenstein) also describing the standards we use in employing such claims. Yet it is difficult to state these matters in ways that do not suggest a strong evaluative com-

mitment. In various places, Wittgenstein seems to encounter the same difficulty. Consider the following remark:

OC, 166. The difficulty is to realize the groundlessness of our believing.

Usually, to say that something is groundless is to offer a criticism: it's to say that a ground that ought to be present is missing. That is to reverse Wittgenstein's point. It is also, as noted earlier, misleading to speak of ignoring legitimate doubts, since a legitimate doubt is precisely one that should not be ignored. Talking of *taking things for granted*, or saying that in making a knowledge claim *we almost always assert more than we have a right to*, is misleading in the same way. All these expressions are typically used to formulate specific criticisms from within a justificatory framework, and they require quite specific responses. It thus can be misleading when they are intended merely descriptively. Unfortunately, it is difficult to find alternative expressions that are free from this difficulty.

Summary of Part I

I hold that the theory of knowledge, in its traditional form, has been an attempt to find ways of establishing knowledge claims from a perspective where the level of scrutiny has been heightened by reflection alone. The dogmatic skeptic privileges this heightened perspective and then, skirting self-refutation, claims that nothing is known. The Pyrrhonists (as I have described them) resist both responses because they refuse to privilege the philosophical perspective that the dogmatic skeptics and their opponents share. When they hypothetically enter the philosophical perspective, they will be inclined to say that nothing is known. Here they simply report how things strike them. For the most part, however, they will occupy a normal perspective where skeptical scenarios and remote (and not so remote) defeaters are simply ignored. They will then speak and act in common, sensible ways.

One way of showing that this Pyrrhonian description of our epistemic situation is wrong would be to show that certain knowledge claims can be certified in the face of unrestricted levels of scrutiny. This would not be a particularly interesting result if this could be done only for a small number of boring knowledge claims. We want more from the epistemologist than a bare refutation of skepticism. We want a vindication of a suitably wide range of knowledge claims that are worth making. The attempts to carry out such a program are

called theories of justification. In the second part of this study I will argue that there is little prospect that such a program can be carried out successfully. But before turning to this largely negative project, let me list what I take to be the positive results of this inquiry thus far. I think I have given an account of knowledge claims that is adequate in the following ways: (1) It not only avoids Gettier problems, but also gives us a correct diagnosis of the origins of those problems. (Derivatively, it explains why discussions of Gettier problems have taken the form they have.) (2) It does this without *generating* skeptical problems. It does not generate skeptical problems because such problems depend on privileging unrestricted levels of scrutiny—a move quite independent of the analysis I have presented. (Derivatively, the analysis explains why, under certain circumstances, skeptical doubts can have genuine force through taking advantage of the shifting levels of scrutiny already present in our common justificatory practices.)

Notes

1. For a critical examination of transcendental arguments, see Stroud's "Transcendental Arguments" (1968).

2. When Wittgenstein speaks of logic in the final sentence, he is using this term in a traditionally wide sense that includes, among other things, theory of meaning. Wittgenstein sometimes uses the term "grammar" in this wide sense as well. The reference to *another* arithmetic is presumably intended to illustrate the point that even something as seemingly unalterable as arithmetic has coherent alternatives, and therefore, at bottom, rests on a kind of brute acceptance as well.

3. We might add that this is not unreasonable, since often the transaction costs of eliminating all legitimate doubts are higher than the penalties incurred by making faulty commitments, which, sometimes at least, can simply be taken back, and adjustments made. I'll return to this point.

4. The passage continues, "For to say one knows one has a pain means nothing." To many this may seem a peculiar way for Wittgenstein to illustrate his point, for his claim that it is meaningless to assert that one knows one has a pain, far from being meaningless, may seem a paradigm of knowledge. We can at least see what Wittgenstein was up to if we recall that it was his view that first-person utterances of pain were not assertions that a pain had occurred, but rather expressions of that pain. That's why Wittgenstein thought it made no sense to ask a person to produce his reasons for saying he is in pain. This may not be a persuasive account of the way in which first-person utterances of pain function, but at least it explains why Wittgenstein thought the illustration apt.

5. Dretske would deny this. Wittgenstein might have as well, since he held, quite explicitly, that skeptical doubts are meaningless. This is not a view I share with Wittgenstein. I will return to this topic in Part II of this study.

6. I'll ignore the BonJour problem that might arise if S also believes further things that would undercut his justification.

7. This is not a self-sealing (and hence an empty) demand, for I might come to see, on my own terms, that a claim of mine is not justified on the grounds that I have used to support it.

Appendix A: The Lottery Paradox and the Preface Paradox

The account of knowledge claims in Part I bears in interesting ways on the so-called lottery and preface paradoxes. I will begin by examining the lottery paradox.

The Lottery Paradox

The lottery paradox emerges from reflections of the following kind: Since most of our empirical knowledge claims are inductively based, their probability typically falls short of 1.[1] It seems, then, that we are willing to say that something counts as empirical knowledge provided the level of probability is suitably high. But however high we fix the suitable level of probability, it is possible to show that fixing the probability at that level leads to paradoxical results. To see this, suppose, being very cautious, we count something as empirical knowledge only if it has no more than one chance in ten million of being false. Next, imagine a lottery containing ten million tickets, where one will win, and each has an equal chance of winning. Under these conditions, ticket #1 has no more than one chance in ten million of winning. Therefore, given the above stipulations, it follows that we *know* that this ticket will not win. The same reasoning holds for each of these ten million tickets, so we seem to know of *each* of them that it will lose, and hence that *all* of them will lose. Schematically:

> We know that ticket #1 will lose.
> We know that ticket #2 will lose.

> We know that ticket #10,000,000 will lose.
> Therefore:

> We know that tickets #1 and #2 and . . . and #10,000,000 will lose.
> We know that that's all the tickets there are.
> Therefore:
> We know that all the tickets will lose.

This result, of course, is paradoxical, for we cannot know that all the tickets will lose, since we know from the setup of the example that it is false that all the tickets will lose.

Our response to this paradox will depend on whether we (i) continue to hold that we know of each ticket that it will lose, or (ii) give this up.

(i) If we continue to hold that we know of each of the tickets that it will lose, then it seems there are only two ways to avoid the paradox. First, denying the principle of closure for knowledge, one could reject the final inference in the argument, that is,

> We know that tickets #1 and #2 and . . . and #10,000,000 will lose.
> We know that that's all the tickets there are.
> Therefore:
> We know that all the tickets will lose.

This strikes me as desperate in itself, and, in any case, I have already argued that there are no good reasons for abandoning the principle of closure for known implication.

A more interesting suggestion is that we abandon the so-called conjunction principle (or principle of agglomeration) for knowledge, namely, we give up the following principle:

$$(Kp \mathbin{\&} Kq \mathbin{\&} \ldots \mathbin{\&} Kt) \Rightarrow K(p \mathbin{\&} q \mathbin{\&} \ldots t)$$

Of course, we should expect this principle to fail on a probabilistic approach to empirical knowledge, since the conjunction of two propositions each with a suitably high level of probability need not itself have a suitably high level of probability to count as knowledge. The difficulty with this solution is that the conjunction principle seems obviously correct, and anyone impressed with this may take the necessity of denying it as adequate grounds for rejecting probabilistic accounts of knowledge.

(ii) The second main strategy for avoiding the lottery paradox is to give up the claim that we do, in fact, know of each of the tickets that it will lose. We don't know that ticket #1 will lose, because we have been told that it has some chance of winning. In the language

of the theory presented in this work, we do not know that ticket #1 will lose since being told that there are nine million, nine hundred ninety-nine thousand, nine hundred ninety-nine chances in ten million that it will lose does not establish the truth of the claim that it will lose. I think this straightforward response squares with common opinion on this matter.

There is, however, a difficulty with this straightforward solution: it seems that we sometimes claim to know things when the probability of their not being true, though remote, is no more remote than the probability that a particular ticket will lose in the lottery. I don't know how to estimate it, but it seems to me that the probability that I spoke to Fred Dretske in his office rather than to a twin brother could be of the same order as the probability that ticket #1 will not lose. If not, we can add more tickets to the lottery. How is it, then, that I know that it was Dretske who was in his office, but do not know that ticket #1 will lose? There are two ways we might deal with this problem: (iia) we can try to find some way of separating the cases that explains why we fail to know in the lottery case, but do know in the Dretske-twin case, or (iib) we can abandon the claim that we know in the Dretske-twin case.

(iia) A tempting way to pursue the first line is to claim that knowledge is not tied to probability, but rather to something like expected value. When the payoff is high enough, it may be reasonable to buy a ticket in a lottery, and it does seem strange to say that it is reasonable to buy a ticket when you know you will lose.[2] But I do not think this is the correct diagnosis of the lottery paradox, for my intuitions do not change when the probabilities stay the same and the payoff is reduced to a penny—or to nothing. Even when it becomes wholly unreasonable for me to buy a ticket relative to expected values, it still strikes me that I do not know that I will lose.

(iib) Reflections in chapter 5 concerning levels of scrutiny suggest that we adopt the second approach and reject the claim that we possess knowledge in cases where our chances of being wrong are on a par with the probability that a particular ticket will lose. Thus, concerning the Dretske-twin example, once the possibility of an identical twin has been introduced, then I, at least, am unwilling to say that I know it was Dretske whom I saw. The difference between the lottery example and the Dretske-twin example is that in the lottery example a heightened level of scrutiny is triggered by practical considerations concerning the reasonableness of a wager; in the Dretske-twin example, it is triggered by nothing more than reflection on an unexcluded defeating possibility.

Returning to our original question, "How is it, then, that I know that it was Dretske who was in his office, but do not know that ticket #1 will lose?" we now see that it was badly put. Once the level of scrutiny is staked at the same levels, we give up the claim to know in both cases.

The Preface Paradox

Robert Stalnaker offers the following statement of the preface paradox:

> In his preface to his historical narrative, the author admits that he has undoubtedly made some mistakes—that some of the statements he made in his narrative are false. He is not confessing to insincerity—he continues to *believe* everything he wrote—he is just confessing fallibility. It does not take excessive modesty to believe that *some* of one's many beliefs are false. This is only reasonable. Yet to believe this is to believe each member of a set of propositions that are recognized to be inconsistent. If these propositions are conjoined and their consequences accepted, the result would be to accept the truth of every proposition. (Stalnaker, 1984, 92)

Stalnaker goes on to explain the difference between the lottery paradox and the preface paradox as follows:

> The assumption that high probability is sufficient for acceptance is essential to the argument of the lottery paradox. . . . But the paradox of the preface does not depend on that assumption and cannot be answered in the same way. It cannot be plausibly denied that the author *accepts* the truth of each of the statements made in the narrative, nor can it be denied that he accepts that at least one of those statements is false. But it also seems plausible to say that the author accepts, or at least commits himself to, any conjunction of the statements in the narrative. (92)

The difference between the two paradoxes is that the lottery paradox concerns *knowledge* whereas the preface paradox concerns *rational acceptance, justified belief,* or some such notion. Different accounts have been given of these concepts, but for our present purposes, we cannot treat the preface paradox as simply the lottery paradox unless it can be rational to accept something, be justified in believing it, and so forth, without that belief thereby being true. This is a feature that all these notions share with the idea I have tried to capture in the (iii$_p$) clause in the analysis I have presented. We can, as I have noted, justifiably come to believe something or be epistemically responsible in believing something although what is believed is false.

In fact, if we read the preface paradox as simply concerning epistemic responsibility, then nothing paradoxical emerges. Our historian may have been completely responsible in making each of his individual claims, yet it would be irresponsible for him to claim that all of them are true. It would be irresponsible for him to claim this because of the well-known fact that however responsible you are, if you make enough historical claims, the chances are great that at least one of these claims will be mistaken. Thus the conjunction principle does not hold for epistemic responsibility. The same point can be made with respect to a lottery example. It would not be epistemically irresponsible to believe that ticket #1 will lose, and base one's plans on this belief. The same is true for each of the other tickets. It would, however, be epistemically irresponsible to believe that all of the tickets will lose. We would, for example, welcome the opportunity to make an even-money bet on each of the tickets that it will lose, whereas it would be foolish to make an even-money bet that all of them will lose. Again we have a failure of the conjunction principle for epistemic responsibility and related notions.

In contrast, if we treat rational acceptance on an analogy with the truth-committing notion of adequate grounds (or conclusive reasons), then, as already indicated, the preface paradox becomes indistinguishable from the lottery paradox. Imagine the historian saying, quite explicitly, "For each proposition in this book, I have grounds adequate to establish its truth, but I do not have grounds adequate to establish their conjunction." Does this sound plausible? Not to me. If that is correct, then the conjunction principle does hold for adequate grounds (or conclusive reasons).

Although the conjunction principle holds for adequate grounds, a failure to understand the phenomenon of shifting levels of scrutiny can lead us mistakenly to reject that principle, just as it can lead us mistakenly to reject the principle of epistemic closure for known implication. To see how this can happen, suppose our historian, who honestly believes he has grounds that establish the truth of each of his assertions, realizes the fact that his book, like others that have been carefully written, is bound to contain some mistakes. We now have an instance of the lottery paradox, for it seems that the historian has adequate grounds for each member of a set of propositions, but does not have adequate grounds for their conjunction. This result suggests, contrary to what I have said, that the conjunction principle fails for adequate grounds (or conclusive reasons). Over against this, my suggestion is that reflecting on the almost inevitable occurrence of error in his work will lead the historian to raise his

level of scrutiny concerning the *individual* claims he has made; for this reason he will revise downward his original epistemic assessment of them. In other words, his reflections on the almost inevitable occurrence of error will lead him to withdraw his claim that he had grounds to establish the truth of each of the assertions he has made, so the failure of the conjunction principle is avoided.

My suggestion for preserving the conjunction principle for adequate grounds may seem implausible—even desperate. That it is neither can be shown by taking any historical work from the shelf and opening it at random. Frederick C. Beiser's fine work, *The Fate of Reason*,[3] comes to hand. On page 65 we read: "On September 1, 1773, Reimarus duly sent a summary of Mendelssohn's August 16 letter to Jacobi." Beiser cites the Jacobi Werke, IV/1 43–46, to support this claim. Without impugning the author's scholarship, it is easy enough to imagine how this claim could turn out to be false. Perhaps Reimarus wrote the letter on September 1, but did not get around to sending it until September 2. Perhaps Reimarus had the day wrong. And so forth. This is all silly nit-picking, but it is precisely because of the possibility of mistakes of this kind that historians will admit that some errors are bound to have slipped into their work. Furthermore, with the level of scrutiny raised in this way, the historian himself will have no difficulty in identifying any number of candidates as the source of possible error. Of these claims he will acknowledge that—strictly speaking—he lacks grounds establishing their truth, though he is not likely to count this as a defect in his work.

The conjunction principle holds for adequate grounds; it does not hold for epistemic responsibility. One last question: Does it hold for knowledge? Before we can get to the core of this question, some side issues paralleling those that arose for epistemic closure for known implication must be set aside. Take at random any hundred (or thousand) propositions that S knows to be true and conjoin them; does S know this proposition to be true? The difficulty here is that S has never entertained this proposition and may be incapable of doing so. It is also possible that S is willing to accept the conjunction of two propositions that she accepts because she thinks every inference is valid. To avoid difficulties of this kind, let us suppose that S is a logically competent reasoner and let us suppose that her intellectual capacities are up to the task at hand. We will also suppose that she reasons quite explicitly, from her knowledge that p and her knowledge that q, to the conclusion that she knows that p & q. Could there be anything wrong with this reasoning? It seems to me that

there couldn't. If that is correct, then the conjunction principle for knowledge holds, at least for competent reasoners dealing with propositions within their ken.

The Conjunction Principle for Knowledge

If these reflections are correct, they show that one way of dealing with the lottery paradox, namely rejecting the conjunction principle for explicit competent reasoning, fails, but they also seem to raise a difficulty for the analysis of knowledge claims that I have presented. Stated abstractly, the difficulty is as follows. The analysis that I have presented contains two clauses: (i) an adequate-grounds (or conclusive-reason) clause, which obeys the conjunction principle, and (ii) an epistemic responsibility (or rational-acceptance) clause, which does not. The legitimacy of the conjunction principle in the first case and its failure in the second allow us to deal, respectively, with the lottery paradox and the preface paradox. That is a nice result, but unfortunately the analysis suggests that the conjunction principle for knowledge should also *fail*. It should fail, it seems, because it contains a component, the epistemic responsibility clause, for which the conjunction principle does not hold. Furthermore, the failure, if it does occur, will not depend on attributing logical incompetence or any other intellectual failing to S.

To see the way out of this difficulty, we can recall why the conjunction principle fails for epistemic responsibility. We saw that S could be responsible in accepting p and responsible in accepting q because each had a suitably high level of probability, but S would not be responsible in accepting p & q because this conjunction lacks a suitably high level of probability. We can now resolve the problem raised in the previous paragraph simply by noting that this circumstance cannot arise if S's beliefs satisfy the adequate-grounds (or conclusive-reason) clause in the analysis I have presented. The adequate-grounds clause forecloses the possibility of such a diminution in probabilities.

We thus arrive at the following results concerning the conjunction principle, at least when it is applied to explicit competent reasoning:

It holds for adequate grounds (or conclusive reasons).
It can fail for epistemic responsibility (or rational acceptance).
It holds for knowledge (or belief justifiably arrived at on grounds that establish its truth).

Notes

1. Kyburg, 1961.
2. At one time I found this approach attractive. My colleague Walter Sinnott-Armstrong independently developed similar views in an unpublished essay he wrote a number of years ago. Others, I think, have been attracted to it as well.
3. Beiser, 1987.

This page intentionally left blank

II

AGRIPPA AND THE PROBLEM OF JUSTIFICATION

This page intentionally left blank

6

Agrippa and the Problem of Epistemic Justification

In the *Outlines of Pyrrhonism*, Sextus Empiricus offers the following description of Agrippa's Five Modes Leading to the Suspension of Belief:

> The later Skeptics laid down Five Modes leading to suspension [of belief], namely these: the first based on discrepancy, the second on regress *ad infinitum*, the third on relativity, the fourth on hypothesis, the fifth on circular reasoning. That based on discrepancy leads us to find that with regard to the object presented there has arisen both amongst ordinary people and amongst the philosophers an interminable conflict because of which we are unable to choose a thing or reject it, and so fall back on suspension. The Mode based upon regress *ad infinitum* is that whereby we assert that the thing adduced as proof of the matter proposed needs further proof, and this again another and so on *ad infinitum*, so that the consequence is suspension, as we possess no starting-point for our argument. The Mode based upon relativity . . . is that whereby the object has such or such an appearance in relation to the subject judging and to the concomitant percepts, but as to its real nature we suspend judgment. We have the Mode of hypothesis when the Dogmatists, being forced to recede *ad infinitum*, take as their starting-point something which they do not establish by argument and without demonstration. The Mode of circular reasoning is the form used when the proof itself which ought to establish the matter of inquiry requires confirmation derived from that matter; in this case, being unable to assume either in order to establish the other, we suspend judgment on both. (*PH*, 1: 164–69)

The Problem

As stated in the Introduction, I will define the philosophical problem of epistemic justification as the attempt to solve the Agrippa problem, that is, as the attempt to meet the skeptical challenge presented by Agrippa's Five Modes Leading to the Suspension of Belief. It says nothing against this way of characterizing contemporary discussions of epistemic justification that most recent writers on this subject seem never to have heard of Agrippa and his Five Modes. As we shall see, the dialectical structure of Agrippa's Five Modes and the dialectical structure of the contemporary debate on epistemic justification are strikingly similar.

In recent literature, what I am calling the Agrippa problem is often referred to as the infinite regress problem. I find this characterization too narrow, for the problem that presents itself is not simply that of avoiding a bad infinite regress; the challenge is to avoid this regress without falling into a bad form of circularity or a bad form of unjustified acceptance. Indeed, I think speaking of the infinite regress problem exerts a subtle, but strong, pressure on the discussion. If we think the threat of an infinite regress of reasons as the central challenge to justified belief, then theories, despite their own difficulties, may lay claim to our acceptance just because they seem to deal with this single aspect of the Agrippa problem. If, however, we begin with an antecedent horror of circularity, an appeal to an infinite regress might recommend itself as a way out. It is important, then, not to grant unwarranted dialectical advantages, but to insist, instead, that a philosophical theory of justification must simultaneously avoid involvement in a bad infinite regress, in a bad form of circularity, and in a bad appeal to unwarranted assumption. The Agrippa problem poses these challenges in an evenhanded way.

Before looking in detail at contemporary attempts to solve the Agrippa problem, it is worth asking how it can arise as a serious philosophical challenge. How does the Agrippa problem take on importance? There seem to be two factors involved. The first is a commitment to a strong *normative* principle of epistemic justification. W. K. Clifford expressed such a commitment in these words made (in)famous by William James: "It is wrong always, everywhere, and for anyone, to believe anything upon insufficient evidence" (2:186).[1] A similar commitment is expressed by Paul K. Moser in terms of epistemic irresponsibility: "To accept a proposition in the absence of good reason is to neglect the cognitive goal of truth. Such acceptance, according to the present normative conception of justifi-

cation, is epistemically irresponsible" (4).[2] Laurence BonJour makes essentially the same claim: "To accept a belief in the absence of . . . a [good] reason, however appealing or even mandatory such acceptance might be from some other standpoint, is to neglect the pursuit of truth; such acceptance is, one might say, *epistemically irresponsible*. My contention here is that the idea of avoiding such irresponsibility, of being epistemically responsible in one's believing, is the core of the notion of epistemic justification" (8).[3] Moser and BonJour have competing views concerning epistemic justification: With respect to empirical knowledge, Moser is a foundationalist whereas BonJour is a coherentist, but in their commitment to very high standards of epistemic responsibility, they are in complete agreement.

But Cliffordism, if I may use the term, does not by itself convert the Agrippa problem into a serious challenge. Confronted with the Agrippa problem, the Cliffordian might decide, in Berkeley's words, "to sit down at last in forlorn skepticism"—and let it go at that. For the Agrippa problem to emerge as a serious problem, the Cliffordian must further believe that knowledge does exist or at least could exist. This is just the stance that, with more than a little bravado, many contemporary epistemologists take. Here is John L. Pollock:

> In typical skeptical arguments, we invariably find that we are more certain of the knowledge seemingly denied us than we are of some of the premises. Thus [shades of G. E. Moore] it is not reasonable to adopt the skeptical conclusion that we do not have that knowledge. The rational stance is instead to deny one or more of the premises. In other words, a typical skeptical argument is best viewed as a *reductio ad absurdum* of its premises, rather than proof of its conclusion. (6)[4]

In the common way of speaking, we do know all sorts of things: our own names, the capitals of various states, and the like. But it is an altogether different question whether we know such things (or anything) relative to Cliffordian standards. The assumption, then, that drives justificationalist programs in both their foundationalist and nonfoundationalist modes is that we do (or could) possess knowledge that conforms to Cliffordian standards. The task of a theory of empirical justification is to show how this is possible. I'll call this the Cliffordian project.

It is important to see that the Pyrrhonists themselves are not engaged in the Cliffordian project, for the Pyrrhonist does not hold that it is epistemically irresponsible to believe things on insufficient evidence. As always, the Pyrrhonist simply takes the standards of

the dogmatist at face value and holds the dogmatist to them. The Pyrrhonist invokes the Five Modes and similar devices for dialectical purposes.

Specifically, the Five Modes present a challenge to the Cliffordian in the following way. Two of Agrippa's modes, discrepancy and relativity, trigger a demand for justification by revealing that there are competing claims concerning the nature of the world we perceive. Given this competition, it would be epistemically irresponsible for the Cliffordian to choose *without argument* one of these competing claims over the others. Thus the modes of discrepancy and relativity force anyone who makes claims beyond the modest expression of opinion to give reasons in support of these claims. I'll call these two modes—discrepancy and relativity—the *challenging* modes.

The task of the remaining three modes—those based on regress *ad infinitum*, circularity, and (arbitrary) hypothesis—is to show that it is impossible to complete this reason-giving process in a satisfactory way. If the Pyrrhonists are right, no argument, once started, can avoid falling into one of the traps of circularity, infinite regress, or arbitrary assumption. I'll call these three modes the *dialectical* modes.[5]

We can now compare the way Agrippa employs his three dialectical modes with a sketch that BonJour gives of the options open to one attempting to provide a theory of empirical justification:

> *Prima facie*, there are four main logical possibilities to the eventual outcome of the potential regress of epistemic justification . . . (1) The regress might terminate with beliefs which are offered as justifying premises for earlier beliefs but for which no justification of any kind, however implicit, is available when they are challenged in turn. (2) The regress might continue indefinitely "backwards," with ever more new empirical premise-beliefs being introduced, so that no belief is repeated in the sequence and yet no end is ever reached. (3) The regress might circle back upon itself, so that if the demand for justification is pushed far enough, beliefs which have already appeared as premises . . . earlier in the sequence of justificatory arguments are again appealed to as justifying premises. (4) The regress might terminate because "basic" empirical beliefs are reached, beliefs which have a degree of epistemic justification which is not inferentially dependent on other empirical beliefs and thus raises no further issues of empirical justification. (21)

A similar passage is found in Moser: "We have, then, at least four possible accounts of inferential justification: inferential justification via infinite regress, via justificatory circles of some sort, via the

unjustified, and via immediate justification" (24).[6] The four options found in BonJour's and Moser's writings match up with Agrippa's three dialectical modes, with one variation. They share the modes of circularity and regress *ad infinitum*, and, depending on how one looks at it, BonJour and Moser either split Agrippa's mode of hypothesis into two components or they add, as a fourth option, the foundationalist maneuver.

BonJour/Moser	Agrippa
(1) Arbitrary Assumption	Hypothesis (1)
(2) Indefinite Regress	Infinite Regress
(3) Circularity	Circularity
(4) Foundationalism	Hypothesis (2)

Both BonJour and Moser, as justificationalists, are opposed to Agrippa's Pyrrhonian program of showing that justification is not possible, but they both, in their own way, exploit the fundamental structure of the Agrippa problem. As a defender of a coherence theory of empirical justification, BonJour argues that the unacceptability of options (1), (2), and (4) leaves (3) as the only possible option for vindicating empirical-knowledge claims. Following this, his positive task is to show that circularity (properly understood and qualified), far from being a defect, is an epistemic virtue in a system of beliefs. The same general strategy is employed by foundationalists, for example Chisholm and Moser, who, again using a process of elimination, argue from the unsatisfactory character of circularity, infinite regress, and arbitrary assumption to the conclusion that some beliefs must be justified without being justified by appeal to other beliefs. The foundationalist's task is then to identify some class of justified beliefs that are not justified by other beliefs, explain how they acquire their justified status, and then show how these beliefs can provide a foundation for at least a tolerably large portion of those other beliefs we count as knowledge.

This, then, is the general form of a treatise on empirical justification: a complex disjunctive syllogism eliminating all contending positions, followed by a constructive attempt to show that the remaining Agrippan mode (properly understood and qualified) provides a way of solving the Agrippa problem.

Success Conditions on Theories of Justification

Before I turn to an examination of these competing justificationalist theories, it will be useful to set down minimal success conditions that we can apply evenhandedly to them all. *First,* the presentation

of a theory of justification should be governed by what I shall call the principle of *philosophical candor.* In conformity with this principle, the author should specify, as desiderata, just which beliefs she takes to be justified, and which not, for without this specification the *scope* of the theory will be indeterminate. It is sometimes a conse-quence—often played down—of a particular theory of justification that we are justified in believing very much less than we initially supposed we were. For example, traditional idealistic versions of coherence theories often have the consequence that we are justified in believing almost none of the things that, in common life, we think we are justified in believing. The traditional coherence theo-rist held that the isolated, fragmented, partial beliefs of everyday life must be run through by reason and utterly transformed before they can be taken up into that absolute whole where justified true belief *alone* resides. As we shall see, recent (non-idealistic) coherence the-ories have similar consequences with respect to everyday belief, though, perhaps, not in the same extravagant way. In any case, unless we are told what sorts of beliefs a theory of justification will yield as justified, we will have little understanding of the scope of the theory in question. It is entirely possible that a theory of justifi-cation may, in fact, be equivalent to radical skepticism.

   In addition to the first demand, that we be told what beliefs the theory of justification takes to be justified, we can insist, as a *second* success condition, that the theory show in some detail just how these same beliefs are justified. If we examine the actual writings of philosophers who have produced theories of justification, we find that, often enough, no serious attempt is made in this direction. For the most part we merely find elaborate discussions of certain *ante-cedent* issues. These discussions of antecedent issues arise because each of the standard positions has its standard difficulties that have to be resolved before the theory can—as the saying goes—get off the ground. Thus the coherentist must say something about the possi-bility of mutually exclusive yet equally coherent systems. The foun-dationalist must tell us how a belief can be justified by something that is not itself another belief. And so on. The tacit assumption that dominates the philosophical literature on epistemic justification is that, if it is possible to show that one's own theory can "get off the ground" whereas all others remain grounded, then the battle is won. But this is wrong because it begs the question. A philosopher who attempts to address the Agrippa problem cannot reason in the fol-lowing way: There must be some answer to the Agrippa problem, and, since my answer is the only conceptually coherent answer, *it*

must be the correct answer. There is, however, no antecedent reason to suppose that there is an adequate response to the Agrippa problem; indeed, on its face, the Agrippa problem seems quite unanswerable.

These reflections lead to a *third* success condition that we can place on any theory of epistemic justification: An answer to the Agrippa problem may not beg the question by assuming *for argumentative purposes* that there *must* be some positive solution to it. It is precisely this assumption that underlies the tactic of defending a particular theory of justification through showing the weaknesses of its competitors. If the Agrippa problem has no solution, such dialectical arguments have no force. Nothing short of exhibiting the solution will count as an answer to the Agrippa problem—which is what the second success condition demands.

These three demands on an adequate theory of epistemic justification strike me as being altogether reasonable, but if we examine the philosophical exchanges between competing epistemologists, we discover that they are rarely insisted upon. There exists what might be called "the Epistemologists' Agreement" not to hold each other to such standards, perhaps because it is tacitly understood that no theory can meet them.

The thesis of Part II of this study is that, in fact, no theory of justification has satisfied these three success conditions. More strongly, I will argue that no theory of justification has come close to doing so. It is not possible to deal in detail with all theories of justification that have been presented—there has been rather an explosion of them in recent years—but an examination of what I take to be the strongest representative samples of the main types of theories of justification may create at least a strong presumption in favor of the negative thesis that this part of this work is intended to further.

Theories of Epistemic Justification

There is no completely natural way to classify theories of epistemic justification, since they can vary in several dimensions. In the literature, we often find a broad contrast drawn between theories that are internalist and those that are externalist. Unfortunately, this contrast can be taken in a number of ways, and it is not always clear what contrast an epistemologist has in mind when she describes (or accuses) another philosopher of being either an internalist or an externalist.

The words "internalism" and "externalism" most naturally suggest two domains of objects: those that are in the mind and those

that are not in the mind. An epistemologist who is an internalist in this sense accepts something like the following thesis:

For S to be justified in believing that p, the grounds that justify this belief must be contents of S's mind.

The usual motivation behind this version of internalism is that only mental contents can provide the immediately accessible evidence needed to provide a secure basis for knowledge. Philosophers who call themselves externalists typically reject this principle. This happens, for example, when externalists contrast their position with Cartesianism.[7] I will say that a position that is internalist in this way is an instance of *ontological* internalism.

Sometimes, however, epistemologists seem to have something altogether different in mind when they draw a distinction between internal and external theories of epistemic justification. They are not concerned with the ontological status of the grounds—mental as opposed to (possibly) nonmental grounds; instead, they are concerned with the relationship between the person who has a justified true belief and the grounds that justify it. An internalist in this sense accepts a principle of the following kind:

For S to be justified in believing that p, S must *base* his belief on the grounds that justify it.

The demand that the justified believer base her belief on grounds that justify it is characteristic of theories that make a strong demand for epistemic responsibility. I will say that theories that embody a principle of this kind exemplify *methodological* internalism. Many philosophers who call themselves externalists reject this form of internalism as well. Here externalists often contrast their positions with those that they say are overly intellectualist.

In fact, even though nothing necessitates this, many philosophers have been internalists in both ways and some have been externalists in both ways. It is easy to see how this can happen. If one is a methodological internalist, insisting that a justified believer must *base* her belief on grounds that justify it, and then holds, for whatever reason, that the only thing accessible to a person is her own ideas (or other mental contents), then it is easy to see how an epistemologist might become both an ontological and a methodological internalist. Epistemologists sometimes reject both forms of internalism and thus become externalists in the two ways I have noted; that is, they hold that it is possible for someone to know something on

grounds that are not mental in character and that are not employed by this person as the basis for the belief.

Though it is easy to understand how these two sorts of internalism and the two sorts of externalism are naturally combined, I think it is important to keep these distinct themes separate. First, they are not always combined. In the next chapter I will show that Chisholm seems to develop a theory of empirical justification that combines ontological internalism with methodological externalism. Second, and more importantly, ontological internalism and methodological internalism generate separate (though admittedly related) philosophical problems. Ontological internalism generates what we might call Cartesian skeptical problems. Put crudely, the fundamental task of Cartesian epistemology is to find some way to establish justified belief beyond the mental contents that serve as the privileged epistemic starting point. This problem is typically posed in the form of skeptical scenarios, for example, Descartes' evil-spirit hypothesis and recent innovations involving brains in vats.

Methodological internalism—the doctrine that to be justified in believing something, a person must base her belief on the grounds that justify it—generates a different kind of skeptical challenge, namely, the Agrippa problem. A theory of epistemic justification, as I understand it, is an attempt to solve the Agrippa problem. This is not the same problem as the Cartesian skeptical problem, though the two are often linked. One common way of trying to solve the Agrippa problem is to take as a starting place (supposedly) incorrigible beliefs concerning immediate experience. The task, then, is to show how these beliefs concerning immediate experience can be used to justify claims concerning the "external world." Thus attempts to answer the Agrippa problem can generate the Cartesian problem, and then fail because of an inability to solve it.

Since the distinction between internalism and externalism can mark at least two different contrasts, I will not classify theories of justification in terms of it. Instead, I will classify theories of epistemic justification in terms of their response to the Agrippa problem. We can simply note which of the three dialectical modes—infinite regress, hypothesis, or circularity—a particular theory attempts to neutralize, then exploit. Since we are concerned with human knowledge, it is hard to see how the mode of infinite regress can be made innocent of skeptical consequences. This leaves the modes of circularity and hypothesis as the remaining options. Coherentists embrace the first option, foundationalists the second. In the next

chapter I will examine foundationalism; in chapter 8, I will examine a traditional version of coherentism. Chapter 9 examines an original theory recently developed by Donald Davidson that, remarkably enough, combines coherentism with (ontological) externalism. Chapter 10 concludes this work with some general reflections on Pyrrhonism.

Notes

1. Clifford, 1879.
2. Moser, 1985. All page references are to this source.
3. BonJour, 1985. All page references are to this source.
4. Pollock, 1986. Some of the force is taken out of this argument in a note appended to this passage, where Pollock tells us, "The claim I am making here is a contingent one about those skeptical arguments that have actually been advanced in philosophy" (6n).
5. Unlike his discussion of the Ten Modes attributed to Aenesidemus, where Sextus sometimes uses an individual mode in isolation to derive skeptical consequences, with the Five Modes, he explicitly uses them in concert in the way I have described. (See *PH*, 1: 169–74.)
6. Another modern restatement of the Agrippa problem—though again without reference to its ancient source—is found in Chisholm, 1982b, 129.
7. See, for example, Alvin Goldman's "Causal Theory of Knowing," discussed in chapter 3.

7

Foundationalism

Foundationalist theories of justification attempt to solve the Agrippa problem by finding some way of bringing the infinite regress of reasons to a nonarbitrary halt. Beliefs that serve this function can be called foundational beliefs. If we set aside attacks on competing theories, the foundationalist program will fall into two parts: (a) the set of foundational beliefs must be identified and their foundational status justified, and (b) given this specification of foundational beliefs, a further construction is needed to show how they can serve to justify a suitably wide range of other beliefs.

Forms of Foundationalism

Foundationalism comes in various forms. First, the theories can vary concerning the epistemic status of the foundational beliefs themselves. On some accounts, these beliefs satisfy the highest epistemic standards in being infallible, incorrigible, indubitable, certain, and so forth. Until relatively recently, all foundationalist theories took this form. At the other extreme, we can imagine the foundational beliefs possessing, perhaps, only a low degree of warrant, that is, the foundational beliefs *by their nature* carry some presumption of truth, but a presumption too weak to stand on its own. Finally, we can imagine a position somewhere between these two, where the foundational beliefs fall short of infallibility, incorrigibility, and so on, yet possess a sufficient degree of warrant to be justified unless overridden.

 A second way foundationalist theories can differ concerns the manner in which the foundational beliefs get their warrant, or at least their initial warrant. Being foundational, such beliefs cannot

get their initial warrant from another belief. This leaves two options: either the belief must be *self*-justifying, or the belief must be *immediately* justified by something that is not a belief. Pollock calls this the difference between a doxastic and a nondoxastic version of foundationalism.[1] In the traditional nondoxastic theories, certain items in consciousness that are not themselves beliefs, for example, sense-data, are what give basic beliefs their warrant.[2] In opposition to these *nondoxastic* theories, some philosophers have maintained that foundational beliefs, while not justified by other beliefs, are not justified by anything that is not a belief. Thus, to the extent that such beliefs are justified, they must be self-justified.

Foundational theories can vary, then, with regard to the degree of warrant assigned to the foundational beliefs, and they can vary in being either doxastic or nondoxastic. They can vary in other ways as well. Thus foundationalism is not a single theory, but a complex family of theories held loosely together by a common strategy for dealing with the Agrippa problem. Because of this complexity, a complete examination of the foundationalist position—one responsive to the extensive literature in this area—would be an enormous task. It is not a task I will undertake here. Instead, I will concentrate on a single version of foundationalism: that developed by Roderick Chisholm. My basic claim is that Chisholm's theory does not satisfy—or even come close to satisfying—the success conditions for a theory of justification laid down in chapter 6. Showing this with respect to his theory will, I hope, provide a pattern for evaluating all versions of foundationalism, for each of them, I think, shares at least one of the basic flaws found in Chisholm's theory.

Chisholm's Version of Foundationalism

Roderick Chisholm has, to my mind, developed the foundationalist theory of justification in more detail, with more precision, and with a deeper systematic understanding of the issues involved than any other defender of this general position. At the same time, he has produced a theory of such technical intricacy that the reader lacking Providential guidance sometimes feels like Herr K striving to reach the castle, occasionally catching glimpses of it, but always being shunted into side streets. Beyond this, Chisholm often introduces principles and definitions (and subtle qualifications of such principles and definitions) without explaining their underlying motivation. Chisholm's philosophy is played close to the vest. Furthermore,

he often assigns ordinary words technical senses that depart in significant ways from their ordinary senses, and then uses these terms to define other terms. For all these reasons, Chisholm is often hard to understand and easy to get wrong.

I will draw the statement of Chisholm's position mainly from the third edition of his *Theory of Knowledge* (*TK3*),[3] occasionally supplementing it with citations from other sources where he is more expansive in his explanations.

Levels of Justification

For Chisholm, knowledge is a species of justified true belief, but not every species of justified true belief counts as knowledge. For justified true belief to count as knowledge, the justification must be of the right kind or level. This thought leads Chisholm to introduce an extraordinarily complex hierarchy of terms of epistemic assessment. Chisholm took this aspect of his theory seriously, and over the years revised it numerous times, sometimes in fundamental ways. Since his theory is expressed in terms of these levels of justification, we must examine them in at least some detail.

In *TK3* Chisholm distinguishes thirteen levels of epistemic evaluation ranging from the *certain* at the top of the hierarchy to the *certainly false* at the bottom. To give this hierarchy a systematic structure, he defines each epistemic concept in this hierarchy using the single relational concept: "So-and-so is *at least as justified* for *S* as such-and-such." For example, the counterbalanced, which lies at the neutral midpoint of the epistemic hierarchy, is defined as follows:

> *p* is counterbalanced for *S* =Df *S* is at least as justified in believing *p* as in believing the negation of *p*; and *S* is at least as justified in believing the negation of *p* as in believing *p*. (9)

That sounds right.

Some of Chisholm's definitions, though intelligible, involve departures from our ordinary ways of speaking. Here is his definition of the probable:

> *p* is probable for *S* =Df *S* is more justified in believing that *p* than in believing the negation of *p*. (10)

First, "probable for" is not idiomatic English, but setting that aside, generally when we say that something is probable we are indicating that it has a fairly good chance of being true. Chisholm's definition

carries no such implication. This is something we must keep in mind when we encounter this term, and also try to remember when this concept is embedded in the definition of other terms.

Chisholm's definition of beyond reasonable doubt takes the following form:

p is beyond reasonable doubt for S =Df S is more justified in believing p than in withholding p. (11)

Earlier Chisholm explained the notion of withholding in these words: "A person may be said to *withhold* a proposition h provided he does not believe h and does not believe the negation of h. The proposition that God exists is such that the theist accepts it, the atheist accepts its negation, and the agnostic withholds it" (8). In other words, if it is more reasonable to accept a proposition than to remain on the fence concerning it, then that proposition is beyond *reasonable* doubt. That too sounds right.

Things become problematic when we turn to Chisholm's definition of the *evident*. This definition demands close examination for, according to Chisholm, "the evident is that which, when added to true belief, yields knowledge" (11). His definition of the evident reads as follows:

p is evident for S =Df For every proposition q, believing p is at least as justified for S as is withholding q.

Chisholm illustrates the force of this definition using the following example: "If it is now evident to you that the sun is shining, then, given this definition, we may say that you are at least as justified in believing that the sun is shining as you are in withholding any contradiction or in withholding what is epistemically impossible (say, the proposition you would express by saying 'I am not thinking')" (11–12). I confess that I find Chisholm's definition of the evident and his illustration of it totally baffling. Earlier we were told that a person withholds a proposition h if "he does not believe h and does not believe the negation of h." Under what conditions will someone be *justified* in withholding h? A natural suggestion is that someone is justified in withholding h if, for that person, neither h nor its negation is beyond reasonable doubt. That, however, cannot be what Chisholm has in mind, for it has the consequence that no one is ever justified in withholding a contradiction, since the denial of a contradiction is beyond reasonable doubt. What, then, does Chisholm have in mind when he speaks of someone being justified in withholding h, where h is recognized to be necessarily false?[4] Chisholm never

explains how this is possible, and it is not obvious to me how this could be done.

It seems, then, that Chisholm's definition of the evident lacks the transparency we have a right to expect, and, if that is correct, we cannot reach the question whether the evident, so defined, provides the right kind of justification for knowledge. Though this strikes me as an extraordinary shortcoming in Chisholm's exposition of his position, this is not a point I wish to stress, for it casts criticisms at the wrong level. If his definition of the evident is inadequate in the way that I suggest, then perhaps it can be suitably amended, or perhaps replaced by a definition of another kind.[5] The correct characterization of the evident may be an important issue in the debate among foundationalists, but the fundamental demand of Pyrrhonism is that philosophers show how the elevated standards they *impose on themselves* can be satisfied. Setting aside technical details, we know broadly what Chisholm is trying to do: He is trying to provide a foundation for knowledge by showing that some of our knowledge claims can meet suitably high standards of justification. I do not think we need to challenge points of technical detail to show that he has not done this.

Before we proceed, however, it is important to note three features of Chisholm's notion of the evident. First, for Chisholm, a proposition can be evident to someone and yet be false. Chisholm is quite explicit on this matter: "We have noted that a proposition may be beyond reasonable doubt and also false. We will find that the same is true of the evident. It is possible that there are some propositions that are both *evident* and false." The reason he gives for this claim is instructive: "If we do in fact know some of those ordinary things that we think we know (for example, that there are such and such pieces of furniture in the room, that the sun was shining yesterday, that the earth has existed for many years past) then we must reconcile ourselves to the possibility that on occasion some of those things that are evident to us are also false" (12). It is then an important feature of Chisholm's position that it does not contain, indeed explicitly rejects, what I have labeled the (iii$_g$), or adequate-grounds, clause. In line with this, we might reformulate Chisholm's analysis of knowledge in these words: "Knowledge is evident belief that also happens to be true." For Chisholm, like Wittgenstein, it is only by the grace of nature that we know anything, for, on Chisholm's analysis, believing only that which is at least evident is all that knowledge demands of *us*, and that by itself is not enough to guarantee truth. In this one respect, at least, the neo-Pyrrhonian approach sketched in

chapter 5 and Chisholm's approach are in agreement. There is, how-
ever, little agreement beyond this single point.

A second feature of Chisholm's description of the evident, at
least taken at face value, is that something can be evident for *S* even
if *S* does not recognize that it is evident for him. This comes out in
the following curious passage:

> If a person *S* is *internally justified* in believing a certain thing, then
> this may be something he can know just by reflecting upon his own
> state of mind. And if *S* is thus internally justified in believing a cer-
> tain thing, can he also know, just by reflecting upon his state of
> mind, that he *is* justified in believing that thing? This, too, is possi-
> ble—once he has acquired the concept of epistemic justification. (7)

In this passage Chisholm speaks about what a person *S* *may* be able
to do, but there is no indication here, or elsewhere, that, to be evi-
dent for *S*, *S* must *base* her belief upon the grounds that make it evi-
dent for her. Thus, as stated, Chisholm's account of knowledge con-
tains nothing corresponding to my (iii$_p$), or performance, clause.
Read off the page, Chisholm's account of empirical justification is
ontologically internalist, but methodologically externalist. On the
other hand, as we shall see later in this chapter, Chisholm finds it
important to insist that "one can find out directly, by *reflection,*
what one is justified in believing" (7). Perhaps his theory comes to
this: In order to be justified in believing something, one must *be able*
to find out by reflection that one is justified, but this need not be
something that one *has actually* done. If that is what Chisholm has
in mind, he has, we might say, replaced a performance clause with a
capacity clause. Even if he has done this, Chisholm remains, in a
way, Cliffordian, since the expected capacity is very demanding.[6]

A third feature of Chisholm's notion of the evident is this:
Though the evident introduces a high standard of epistemic justifi-
cation, it is not at the top of Chisholm's epistemic hierarchy. That
honor goes to the *certain,* which Chisholm defines as follows:

> *p* is certain for *S* =Df For every *q,* believing *p* is more justified for *S*
> than withholding *q,* and believing *p* is at least as justified for *S* as is
> believing *q.* (12)

With certainty we can say: "It doesn't get any better than this."[7]

With the notion of the certain in hand, we can sketch Chis-
holm's overall strategy. It falls into two parts. He will first argue for
the existence of beliefs that are certain. He will then argue that these
beliefs that are certain can make other beliefs that are not certain at

least evident. If the evident is, as he says, "that which when added to true belief, yields knowledge," then completing this double task will, at least on his own terms, constitute a constructive response to skepticism.

Certainty and the Self-Presenting

Chisholm's candidate for beliefs that are certain are those that arise from *self-presenting* properties. Since the account in *TK3* is compact and austere even by Chisholmian standards, it will be useful to turn to an earlier source where his explanation is somewhat more forth-coming. In "A Version of Foundationalism,"[8] he describes a self-pre-senting property in the following way: "Every self-presenting prop-erty . . . is a property that is such that, if while having it, you consider your having it, then you will *believe* yourself to have it" (11). Chisholm offers an interestingly varied list of properties that he claims can be self-presenting for the subject who has them: "One example is feeling sad; another is thinking about a golden mountain; another is believing oneself to be wise; and still another may be sug-gested by the awkward locution, 'is appeared redly to'" (10).

With this as a background, it is possible, I think, to see what Chisholm is getting at in his remarkably obscure definition of the self-presenting in *TK3:*

> P is self-presenting =Df Every property that P entails includes the property of thinking. (19)

Here the terms "entails" and "includes" are used in special techni-cal ways, but, tutored by his earlier definition and examples, I think we can get the gist of what Chisholm is saying without going into these technical details. Roughly, for Chisholm, P is self-presenting if it is fully given in thought, that is, if there is not anything about it that is not immediately accessible to thought. (Those things that are self-presenting, we might say, have no crevices, just as sense-data were thought to have no reverse sides.)

For Chisholm, self-presenting properties are "a source of cer-tainty" (*TK3*, 19):

> If the property of being F is self-presenting, if S is F, and if S believes himself to be F, then it is certain for S that he is F. (19)

Chisholm cites sadness as one (possible) example of a self-presenting property:

If *S* feels sad, then *S* is at least as justified in believing that he feels sad as he is in having any other belief [i.e., it is certain for *S* that he is sad]. (19)

Before turning to the self-presenting property suggested by Chisholm's awkward locution, "ways of being appeared to," we might ask whether it is true that if a person feels sad and believes she feels sad, must it be certain for her that she is sad? We can ask a parallel question concerning a person's belief that she is wise: If a person believes that she is wise and believes she believes this, must it be certain for her that she believes that she is wise? Will these beliefs be as justified for her as her belief, for example, in her own existence?

The answer to these questions is not uncontroversial. We can imagine a philosopher challenging Chisholm in the following way:

When I attribute sadness to myself, I am bringing this particular feeling under a general concept and thus implicitly *comparing* this feeling with other feelings. There is no reason why I cannot be mistaken in making such a comparison between my present experience and something outside of my experience. So, even if I am right in this particular case in attributing sadness to myself, it need not be certain for me that I feel sad.

It seems that a parallel point can be made with respect to *being appeared redly to*,[9] or, more idiomatically, with respect to something appearing red to us. Isn't it possible for me to misclassify a sensation, taking it to be red when it is not really like other things I consider red? If this is possible, doesn't my present belief that I am being appeared redly to, even if true, fall short of certainty?

Chisholm's response is that this loss of certainty is not possible in the noncomparative sense in which he understands appeared-to statements. In their noncomparative sense, "'appear'-words," he tells us, "[are] used to describe . . . ways of appearing" (21). They describe what might be called the "phenomenological" content of an experience, that is, that which remains when all reference to that which is external is bracketed out.

To arrive at what is self-presenting . . . we must remove the reference to the external things—to wine in "The wine tastes sour to me," and to the appearing thing in "That thing appears red to me." This, however, is very difficult to do, since our language was not developed for any such philosophical purpose. (22)

Removing references to external objects, if this can be done, would eliminate one source of possible error, but this maneuver seems

unresponsive to the possibility of misclassification. Chisholm, recognizing this, states the criticism this way:

> (a) In saying "something appears white," you are making certain assumptions about language; you are assuming, for example, that the word "white," or the phrase "appears white," is being used in the way in which you have used it on other occasions, or in the way in which other people have used it. Therefore (b), when you say "This appears white," you are saying something, not only about your present experience, but about all of those other occasions. But what you are saying about those other occasions is not certain. And therefore (d), "This is [appears][10] white," does not express what is certain. (24)

Chisholm responds as follows:

> The false step in this argument is the inference from (a) to (b). We must distinguish the belief that a speaker has about the *words* he is using from the belief that he is using those words to express. What holds true for the former need not hold true for the latter. A Frenchman, believing that "potatoes" is English for apples, may use "There are potatoes in the basket" to express the belief that there are apples in the basket; from the fact that he has a mistaken belief about "potatoes" and "apples," it does not follow that he has a mistaken belief about potatoes and apples. Similarly, it may be that what a man believes about his own use of the expression "appears white" is something that is not certain for him—indeed what he believes about his own language may even be false and unreasonable; but from these facts it does not follow that what he intends to assert when he utters "This appears white to me" is something that cannot be certain. (24–25)

We can first note that the comparison with the Frenchman is not responsive to the point at issue. We can imagine the Frenchman going wrong in two ways. He not only mistakenly uses the word "potato" where he should use the word "apple," he may sometimes mistakenly identify something as an apple (using the expression "That's a potato") when, in fact, it is a pear. What we need is a guarantee that this second sort of mistake is not possible with respect to reports of being appeared to. What we need, that is, is a guarantee that, with respect to appearance statements, we can never confound one sort of appearance with another. The price of achieving that, I believe, is unacceptably high, for it amounts to removing all *contrastive* content from appearance statements, for in the noncomparative sense, believing that something appears white to me, I am not grouping this appearance with other appearances of the same kind,

and I am not contrasting it with appearances of a different kind. The upshot of this is the reduction of "This appears white to me" to something like "This appears this way" (where I absorb myself in the present experience). Appearance statements have, it seems, become what we might call semantically atomic. If this is correct, Chisholm's defense of the noncomparative sense of appearance statements may have so thinned them of content that they will be unsuited to serve a foundational role. In particular, we must be told how the transition is made from knowledge claims involving non-comparative ascriptions to those involving the comparative ascriptions found in our everyday knowledge claims. When we examine Chisholm's constructive efforts to show how empirical knowledge has its foundation, we will see that he does not address the question. Furthermore, given the resources available to him, it will be hard to see how he could give a satisfactory answer to it.

Presumptions

It is sometimes breathtaking the way in which epistemologists beg the question against the skeptical position they are claiming to refute. For example, the fifth chapter of *TK3* begins with these words: "It is primarily by means of perception that we know about the external things around us. Our senses, somehow, provide us with evidence about the things that stimulate them. No theory of knowledge can be taken seriously that is not adequate to this fact" (39). My objection is not that this passage begs the question concerning the existence and knowability of an external world, for Chisholm will make an honest (though, it seems to me, unsuccessful) attempt to solve that problem. The passage is question-begging in its tacit assumption that we *are* in possession of secure perceptual knowledge. The philosophical question takes the following form: Given the nature of our senses (their fallibility, etc.), how is it (still) possible to have the perceptual knowledge we have?

The Pyrrhonist, or neo-Pyrrhonist, will have none of this. If asked how our senses provide us with knowledge of external things around us, he might point to various common ways of using our senses to find things out. A person can know that dinner is being prepared by smelling it cooking. This is how the word "know" is used in everyday life. If pressed hard enough whether he *really* knows that dinner is being prepared, he will simply retreat from the claim. The Pyrrhonist does not grant the existence of the kind of knowledge that the epistemologist attempts to vindicate.

But isn't this rather beside the point? If Chisholm can vindicate our ordinary knowledge claims, wouldn't this just show that the Pyrrhonist is fainthearted in retreating from them when challenged? It would show that, but, as we shall see, begging the question against Pyrrhonism is built into the very machinery of Chisholm's theory. This occurs through his use of the notion of *presumption,* which plays a central role in Chisholm's theory. Under various titles, it plays a central—and question-begging—role in a great many other contemporary theories as well.

Chisholm cites Carneades as one source of the view he accepts: "Carneades knows that we cannot make any inductive inferences about external things until we have some *perceptual data* about such things. And it follows from this that, if we are to have any positive justification for what we believe about the external world, our experience must provide us with a probability that is *not* derived from an induction" (46). This leads Chisholm to attribute the following doctrine to Carneades:

> Taking something to be F tends to make it probable that there is something that one is taking to be F. (47)

Chisholm finds a similar view in the writings of Meinong, who held, Chisholm tells us, that "when one takes there to be a tree then the judgment that there *is* a tree that one is perceiving may have presumptive evidence" (47).

Finally, Chisholm cites a passage from H. H. Price that will be particularly useful for our purposes: "The fact that a material thing is perceptually presented to the mind is *prima facie evidence* of the thing's existence and of its really having that sort of surface which it ostensibly has; . . . there is *some presumption in favour of* this, not merely in the sense that we do as a matter of fact presume it (which of course we do) but in the sense that we are entitled to do so" (48).[11] The passage I have underscored marks the basic epistemological assumption that the Pyrrhonist will not grant.

Why accept this dogma that we are *entitled* to a presumption in favor of our senses? It is not needed for an adequate *description* of our use of epistemic terms: The mere fact that we do (for the most part) presume that our senses are reliable is all we need for that. We need the notion of entitlement only if we are trying to find a general *vindication* of our reliance on our senses. That, of course, is the very point at issue in trying to refute Pyrrhonism. What we need, then, is some *independent* reason for accepting the claim that we are entitled to the presumption that our senses are reliable. Without this

independent argument, the response to Pyrrhonism simply fails. I will return to this point at the close of this chapter.

The Transfer of Justification

At this stage the following broad picture of Chisholm's program has emerged. At the foundation of the edifice of knowledge, there is a set of beliefs that are certain, generated by self-presenting properties. The following, then, is Chisholm's first principle of epistemic justification.

MP1 If the property of being F is self-presenting, if S is F, and if S believes himself to be F, then it is certain for S that he is F.

Chisholm has given a number of examples of self-presenting properties. Some involve intentional attitudes; others involve ways of being appeared to. Chisholm would hold that both of these instantiations of MP1 are true:

If S is thinking about a bicycle, and believes that she is thinking about a bicycle, then it is certain for S that she is thinking about a bicycle.

If S is being appeared to bluely, and believes that she is being appeared to bluely, then it is certain for S that she is being appeared to bluely.

The first example illustrates a self-presenting property that is intentional; the second a self-presenting property that is sensible.

It is important to see that Chisholm's foundationalist program relies on both kinds of self-presenting properties. He needs intentional self-presenting properties to support his belief that a person can determine, *by reflection alone*, whether a belief is justified for her or not. Chisholm is quite explicit on this matter: "One can find out directly, by *reflection*, what one is justified in believing at any given time" (7). Beyond this, Chisholm also credits human beings with the capacity to rank their beliefs—any two of their beliefs—correctly relative to their level of epistemic justification. If Chisholm holds, as he seems to hold, that possessing this ability is a necessary condition for being justified in one's belief, then this is a strong commitment indeed. It is also a commitment that seems, on its face, to be empirically implausible.[12] If this is correct, then Chisholm is developing a theory for idealized cognizers, not for human beings. Given a theory of epistemic justification, we should always ask: Does this theory demand cognitive powers that human beings do not

possess? If the answer to that question is yes, then the theory leads to skepticism rather than avoiding it.

Chisholm needs *sensible* self-presenting properties for his position to count as a theory of *empirical* justification. These sensible properties perform two services: They supply the empirical content for empirical knowledge, and they supply the certainty Chisholm believes must undergird all other levels of empirical justification. As we examine the way in which Chisholm's theory unfolds, it will be important to keep in mind that these sensible self-presenting properties provide the *sole* content for empirical judgments and, further, this content is the thinned-out content of noncomparative attribution.

Broadly speaking, for Chisholm, the transfer of justification takes place in the following way. We start with the certainty of the self-presenting. If I am appeared to redly (and believe this of myself), then it is certain for me that I am appeared to that way. Here the level of epistemic justification is unsurpassed, but the content of the belief is subjective and thin. Given thin certainties of this kind, the task is to show that other, more robust judgments (for example, that there are two chairs in this room) are evident for me, for if such judgments are evident for me (and also true) then, according to Chisholm, they are known to me. The task, we might say, is to trade high epistemic quality for richer informational quantity, keeping the quality high enough to count as knowledge, while getting the content rich enough to count as something worth knowing.

Material Epistemic Principles

In the first half of this century, philosophers often posed the foundational problem as some variation of the following question: How are sense-data statements related to material-object statements, such that they are capable of giving material-object statements adequate evidential support? Attempts to answer this question involved giving an analysis (or partial analysis or explication) of material-object statements, an enterprise that proved wholly unsuccessful. Chisholm offers an attractive alternative to this "analytic" foundationalism, one that does not depend on an antecedent analysis of material-object statements. Exploiting an analogy with ethical principles, Chisholm introduces what he calls *material* epistemic principles, which he describes as follows:

> A *material* epistemic principle is a principle relating non-epistemic concepts to epistemic concepts. An example is, "If a person believes

himself to be talking with someone, then it is certain for that per-
son that he believes himself to be talking with someone." The
antecedent of the principle tells us that the applicability of a certain
non-normative concept (believing oneself to be talking with some-
one) constitutes a sufficient *logical* [emphasis added] condition for
the application of a certain normative concept (being certain that
one believes oneself to be talking with someone). (61–62)

Chisholm offers *ten* such material epistemic principles. His
account of the foundations of empirical knowledge is an attempt to
show how, through progressive stages, the certainty of the self-
presenting can be used to make evident ordinary claims about the
world around us. I'll simply follow this development, adding brief
explanatory comments along the way.

MP1 If the property of being F is self-presenting, if S is F, and if
S believes himself to be F, then it is certain for S that he is F.

MP2 Accepting h tends to make h probable.
(The important point here is that MP2 does not say that accept-
ing h *makes* h probable; it asserts the weaker claim that accept-
ing h *tends to* make h probable. Because of this contrast, Chis-
holm says that MP2 embodies "epistemic commonsensism"
rather than "epistemic conservatism" (63). For our purposes, it
is more important to note that the expression "tends to make h
probable" introduces the notion of presumption already chal-
lenged.)

MP3 If S accepts h and if h is not disconfirmed by S's total evi-
dence, then h is probable for S.
(Here the word "tends" is dropped, thus replacing MP2 with an
absolute claim to probability. This strengthening is justified by
the stronger demand that h is not disconfirmed by S's total evi-
dence. This is an important development, for now the epistemic
status of a given proposition is made a function [in part] of the
epistemic status of other propositions. This standard that takes
us from what tends to be probable to what is probable is some-
times called the negative coherence criterion: Roughly, if some-
thing is prima facie [or presumptively] justified, and does not
conflict with other beliefs that are justified, then that belief is
justified as well. MP3 is an instance of such a principle for a par-
ticular level of justification.)

MP4 If S accepts h and if not-h is not probable in relation to the
set of propositions that are probable for S, then h is epistemi-
cally in the clear for S.

(This is essentially another application of a negative coherence criterion, now using the set of propositions shown to be probable through satisfying MP3.)

MP5 If *S takes* there to be an *F*, and if it is epistemically in the clear for him that there is an *F* which he takes to be *F*, then it is *beyond reasonable doubt* for *S* that he is perceiving something to be *F*.

(This principle contains the strong epistemic operator *perceiving*—for Chisholm, perceiving something to be the case implies that it is the case [40–41]. This operator appears, however, under the scope of the weaker epistemic operator *beyond reasonable doubt*. The meaning of such nested operators is not always easy to grasp. Part of Chisholm's task is to get the strong epistemic operator out from under the scope of this weaker one.)

MP6 If *S seems to remember* having been *F*, and if it is epistemically in the clear for him that he remembers having been *F*, then it is *beyond reasonable doubt* for *S* that he remembers having been *F*.

(This principle concerning memory is, as Chisholm notes, analogous to MP5, which concerns perception.)

MP7 If there is a set of concurrent propositions such that all of the propositions are epistemically in the clear for *S* and one of them is beyond reasonable doubt for *S*, then all of them are beyond reasonable doubt for *S*.

(The introduction of the concept of *a set of concurrent propositions* marks an important departure. For Chisholm, concurrence involves the notion of *mutual support*, which is stronger than the notion of *mutual compatibility*. We have thus passed from the use of a negative coherence standard, first introduced at MP3, to the use of a positive coherence standard. The claim in MP7 is quite subtle. If we have a set of concurrent propositions, all of which are in the clear, then if one of them is beyond reasonable doubt this guarantees that the other members of the set are at least beyond reasonable doubt. This is a strong principle that needs a strong concept of mutual support to render it plausible. It is far from clear to what extent human beliefs can support each other in this strong way.)

MP8 If *being appeared ____ to* is evident for *S*, and if it is epistemically in the clear for *S* that there is something that appears ____ to him, then it is *evident* for *S* that there is something that is appearing ____ to him.

(As Chisholm tells us, "This principle tells us of conditions under which *being appeared to* makes it evident that *something is appearing.*" Metaphorically, it tells us the conditions under which it is evident that something is *out there*, but, so far at least, nothing need be evident concerning what this thing "out there" is like. This principle is intended to give us no more than a referential toehold on what used to be called external reality.)

MP9 If *S* takes there to be an *F* and if it is beyond reasonable doubt for *S* that he is perceiving something to be *F*, then it is evident for *S* that he is perceiving something to be *F*.

(This is complicated, and Chisholm does not explain the complications. For Chisholm, perceiving is itself a strong epistemic operator. If someone perceives that there is a cat on the roof, then, on Chisholm's use of the term "perceive," there is a cat on the roof and it is evident to that person that there is. Perceiving is perceptual knowing. Thus MP9, if accepted, takes us, at last, over the divide into the region of propositional knowledge of the world around us.)

MP10 If there is a set of concurrent propositions such that all of them are beyond reasonable doubt for *S* and one of them is evident for *S*, then all of them are evident for *S*.

(This principle is analogous to MP7. In both cases, a member of a set of concurrent propositions with an elevated epistemic status guarantees that the other members are at least at this same level as well. Playing his cards unusually close to his vest, Chisholm provides no commentary on this principle. Its point, I take it, is to extend empirical knowledge beyond that which is currently perceived or recalled.)

Summary and Evaluation

In the previous chapter I suggested that any philosopher engaged in a justificationalist project, be it foundationalist or antifoundationalist, must meet a number of challenges: (1) The position should be presented with what I have called philosophical candor, that is, we have a right to be told the range of beliefs that the theory takes as justified. (2) We have a right to be told, in some detail, how this justification is secured. Without the first constraint the theory, for all we know, may be the equivalent of a strong version of skepticism. Without the second, we will have no reason to suppose that skepti-

cism has been refuted. Finally, (3) the argument in behalf of the position must not beg any questions against Pyrrhonian skepticism.

1. In "A Version of Foundationalism" (Chisholm, 1982b), Chisholm begins by noting some things that we know and some things that we do not know: "There are certain things I know and certain things I do not know. I can give examples of each. Like Moore, I know that I have two hands and that the earth has existed for hundreds of years. But I do not know whether it will rain here a year from today and I do not know how many people now live in East Jaffrey" (3). This seems fairly bold, for if we can know, for example, that the earth has existed for hundreds of years, then, presumably, we can know many other things of this kind. We might be able to know, for example, that dinosaurs once roamed the earth. Unfortunately, this is the last we hear of such ambitious claims to knowledge in "A Version of Foundationalism."

Chisholm seems to be much more restrictive in *TK3* concerning what we can know. In particular, he seems to deny that we have *any* knowledge of future contingencies. This seems to be the plain implication of the following passage: "You may have very good grounds for accepting the proposition that you will walk tomorrow and the day after that: the proposition may be strongly supported by induction. But it is not *evident* to you or to anyone else that you will walk tomorrow, for no one now *knows* that you will walk tomorrow. . . . [N]othing you can find out *today* will make it *evident* for you today that you will walk tomorrow" (17). Unless Chisholm has a calendar-sensitive theory of knowledge, the same point should hold about things that will happen later today. It seems that it cannot be evident for me now that it will still be raining, say, two minutes from now. Speaking in a common way, people do, of course, suppose that they know such things about the relatively near (and sometimes not-so-near) future. In denying this possibility, as he plainly seems to do, Chisholm seems to deprive us of much that counts as common knowledge. Furthermore, to the extent that knowledge claims about the present and past entail propositions about the future, Chisholm's position may be even less generous in granting us knowledge. There is a general tendency in theories of knowledge to drift toward skepticism without acknowledging that this is happening. Chisholm, in fact, does not provide even a broad sketch of what he takes to be the scope or range of empirical knowledge. Since we are not told what edifice the foundation is supposed to support, it is hard to decide whether the theory is more skeptical or antiskeptical in its tendencies. This

omission is a fundamental defect in Chisholm's position—one, I think, that is shared by virtually everyone currently presenting theories of empirical justification.

2. Taking Chisholm on his own terms, that is, confining our discussion to the perception of trees and pieces of furniture, we can turn to our second criterion of adequacy and ask if he has shown, in suitable detail, how such knowledge is possible. One cluster of problems centers on Chisholm's use of the notion of the noncomparative attributions of properties. Earlier I remarked that "we have to be told how the transition is made from knowledge claims involving noncomparative ascriptions to those that involve the comparative ascriptions found in our everyday knowledge claims." Having examined Chisholm's ascent to empirical knowledge, we have not found a whisper concerning how this might be done. If it cannot be done— and the prospects hardly look promising—then Chisholm has defended a very small edifice of empirical knowledge indeed. To see how small, consider the following dialogue.

A. I know that bird; it is a skylark.
B. Actually, it is a crested lark.
A. Whatever, at least I know that it is like that.

A's final remark illustrates what Chisholmian knowledge comes to. We certainly thought we were going to get more than this.

Chisholm's use of the notion of the noncomparative attribution of properties raises another problem that he seems not to have noticed. His program depends, in various places, on the notions of disconfirmation and concurrence. The first is employed in what amounts to a negative coherence test, the second in a positive coherence test. The difficulty is that it is hard to see how propositions expressing noncomparative predications can enter into either of these relations with one another. If the transition could be made from the noncomparative to the comparative attribution of properties, then something might be done about this problem, but, as noted, Chisholm has not addressed this issue.[13]

On a related, but different, matter, Chisholm's combination of coherentist ideas with a conception of basic beliefs having very thin content leads directly to one of the standard problems for coherence theories. Coherentists are often challenged to show why two equally coherent, but mutually incompatible, systems of belief could not exist. The supposed inability of coherentists to solve this problem is thought to show that at least some beliefs must have at least some degree of epistemic justification that is independent of the epistemic

justification bestowed on them through relationships with other beliefs. I will discuss this problem with respect to coherence theories in the next chapter; here I think it is sufficient to note that Chisholm's version of foundationalism does not seem to provide any solution to it. Nothing in his theory excludes the possibility that two mutually exclusive but equally coherent systems of belief might have as their foundational basis the very same set of attributions concerning self-presenting properties. In borrowing from the coherentists, Chisholm has borrowed trouble.

Finally, I have one general criticism: I find many of Chisholm's material epistemic principles intuitively unpersuasive. In particular, I find nothing that motivates the acceptance of the epistemic material principles 8, 9, and 10. Chisholm, for his part, has not provided such a motivation, and, for me at least, these principles do not rest on their own bottoms.

3. The third success condition for a theory of epistemic justification is that it not beg the question against Pyrrhonism by making the argument depend on assuming its falsehood. It is remarkable how often epistemologists do this, quite explicitly, without a blush. The following specimen comes from Chisholm:

> There is the Aristotelian argument to the effect that some of the things I'm justified in believing are self-justifying. The argument is easier to ridicule than to refute. If my justification for accepting a certain proposition q requires me to go beyond and to appeal to a certain other proposition p, then I'm also justified in accepting q. Therefore these are the three possibilities: either there is an infinite regress; or there is a circle; or some of the propositions I'm justified in believing are self-justifying. But the first two of these three possibilities are inconsistent with the fact that I do know something. Therefore some propositions are self-justifying. (598)[14]

The underlying assumption of this passage is that beliefs that we take to be justified really are justified. The task is to show how. Pyrrhonists, as I have characterized them, can admit that certain things possess a relative justification, that is, they are justified relative to other things we take for granted. They can also admit that certain things often *strike* them as wholly justified. Furthermore, their feelings can be as strong as anyone's and in everyday speech they may express these feelings. Yet they hold—and Hume saw this, as did Wittgenstein—that this sense of justification can evaporate when we step back and examine our *actual* grounds for taking something to be justified. The justificationalists see this as well, for justificationalism is simply the attempt to find philosophical rejoinders

to the doubt that has been cast on various portions of our belief by philosophy itself. Indeed, it is tempting to view justificationalism (in both its foundationalist and nonfoundationalist forms) as an attempt to return to us, now with sound philosophical credentials, the original sense of epistemic security that philosophy itself has taken away.

The justificationalist cannot, without begging the question against the Pyrrhonist, argue in support of his position that it gives the best account—or even the only account—of how things that we take to be justified really are justified. In the sense in which *philosophers* are seeking justification, there is no antecedent reason to suppose, and no right to assume, that anything is justified. Yet Chisholm and many others constantly make this assumption.

Furthermore, Chisholm not only makes this assumption in a broad, programmatic way, as in the passage last cited; he also builds it into the technical machinery of his theory. This is immediately apparent if we examine Chisholm's material epistemic principles. Setting aside problems concerning noncomparative attributions, the Pyrrhonist might grant the first principle.

MP1 If the property of being F is self-presenting, if S is F, and if S believes himself to be F, then it is certain for S that he is F.

The traditional Pyrrhonists did not reject the possibility that we can be certain about how things appear to us—though they probably should have. They would, however, challenge Chisholm's next epistemic principle and the eight that follow it. To Chisholm's second epistemic principle,

MP2 Accepting h tends to make h probable,

they would oppose another,

MP2* Accepting h strikes us as making h probable.

Setting aside other difficulties, they would suggest a similar alternative to Chisholm's remaining eight epistemic principles. It should be clear that Chisholm's acceptance of principles MP2 through MP10 without giving good reasons for preferring them to their Pyrrhonian alternatives begs the question against Pyrrhonism. It begs the question against Pyrrhonism in the same way that a moral philosopher would beg the question against ethical skepticism by assuming the existence of prima facie obligations or prima facie goods. Chisholm, in fact, explicitly relies on this comparison, not realizing the problems it raises:

There is a fundamental analogy, then, between the logic of moral requirement and that of confirmation. In setting forth the principles of moral requirement, we have used "*prima facie* duty" and have contrasted *prima facie* duties with what Ross called "*absolute duties.*" In setting forth principles of confirmation, we could have used an analogous terminology, contrasting "*prima facie* probability" with what, following Bolzano, we could call "absolute probability." (*TK3*, 58)

In what we might call an internal criticism of a position relying on the distinction between the prima facie and the absolute, we might challenge the attempt to make the transition from the first to the second. That is, granting the existence of prima facie justification, we might argue that the transition to absolute justification has not been shown. Some of my criticisms of Chisholm's position are of this form, for example, my claim that Chisholm has not shown how comparative attributions can be made evident in virtue of the certainty of noncomparative attributions. A deeper criticism is to challenge Chisholm's right to make use of the notion of the prima facie probable at all. That, I have argued, already begs the question against Pyrrhonism. Chisholm is not alone in taking for granted the notion of the prima facie probable or some other form of the prima facie justified. This is a feature, expressed in various ways, of a great many contemporary theories of epistemic justification. *All* such theories beg the question against Pyrrhonism, and do so blatantly.

Finally, it should be conceded that many of the specific criticisms I have made against Chisholm's version of foundationalism do not apply to all versions of foundationalism, although some of the more general criticisms have wide application. In any case, this examination of Chisholm's foundationalism shows that the foundationalist program is much more difficult to promote when the Pyrrhonist is on hand and made party to the dispute. Whatever its merits relative to competing theories of justification, it seems to me that Chisholm's position fails utterly in meeting the Pyrrhonian challenge. I am not acquainted with any other version of foundationalism that does better.

Notes

1. Pollock, 1986, 19.
2. We could also imagine an ontologically externalist version of nondoxastic foundationalism where basic beliefs are justified in virtue of standing

in the right sort of relationship (perhaps a causal relationship) to something outside consciousness.

3. Chisholm, 1989. Unless otherwise indicated, all page references are to this source.

4. I assume that Chisholm has in mind recognized contradictions; if not, his account of the evident makes even less sense.

5. Chisholm himself defined the evident in an entirely different way in his earlier work "A Version of Foundationalism." Using the technical notion of a direct attribution, which I will not here explain, he there defined the evident as follows: "We may say that the direct attribution is *evident* for a given person provided that the attribution is beyond reasonable doubt and is one of those attributions on which it is reasonable for her to base her decisions" (Chisholm, 1982b, 9). I don't find this definition plausible either, for I cannot see how the level of justification (or reasonableness) of a belief is elevated by the reasonableness of acting on that belief. But I will not press this point either, not only because Chisholm abandoned this characterization of the evident, but because, again, it pitches the level of criticism too low.

6. Here I have been helped by Paul K. Moser. See, for example, Moser, 1985, 147, where he charges Chisholm with just this form of externalism.

7. Between the certain and the evident, Chisholm interpolates a level of epistemic appraisal he calls the *obvious*, which he defines as follows: "A proposition is said to be obvious for a subject S provided that, for every proposition q, S is more justified in believing p than in withholding q" (16). The reference to withholding makes this definition obscure in the same way that the definition of the evident is obscure, but from a purely formal point of view, the top three levels of the hierarchy can be represented as follows. Letting "Bp" mean "S's justification for believing" and "Wp" mean "S's justification for withholding," then, with the quantifiers understood, the top three levels of Chisholm's hierarchy have this form:

> Certain: $(Bp > Wq)$ & $(Bp \geq Bq)$
> Obvious: $Bp > Wq$
> Evident: $Bp \geq Wq$

Using this representation, it is clear that these "categories are such that each includes but is not included in the category listed immediately below it" (16). It is tempting to think that Chisholm introduced the notion of being justified in withholding a belief simply to generate this formal result, for this notion does not seem to have any independent intuitive force in specifying the proper level of justification for knowledge.

(With this formal structure before us, it is also tempting to ask whether there might be higher levels of epistemic justification beyond even the certain. The following is at least a formal possibility: There might be some proposition that we are more justified in believing than any other proposition. Such a proposition would be better known to us than any other proposition. Descartes held that his belief in God's existence had this status. But I digress.)

8. "A Version of Foundationalism" is found in Chisholm, 1982b.

9. The use of the expression "being appeared to" is a reflection of Chisholm's commitment to an adverbial rather than an objectual account of sensing. For Chisholm, to have a red sensation is to sense in a certain way, just as to dance a waltz is to dance in a certain way. In "A Version of Foundationalism" he puts the matter this way:

> The subject, who was said to experience a red sensation, does not stand in a sentient relation to an *object* that is a red sensation; rather [it] is sentient *in a certain way*—a way that we could describe as "redly." (Compare "she experiences sadness" and "she feels sad": the former suggests, misleadingly, that sadness is one of two things that are related by experiencing; the latter suggests, more accurately, that being sad is a *way* of experiencing.) (15–16)

Although Chisholm carries over the phrasing of the adverbial theory, he does not expound or defend it in *TK3*. I do not think there is any need to examine this aspect of his position here.

10. This crucial word somehow got lost.

11. Originally in Price, 1935, 185.

12. For arguments to this effect, see Stich and Nisbett, 1980; Stich, 1985; and Cherniak, 1986.

13. That all predications seem implicitly comparative is a bone in the craw of all versions of foundationalism. Moser, who does not accept Chisholm's notion of noncomparative predications, tries to avoid this problem by conceding that the acceptance of foundational beliefs presupposes an understanding of the words that express those beliefs. He puts it this way: "It seems clear that semantic information is necessary for believed propositions to be intelligible; and, therefore, we may hold that an understanding of certain semantic information is a necessary condition of a person's genuinely having any given-beliefs [beliefs concerning that which is given]" (Moser, 1985, 185). One of the standard objections to foundationalist programs is that, in order to make the transition from foundationalist beliefs to higher-level, nonfoundationalist beliefs, they must make use of nonfoundationalist beliefs that, supposedly, they are trying to vindicate. Moser's appeal to what he casually calls "semantic information" is a clear example of this error.

14. Chisholm, 1978.

8

Internal Coherentism

Under the heading of coherentism I will consider those theories that attempt to avoid the Agrippa problem by embracing something akin to circularity in a way intended to avoid rather than to generate skepticism. Traditionally, such theories have adopted two justificatory principles:

I. The only thing that justifies a belief is a belief.

II. The only thing that justifies a belief is another belief.

The first principle makes coherentism doxastic; the second separates it from doxastic versions of foundationalism (for example, of the kind that Chisholm adopts).

The second doxastic principle, far from solving the Agrippa problem, actually seems to generate it by triggering an infinite regress. To avoid this infinite regress, traditional coherentists typically reject what they call a linear conception of justification.[1] For a coherentist, justificatory arguments are not viewed as a sequence of steps, the last step being designated the conclusion, such that each step is either self-justified or justified by a rule of inference from previous steps. Since the coherentist does not allow self-justified beliefs, for him a linear conception of justification is bound to generate a bad infinite regress. Nor does the coherentist permit what might be called a circular form of linearity, that is, a structure of reasons that simply loops back on itself. For the standard coherentist, linear circularity is a bad form of circularity. In place of such linear conceptions of justification, the coherentist pictures justification using such metaphors as a network, a mesh, a system, or an organic totality of beliefs. The fundamental idea is that the items in coherent sys-

tems of beliefs must stand in relationships of mutual support. Theories of this kind are often referred to as *positive* coherence theories. Blanshard's *Nature of Thought*[2] and BonJour's *Structure of Empirical Knowledge*[3] exemplify this, the standard form, that coherence theories have taken.[4]

A weaker or negative version of coherentism invokes coherence (of one sort or another) only as a constraint or check on justified beliefs. Typically, the negative coherentist does not accept the principle that a belief can be justified only by another belief, but holds, instead, that a belief—just in being a belief—is presumptively true: innocent until proven guilty. Since, as we saw in chapter 7, presumption theories beg the question against Pyrrhonism, I shall not discuss them here. Thus, this chapter is concerned only with positive coherence theories.

In what follows, I will not attempt to survey all the leading versions of positive coherentism. Instead, in this chapter, I will take Laurence BonJour's position as representative of a standard *internalist* version of coherentism. In the next chapter I will examine Davidson's attempt to develop a coherence theory from the perspective of an *externalist* semantics.

BonJour's Version of Coherentism

Here, in broad outline, is BonJour's description of his position:

> In the first place, our concern is with coherence theories of *empirical justification* and not with coherence theories of truth. (88)

> Second . . . I am concerned here only with coherence theories that purport to provide a response to skepticism. (88)

> Third . . . If coherentism is to be even a dialectically interesting alternative, the coherentist justification must, in principle at least, be accessible to the believer himself. (89)

Taking these remarks one at a time, we can note, first, that BonJour defends a coherence theory of justification but adopts a correspondence, rather than a coherence, theory of truth. In the language of the first part of this work, BonJour is concerned with the performance clause (or the epistemic-responsibility clause) in the traditional definition of knowledge. This has the following important consequence for the development of his position: The internal coherence of a system of beliefs does not guarantee that it is true; that is, it does not guarantee that it squares with or corresponds to reality. For BonJour, the correspondence between a coherent system

of beliefs and the world must be established by a further argument that he calls a metajustification.

Second, from the passage cited in chapter 6,[5] it is clear that Bon-Jour is concerned with skepticism of the kind generated by the threat of an infinite regress. Thus, his primary target is Pyrrhonian skepticism. Since, however, he attempts to solve this problem by appealing to private mental contents, BonJour also encounters, and must respond to, problems generated by Cartesian skepticism. That is the task of his metajustification.

Third, his demand that coherentist justification, to be dialectically interesting, must, "in principle at least, be accessible to the believer himself" shows that he is a methodological internalist as I have defined this notion in chapter 6.[6] BonJour is, then, a methodological internalist and an ontological internalist—features that his theory shares with traditional versions of coherentism.

Standards of Coherence

The central task for an epistemologist developing a coherence theory of justification is to tell us, quite exactly, what is meant by coherence. Coherentists typically begin by saying that to be coherent a system of beliefs must, of course, be consistent, but they then go on to insist, as BonJour does, that "coherence is not to be equated with mere consistency" (95). For coherentists, consistency is a necessary condition for a coherent system of beliefs, but not a sufficient condition. We get different variations on the coherentist approach, depending on the nature of the constraints coherentists impose beyond that of mere consistency. A system is sometimes said to be coherent just to the extent that the beliefs in the system support (or probabilize) one another. Sometimes explanatory relationships are made central. And so on.

A distinctive feature of BonJour's coherentism is that it invokes virtually *all* of the standards of coherence that defenders of the coherence theory have employed:

(1) A system of beliefs is coherent only if it is logically consistent. (95)

(2) A system of beliefs is coherent in proportion to its degree of probabilistic consistency. (95)

(3) The coherence of a system of beliefs is increased by the presence of inferential connections between its component beliefs and increased in proportion to the number and strength of such connections. (98)

(4) The coherence of a system of beliefs is diminished to the extent to which it is divided into subsystems of beliefs which are relatively unconnected to each other by inferential connections. (98)

(5) The coherence of a system of beliefs is decreased in proportion to the presence of unexplained anomalies in the believed content of the system. (99)

Aside from invoking a plurality of coherence-enhancing relationships, a second central feature of BonJour's coherence theory of empirical justification is that he does *not* hold that a belief's membership within a suitably coherent system, of itself, constitutes its justification. To be justified in believing something, a person must *recognize* that the belief stands in this relationship to other beliefs— or at least be able to do so. BonJour discusses this internalist commitment under the heading of the *doxastic presumption*. Furthermore, as already noted, BonJour does not hold that a fully coherent system will be self-justifying. For him, avoidance of what amounts to Cartesian skepticism must be met by what he calls a metajustification. A plurality of standards of coherence, a commitment to the doxastic presumption, and the demand for a metajustification give BonJour's coherentism its distinctive character.

Immediate Problems

Critics of BonJour's position are likely to target his commitment to the doxastic presumption and his proposed metajustification for criticism—and I'll look at them later. But difficulties break out earlier that raise serious problems for most versions of positive coherentism, yet often go unnoticed. Given this list of standards for coherence, we can ask whether any human system of beliefs has ever satisfied them. The answer may be no. BonJour, taking a standard coherentist approach, insists that formal consistency is not a sufficient condition for a system to be coherent, but he does take it to be a necessary condition for a system of beliefs to be coherent. He seems not to have noticed that it is very unlikely that the belief system of any human being satisfies this necessary condition. It seems safe to assume that all mature human beings hold at least some beliefs that are inconsistent with each other or at least imply things that are inconsistent with one another. Of course, inconsistencies in one region of our belief system may not infect other regions because they are suitably isolated from one another. This, however, is not a response open to BonJour, given the strong holistic commitments involved in his fourth criterion of coherence. Any coherence theory

that accepts consistency as a necessary condition for coherence and also makes strong holistic demands will yield the result that human beings may not be justified in any of their empirical beliefs and therefore may be wholly lacking in empirical knowledge.[7] Coherentism, then, in the form in which BonJour develops it, far from refuting radical skepticism, poses it as an unanswered challenge. This, I think, is a fundamental criticism of BonJour's position and of a whole family of coherence theories that resemble it. That is not to say that on BonJour's theory human beings *could not* possess empirical knowledge. It does, however, yield the startling consequence that for all we know we are unjustified in our beliefs concerning even the most common things.[8]

Several responses to this criticism are possible. It might be granted that, *loosely speaking,* we are justified in believing many things, even if *strictly speaking* we are not. That, though a concession to human frailty, is not concession enough. Taking the coherentist standards literally, we are not even loosely speaking justified in our common beliefs. The common justificatory procedures discussed in Part I of this study, that is, those justificatory procedures that we actually employ in making knowledge claims, come nowhere near meeting the coherentist's standards laid down by BonJour and others.

A high-minded coherentist of BonJour's stripe can also claim not to be concerned with actual (vulgar) knowledge claims, but instead with an ideal conception of knowledge. BonJour comes close to saying just this:

> Any nonexternalist account of empirical knowledge that has any plausibility will impose standards for justification that many commonsensical cases of knowledge will fail to meet in any full and explicit way. And thus on such a view, such beliefs will not *strictly speaking* be instances of adequate justification and knowledge. But it does not follow that externalism is correct. This would only follow with the addition of the premise that the judgments of common sense as to which of our beliefs qualify as knowledge are sacrosanct, that any serious departure from them is enough to demonstrate that a theory of knowledge is inadequate. . . . Thus while it would take very strong grounds to justify a strong form of skepticism which claims that the beliefs which common sense regards as knowledge have no significant positive epistemic status at all, not nearly so much would be required to make acceptable the view that these beliefs are in fact only rough approximations to an epistemic ideal which *strictly speaking* they do not satisfy. (52–53)[9]

Here BonJour understates the problem. Our ordinary knowledge claims take place within a plurality of disconnected or only loosely connected justificatory procedures. For this reason, they stand in sharp contrast to the holistic demands laid down in BonJour's standards for coherence. Beyond this, as already noted, what BonJour takes to be a modest demand for consistency in the total system of beliefs is completely devastating to our ordinary claims of knowledge. In the passage just cited, BonJour gives the impression that our ordinary knowledge claims, though *strictly speaking* not legitimate, do have a "significant positive epistemic status." Yet on BonJour's account, it is hard to see how this *significant positive epistemic status* is engendered. His own standards ought to lead him to say that our common knowledge claims are hardly justified at all and, in the end, must be replaced by knowledge claims of a different order. The traditional coherentists of the idealist sort were often willing to take the radical step of declaring that the standards of coherence could only be satisfied in an absolute system where our common, fragmented, and partial beliefs have been transcended. For them, all that we commonly think we do know, we simply do not know. It shows a lack of candor that BonJour does not see that his own position carries similar implications.

Finally, BonJour could try to avoid this difficulty generated by his joint commitment to consistency and holism by relaxing at least one of these strong demands. In a footnote, he remarks that other coherentists have given up the strict demand for consistency provided that the total system is sufficiently rich and the inconsistency suitably trivial. He does not adopt this position himself.[10] Alternatively, we might view our belief system as composed of relatively autonomous subsystems, with particular beliefs finding their justification within the subsystems in which they occur. In this way a person might be wholly justified in believing certain things even though other portions of his belief system are flawed.[11] This modification does not, however, solve the problem at hand. The possibility—and it is a real possibility—remains that a belief in one subsystem could be inconsistent with a belief in another subsystem. There seem to be only two ways of dealing with this problem. We can say that a person can be both justified and unjustified in believing something, or, to avoid this, we can relativize justification to the subsystem in which it occurs. Both approaches would be unsatisfactory to an epistemologist who, like BonJour, embraces high standards of epistemic responsibility.

The Doxastic Presumption

As noted, one of the distinctive features of BonJour's version of coherentism is a commitment to what he calls the doxastic presumption. He introduces this idea as follows: "I have so far considered two of the elements which are arguably essential to a viable coherence theory: the idea of nonlinear justification and the concept of coherence itself. A third essential element is the presumption regarding one's grasp of one's own system of beliefs . . . ; this is required, I suggest, in our coherence theory to avoid a lapse into externalism" (101). The externalism that BonJour speaks of at the close of the passage stands opposed to what I have called methodological internalism. For BonJour, it is not sufficient for an epistemically responsible judger to stand in some external relationship to the grounds that justify her belief; the epistemically responsible judger must, in some quite strong sense, *base* her belief on the recognition of these justifying grounds. The central role of the doxastic presumption is explained in more detail later:

> The Doxastic Presumption plays no direct role in the cognitive system but rather formulates something which, from the standpoint of a coherence theory, is an essential aspect of cognitive *practice:* though questions can be raised and answered with regard to particular aspects of my grasp of my system of beliefs, the *approximate* accuracy of my overall grasp of that system must be taken for granted in order for coherentist justification even to begin. (127)

BonJour's commitment to the doxastic presumption raises a number of difficulties, some of which he candidly faces, some of which he does not. But before we examine these difficulties, it is important to see that BonJour's explicit acknowledgment of the doxastic presumption is a distinctive and admirable feature of his own statement of coherentism. Most traditional versions of coherentism—as well as many contemporary versions—simply take some form of the doxastic presumption for granted. It is to BonJour's credit that he explicitly acknowledges this commitment and attempts to deal with at least some of the problems it raises.

One objection is dialectical in character: it concerns the conflict between competing theories of justification. It might be claimed that BonJour's commitment to the doxastic presumption compromises his coherentist program by introducing a foundationalist element. BonJour spends considerable time trying to show that relying on the doxastic presumption, however anomalous in a coherentist position, does not involve a commitment to foundationalism. From the per-

spective of the present study it is not, in fact, of any importance whether BonJour's commitment to the doxastic presumption sullies the purity of his coherentism. Our question is whether BonJour has produced an adequate theory of empirical justification. If the answer to this is yes, then it really does not matter whether his acceptance of the doxastic presumption yields an impure rather than a pure version of coherentism. The central question is whether his acceptance of the doxastic presumption undercuts his position—however that position is labeled.

In fact, BonJour acknowledges, quite candidly, that making his theory rely on the doxastic presumption opens it to a skeptical attack he cannot answer. He introduces the problem in these words: "But doesn't the Doxastic Presumption, or rather the aspect of the cognitive practice which it reflects, amount to begging the question against a certain form of skepticism, namely, that form which would question whether my representation of my own system of beliefs is in fact accurate?" BonJour answers this question as follows: "What the discussion leading up to the Doxastic Presumption shows is precisely that a coherence theory of empirical justification cannot, in principle, answer this form of skepticism; and this seems to me to count in favor of the skeptic, not against him" (105). He then adds: "Thus the position advocated here holds that such a version of skepticism, though certainly unusual, is perfectly coherent (and thus that it would be desirable to be able to answer it) but also concedes that such an answer is unfortunately in principle not available for a coherence theory" (105).

I find this all very perplexing. With disarming candor, BonJour acknowledges that the coherentist cannot meet one form of skepticism—that form of skepticism that challenges the doxastic presumption. It seems, then, that BonJour's position has foundered on Agrippa's third mode: that of hypothesis or arbitrary assumption. If that's so, then his project has failed! Period! There is nothing to be said next if BonJour's task is, as he indicated earlier, to refute skepticism. BonJour has given away the store, yet continues advertising goods for sale.

Setting aside, but not forgetting, this last criticism, we can consider the doxastic presumption on its own merits and ask how plausible it is. Do we, as individuals, have a largely adequate grasp of our belief systems, including the justificatory procedures we use in validating beliefs? The answer to these questions is, I think, no. Our mental landscape is largely *terra incognita*. Here a comparison with our linguistic abilities may be helpful. Through linguistic training we

learn how to employ various expressions, using them with reasonable skill. Yet, as the interminable disputes concerning the correct analysis of common terms show, we seem to possess little second-order knowledge about how these expressions actually function. Similarly, we are trained from early childhood to form and test judgments in various ways. In this way we come to command various justificatory procedures. This is how we act—this, as Wittgenstein would put it, is how the game of judging is played. Yet we often lack an articulate understanding of how these justificatory procedures function. From the perspective of BonJour's antiexternalist program, our common ways of forming what we take to be justified beliefs will seem to be deeply infected with epistemic irresponsibility. Furthermore, it is not simply that our common justificatory procedures fall a bit short of meeting BonJour's internalist standards; they seem to come nowhere near meeting them.

There thus seem to be two grounds for holding that BonJour's theory yields results equivalent to radical skepticism. The first flows from his joint commitment to the coherentist standards of holism and consistency. The second follows from his strong commitment to methodological internalism.

Standard Objections to Coherentism

The criticisms just given constitute my central reasons for rejecting BonJour's version of coherentism. I give them pride of place because they will carry over to all versions of positive coherentism that either (1) lay down both consistency and holistic standards for coherence, or (2) make a strong demand that to be justified in believing something, the justifying grounds must be accessible to (or accessed by) the believer. I think most positive coherence theories embody at least one of these assumptions, and, if that is correct, then most positive coherence theories are inadequate.

There are, however, more standard criticisms that have been brought against positive coherentism: challenges that coherentists recognize and try to answer. BonJour, for example, tells us that a successful positive coherence theory must provide answers to the three following objections:

> (I) *The alternative coherent system objection.* . . . An appeal to coherence will never even begin to pick out one uniquely justified system of beliefs, since on any plausible conception of coherence, there will always be many, probably infinitely many, different and incompatible systems of belief which are equally coherent. (107)

I'll call this the multiple-choice problem.

> (II) *The input objection.* . . . If coherence is the sole basis for empir-
> ical justification, it follows that a system of empirical beliefs might
> be adequately justified, indeed might constitute empirical knowl-
> edge, in spite of being utterly out of contact with the world it pur-
> ports to describe. (108)

> (III) *The problem of truth.* . . . It must be somehow shown that jus-
> tification as conceived by the [coherence] theory is *truth-conducive*,
> that one who seeks justified beliefs is at least likely to find true ones.
> The objection is simply that a coherence theory will be unable to
> accomplish this part of the epistemological task unless it also adopts
> a coherence theory of truth and the idealistic metaphysics which
> goes along with it. (108–9)

If a slogan is needed, we might call this the demand for *coherentism without idealism.*

BonJour's strategy is to begin with the input objection and then use his response to that objection as the basis for responding to the other two.

Coherence and Observation

What a coherence theory needs in order to avoid empirical vacuity is some kind of content that will serve as a check on beliefs within the system of beliefs. BonJour puts it this way: "What is needed to answer [the input objection], speaking very intuitively for the moment, is beliefs whose assertive content is not simply an inferred product of the rest of the system and which can thus constitute an independent check on that system" (113). It is precisely this demand that encourages philosophers in the direction of foundationalism. Foundationalism is, of course, rejected by coherentists, since, for them, no empirical belief is either self-justifying or immediately jus-tified. So the coherentists seem to have reached an impasse. What they seem to need is something like the foundationalists' beliefs to give their theories empirical content, but it is a central feature of their position to deny that such foundational beliefs exist.

BonJour attempts to avoid this impasse by introducing a coher-entist counterpart of the foundationalist's basic beliefs. Such beliefs are basic in that they *arise* without being inferred from any other beliefs. BonJour calls such a belief a *cognitively spontaneous belief*, which he describes as follows: "[It does not] result from any . . . sort of deliberative or ratiocinative process, whether explicit or implicit.

Rather it simply occurs to me, 'strikes me,' in a manner which is both involuntary and quite coercive" (117).

Rejecting even the mildest version of foundationalism, BonJour does not say that these spontaneous beliefs are to any degree self-justified or immediately justified. Thus they cannot enter into the system as justificatory premises. These spontaneous beliefs enter into the justificatory framework in an indirect way suggested first, I believe, by Wilfrid Sellars.[12] It is *the fact that these beliefs occur*, not their putative warrant, that serves the coherentists' purposes.

The easiest way to explain this theory is to state it first naively, then transpose it into a version suitable for inclusion in a coherence theory. (This, in fact, is how BonJour proceeds.) I look out my window and see a ferry crossing over from Menaggio. Without reflecting on the matter, I spontaneously form the belief that this event is taking place. I realize that I cannot always trust my eyes, but nothing in the present circumstances triggers such suspicions. I suppose (or take it for granted) that the circumstances are normal, that is, they are of the kind where my senses perform reliably. Though I may not have reflected on the matter, according to BonJour, my belief is justified in virtue of the soundness of an argument having the following pattern:

> (1) I have a cognitively spontaneous belief that P which is of kind K.
> (2) Conditions C obtain.
> (3) Cognitively spontaneous beliefs of kind K in conditions C are very likely to be true.
> Therefore, my belief that P is very likely to be true.
> Therefore, (probably) P. (123)

For BonJour, the availability of the first premise rests on the doxastic presumption: The ability to recognize cognitively spontaneous beliefs as cognitively spontaneous is, for BonJour, a necessary condition for entering into the practice of coherentist justification. Borrowing from Sellars, he offers the following defense of the second and third premises:

> I think that it is correct to say . . . that for a belief to be genuinely observational requires that it be objectively reliable. . . . But of course the would-be observer has no epistemologically unproblematic access to such objective reliability, and I have already rejected the externalist appeal to facts beyond the ken of the believer. Thus the immediate concern of a coherence theory of justification must be reliability as judged from within the person's system of beliefs. (122–23)

At first sight, BonJour's strategy may seem completely unsatisfactory. It seems on a par with convicts planning an escape where one of the crucial parts of the plan places one of them outside the prison. This, however, is to misunderstand what BonJour is attempting to achieve. In the present context BonJour is not trying to establish correlation or correspondence between a system of beliefs and a reality outside it; he is trying to solve the input problem. He is trying to show how a coherentist program can have a significant observational component at all. His answer to *that* problem is that the occurrence of cognitively spontaneous beliefs together with an *intrasystematic* account of reliability creates a system of beliefs capable of representing (either truly or falsely) a world outside it. He further claims that it is primarily with these observational beliefs that other beliefs must cohere in order to be justified. All this is nicely summarized in the following passage:

> Relative to this grasp [of the doxastic presumption] . . . it is possible to identify certain beliefs as cognitively spontaneous; and also to determine that certain classes of such spontaneous beliefs are, *as judged from within the system,* reliable, that is, likely to be true.
> . . . Such beliefs will then qualify, judging from within the system, as observational; and it is primarily with these observational beliefs that other beliefs must cohere in order to be justified. (138)

A few pages later, he strengthens this demand: "In order for the beliefs of a cognitive system to be even candidates for empirical justification, that system must contain laws attributing a high degree of reliability to a reasonable variety of cognitively spontaneous beliefs. . . . This requirement . . . I will refer to as the *Observation Requirement"* (141). Then, reflecting his zeal for epistemic responsibility, he adds: "A user of a system must make a reasonable effort to seek out relevant, possibly conflicting observations, if his beliefs are to be justified" (142).

Having introduced the observation requirement, BonJour declares the input problem solved. "Thus understood, the Observation Requirement effectively guarantees that a cognitive system which satisfies it will receive at least apparent input from the world and hence that empirical justification will not depend merely on the internal relations of a static belief system; it thus provides the basic answer to objection (II)" (142).[13]

I have serious reservations about this part of BonJour's program because, as I have already indicated, I think the grasp of the system of beliefs it demands is an epistemologist's myth. Instead of pressing

this point again, I will examine the way in which BonJour attempts to use the observational requirement to answer the two remaining standard objections to coherentism. There will be two steps in this defense. The first is intended to show that a system that satisfies the observational requirement can provide a *unique* representation of an external reality—thus solving the multiple-choice problem. The second step is intended to show that a representation that is unique in this way is very likely to be true of an *external* reality—thus solving the truth problem.

The Multiple-Choice Problem

What I have called the multiple-choice problem takes the following form: Isn't it possible for there to be (at least) two incompatible systems of belief such that each meets the observational requirement, and the systems are equally coherent?

In response, BonJour acknowledges that two competing systems may be tied in the coherence derby at some particular time, but then adds that "the most that it seems reasonable to expect of an epistemological account is that it make it possible for such ties to be broken *in the long run*" (144). That, however, simply invites a further question: What guarantees that ties will be broken in the long run?

BonJour's response provides another remarkable example of an argument that blatantly begs the question against Pyrrhonism: "Once the possibility of observational input is appreciated, it is no longer clear why this claim [that multiple choices of systems may always exist] should be accepted, *or at least why it is thought to be any more plausible in relation to a coherence theory than it is in relation to other theories of knowledge*" (144, emphasis added). The italicized portion of this passage is, of course, question-begging. As long as skepticism is in the field of competitors, one cannot argue, as BonJour here argues, that a criticism that applies equally to all theories of justification has no force against any one in particular. I think this last point is worth insisting on because it is a common practice among defenders of one or another theory of epistemic justification to ignore the skeptical alternative their theory is intended to eliminate. Here is another passage from BonJour exhibiting this same tendency: "As should by now be clear, the main motivation for a coherence theory is not any independent plausibility attaching to the idea that coherence is the sole basis for justification, but rather the untenability of foundationalism in all its forms" (149). The Pyrrhonist couldn't ask for anything more.

A further difficulty with BonJour's appeal to long-run observation is that it is not responsive to the claim, often made, that competing interpretive schemata can organize data in incompatible ways, no matter how much data is given to be organized. It is centrally important for BonJour to answer this criticism, but he merely touches on it in a brief footnote: "The familiar Quinean claim that theory is underdetermined by observation would, if correct, apply as much to foundationalist views as to coherence theories" (243 n. 4). Quite remarkable.

The passage continues: "A full consideration of this [Quinean] claim is impossible in this book. It does seem to me, however, that it depends for whatever plausibility it possesses on a quite sharp observational/theoretical distinction for which no adequate defense has been offered" (243 n. 4). This is doubly odd. First, BonJour's theory seems to depend on a sharp distinction between those beliefs that are cognitively spontaneous and those that are not, so it is hardly open to him to complain about a parallel sharp distinction in Quine's position. Moreover, an abandonment of a sharp distinction between the observational and the theoretical seems to enhance, rather than diminish, the possibility for a range of competing belief systems. Conceptual relativism thus remains an unanswered threat to BonJour's position.[14]

It seems, then, that even if we are generous in accepting BonJour's observational requirement, the multiple-choice problem remains unsolved.

Justification and Truth

To follow BonJour's position to its completion, let us suppose, as he does, that he has given adequate responses to both the input problem and the multiple-choice problem. We will suppose, that is, that it is possible to justify particular beliefs by an appeal to an observationally rich, uniquely adequate coherent system of beliefs. If such an internal justification of a belief is possible, does this show that the belief so justified is true? BonJour's answer is no. The truth connection between the internally justified system of beliefs and the facts comprising the external world is, for BonJour, a matter in need of a further independent argument. He calls this argument the metajustificatory argument.

BonJour's position is in need of a metajustification because, unlike his idealist forebears, he is committed to a twofold version of realism:

> Metaphysical realism is the thesis that *an sich* reality exists; semantical realism is the thesis that our statements (and presumably also

our beliefs) purport, in virtue of their meaning or content, to describe this *an sich* reality. And an immediate corollary of the two themes taken together is a minimal version of the correspondence theory of truth: if *an sich* reality genuinely exists and if our statements (and beliefs) purport to describe this reality, then those statements (and beliefs) will be *true* if and only if the reality they purport to describe is in fact as they describe it. (162)

What BonJour calls a "minimal version of the correspondence theory of truth" is, in fact, a rich version of that theory of truth—at least in contrast to the so-called disquotational variant of that theory. As we shall see in more detail later, a disquotational theory of truth is favored by some contemporary epistemologists precisely because it seems to avoid the need for a metajustificatory argument of the kind that BonJour's conception of truth seems to demand.

The goal of BonJour's metajustification is to establish the truth of the following claim: "A system of beliefs which (a) remains coherent (and stable) over the long run and (b) continues to satisfy the Observation Requirement is likely, to a degree which is proportional to the degree of coherence (and stability) and longness of the run, to correspond closely to an independent reality" (171). BonJour's defense of this claim has the form of an inference to the best explanation. Correspondence with *an sich* reality (as he calls it) provides the most probable explanation of such enduring coherence. In particular, it provides a more probable explanation than alternative explanations offered in skeptical scenarios.

It will be important to get clear what sort of probability argument BonJour is offering. Obviously, it cannot be an argument based on the probability calculus. The question "How likely is it that a coherent, stable, observationally rich system corresponds with *an sich* reality?" is not like the question "If two cards are drawn from a deck, how likely is it that the first is a spade, and the second has a value less than six?" Furthermore, the probability argument cannot rely on empirical correlations between beliefs within the system and a reality outside it, for it is this very correlation that is at issue. In place of these standard forms of probabilistic reasoning, which will not serve his purposes, BonJour presents a novel form of a probability argument, namely an a priori argument intended to show that the *antecedent* probability of a skeptical hypothesis is lower than the antecedent probability of the antiskeptical hypothesis that BonJour is defending.

The argument takes the following general form: "The basic suggestion . . . is that it is the very versatility of skeptical hypotheses . . . , their ability to explain any sort of experience equally well,

which renders them not merely methodologically less satisfactory as explanations but less likely to be true, given the fact of a coherent (and stable) system of beliefs" (181). More fully:

> While the whole point of a skeptical hypothesis like that of the evil demon is to be completely and equally compatible with any resulting pattern of experience, and thus neither refutable nor disconfirmable by any such pattern, this is not true of the hypothesis that beliefs are caused by a spatio-temporal world of a more or less ordinary sort. Such a world, unlike the demon, is not merely a neutral producer of beliefs. On the contrary, having as it does a definite and orderly character of its own, such a world would be expected *a priori* to cause beliefs in ways which reflected that character to some degree, not in a completely random fashion.
>
> Second, and more important, there is available a complicated albeit schematic account in terms of biological evolution and to some extent cultural and conceptual evolution which explains how cognitive beings whose spontaneous beliefs are connected with the world in the right way could come to exist. (187)

Others, for example, Wittgenstein in the *Tractatus,* have attempted to use the invulnerability of skeptical scenarios as a basis for dismissing them. Wittgenstein's claim—and many have followed him in this—was that this invulnerability renders these skeptical scenarios meaningless. BonJour does not adopt this line but argues, instead, that vulnerability to refutation gives the correspondence hypothesis a higher degree of antecedent probability than, say, the Descartes demon hypothesis, which, like other skeptical scenarios, is immune to refutation. It is not, however, explained why this difference should lead to a different assignment of antecedent probabilities.

BonJour points out, quite correctly, that there is something fundamentally arbitrary about the demon hypothesis. Why, one might ask, would a demon want to spend his time fooling people? What would be the point of it? In contrast, the correspondence hypothesis contains as a component a specific theory concerning the conditions under which cognitively spontaneous beliefs will occur—it contains a theory of perception. Furthermore, and this is BonJour's second point, this theory of perception squares, at least in a broad way, with the theory of evolution, which is also a component of our internal system. No similar nonarbitrary conditions exist for the demon hypothesis. I think this may well explain why, from within the system of beliefs currently popular in Western intellectual circles, we take seriously the correspondence hypothesis but not the demon hypothesis. But why any of this should bear on the comparative

antecedent probability of these competing hypotheses, I confess that I cannot see. The assignment of antecedent probabilities can take place within the system of beliefs relative to background assumptions, previous investigations, and so forth, and they can also be made arbitrarily (just to get things going). None of these ways of assigning antecedent probabilities will serve the purposes of BonJour's metajustification. BonJour needs some other method for assigning antecedent probabilities: a noncoherentist, nonarbitrary, a priori method. I confess that I find this all quite incomprehensible. It is hard to see that BonJour has done anything more than take over the old verificationalist line and convert its claim about the cognitive meaninglessness of skeptical hypotheses into a claim concerning the antecedent probability of their falsehood.

An Assessment

I will bring this discussion to an end by asking how well BonJour's version of coherentism satisfies the three success conditions for an adequate theory of epistemic justification given in chapter 6. They will have to be reworded to reflect BonJour's concern with empirical justification rather than the stronger notion of empirical knowledge. First, does his statement of coherentism satisfy the demand for philosophical candor? That is, does he clearly designate the range of those beliefs whose status of being justified his theory is intended to vindicate? The answer to these questions is no. It is left entirely unclear which, if any, of those empirical beliefs that are commonly thought to be justified are, for BonJour, really justified. It is left equally unclear what sorts of beliefs might be justified in the future. Second, has he explained, in at least some detail, how our justified empirical beliefs derive their justification? The answer to that question is no. Finally, has he developed his theory of justification without begging the question against Pyrrhonism? No.

Lehrer and the Isolation Objection

I have argued that BonJour's metajustification fails to secure a correspondence relationship between beliefs and the world they are intended to represent. In effect, BonJour has raised the Cartesian problem of the external world, and, like many before him, failed to solve it. In this section I will examine a way of dealing with the truth problem that avoids—or attempts to avoid—turning it into a variant of the traditional problem of the external world. As we saw, BonJour raised this problem by adopting a robust version of a correspondence

definition of truth. On the current philosophical scene, he is almost alone in willingly accepting the burden of solving the truth problem along these lines.

Older coherentists of an idealist stripe attempted to avoid the truth problem by rejecting the claim that correspondence defines truth. For them, coherence not only lays down the criterion for truth, it also defines it. Truth, then, does not consist of a relationship between thought and that which is not thought: it consists of a relationship among thoughts. The most straightforward way of developing such a theory is to adopt some version of metaphysical idealism. A belief, then, in order to be true, need not match something that is not a belief, for beyond beliefs and other mental entities, nothing exists. Theories of this kind are much out of style, though there have been certain incremental movements in their direction.

A more recent innovation is the attempt to solve the truth problem by invoking the so-called disquotational conception of truth. The following specimen sentence taken from Alfred Tarski illustrates how the truth predicate functions disquotationally:

"Snow is white" is true, if and only if snow is white.

By appending the expression "is true" to a quoted declarative sentence, we produce a sentence truth-functionally equivalent to another, where the quotation marks have disappeared along with the truth predicate itself. We might call this a "disappearance" theory of truth.[15]

Why should such a conception of truth be attractive to epistemologists? The answer, at least for some philosophers, is that it seems to define truth in a way that avoids the need for something like BonJour's metajustificatory argument. If the disquotational conception of truth says all there is to say about the nature of truth, then the idea that the content of a true belief must match or mirror reality can be discarded as a misguided piece of imagery. In Donald Davidson's felicitous phrase, given Tarski's conception of truth, we may speak of "correspondence without confrontation."[16]

An immediate problem in using the disquotational or semantic conception of truth is that it cannot be stated in a simple general way without generating paradoxes. These paradoxes come in a variety of forms, but, for our purposes, this example is as good as any:

The sentence on the thirty-seventh line of this page is false.

As it turns out, this sentence occurs on the thirty-seventh line of this page. So if it is true, since it says of itself that it is false, then it is false; furthermore, if it is false, since that's what it says of itself, then it is true. So if it is true, it is false, and conversely.

Because of the threat of paradox, the development of the seman-
tic conception of truth is complex and elaborate. There is also con-
siderable disagreement concerning how the semantic paradoxes
should be dealt with—for example, whether they can be solved or
not, and if they can be solved, how? For our present purposes, how-
ever, we can set these issues aside. If an appeal to the disquotational
conception of truth could solve epistemological problems in the way
that some philosophers think it can, then we would have to exam-
ine the credentials of that theory itself. As we shall see, the disquo-
tational theory of truth will not serve these purposes. To show this,
I will examine Keith Lehrer's use of this conception of truth in his
book *Theory of Knowledge.*[17]

Lehrer offers a defeasibility analysis of knowledge of the kind
examined in chapter 2:

> DK S knows that P if and only if (i) S accepts that p, (ii) it is true
> that p, (iii) S is completely justified in accepting that p, and (iv) S is
> completely justified in accepting that p in a way that is not defeated
> by any false statement. (147)[18]

Lehrer expends most of his energy unpacking, explaining, and
defending the third and fourth clauses in this analysis. Using termi-
nology introduced in the first part of this study, Lehrer is attempting
to buttress (iii), the *performance* clause, using the side-constraints
presented in (iv), the *externalist* clause. I have argued that this is a
misguided effort, but I do not want to go into these matters again.
Here I am interested only in Lehrer's use of the disquotational theory
of truth in offering a response to the isolation objection: the objection
that an "acceptance system and all that coheres with it could occur
in the mind completely isolated from the external world" (143).

Lehrer's initial discussion of the disquotational theory of truth is
brief, and his use of it in defending a coherentist account of knowl-
edge occurs so innocuously in the text, that it can easily pass by
unnoticed. He introduces the disquotational theory of truth in expli-
cating the second clause—the truth clause—in his definition. He rec-
ognizes that the adoption of an unrestricted version of the disquota-
tional theory of truth leads to the kinds of paradoxes noted above. He
offers no solution to these paradoxes. Yet he finds reasons not to be
daunted by this.

For heuristic purposes, Lehrer presents his position through
developing what he calls *justification games.* A claimant to knowl-
edge is confronted by a skeptic who produces possible defeators to
her knowledge claim. If this claimant cannot, on the basis of the sys-
tem of beliefs she accepts, either defeat or neutralize the proposed

competitor, then the claimant loses the justification game and thus is not personally justified in her belief.

Various sorts of justification games emerge as the skeptic is given more and more latitude in raising objections. The development of these justification games roughly recapitulates the sequence of revisions to the defeasibility theory examined in Chapter 2 of this study. An example of a justification game might help motivate Lehrer's position.

> *Claimant:* I see a table in front of me.
> *Skeptic:* You are isolated from the external world.
> *Claimant:* It is more reasonable for me to accept that I see a table in front of me than that I am isolated from the external world. (I am visually connected with the external world and not isolated from it.) (143)

Provided that the claimant's assertion is true, she would have won this round in the justification game, for she would have defeated the skeptic's challenge from within her acceptance system.

I will not follow Lehrer's progress through various levels of justification, but, instead, go straight to what Lehrer calls the ultra justification game. Its task, as we shall see, is to deal with the truth problem or, as Lehrer calls it, the isolation problem.

In the ultra justification game, the skeptic is allowed the following maneuver: "The skeptic may require the claimant to replace anything the claimant accepts that is false with the acceptance of its denial and at the same time replace anything that logically implies the replaced item with the acceptance of its denial" (141). (This is a counterpart of what I earlier called Swain's strategy of unbenighting the believer: a person is justified in believing something only if that person continues to be justified when all her false beliefs are replaced by true ones, etc.) To see how the ultra justification game functions, suppose that, as a matter of fact, the claimant is not visually connected with the external world, though, given her acceptance system, she can reasonably suppose she is. Under these conditions, the game goes as follows:

> *Claimant:* I see a table in front of me.
> *Skeptic:* Replace "I am visually connected with the external world" with "I am not visually connected with the external world." You are isolated from the external world. (144)

This time the claimant loses, for with her acceptance system so altered, she is no longer justified in her belief that she sees a table in front of her.

Suppose, however, that the claimant is, as a matter of fact, visu-ally connected with the world; then the skeptic can not make the defeating move in the last dialogue, for he is not allowed to ask the claimant to replace something true with something false. Let us fur-ther suppose that the claimant remains undefeated through every round of the justification game, including rounds of the ultra justifi-cation game. In that case we get the following strong result: "For per-sonal justification to remain undefeated, it must be true that one is connected in the way one accepts that one is. The truth connection transforms personal justification into knowledge" (144). This pro-vides the basis for Lehrer's solution to the isolation problem:

> The reply to [the] isolation objection is a dilemma. The claimant must accept that she is visually connected with external reality to win the justification game yielding personal justification. Either she is correct in accepting this, and she is so connected, or she is incorrect, and she is not connected. Suppose that she is connected with the external world as she accepts. In that case, she will be vic-torious in the ultra justification game, her justification will be undefeated, and she will turn out to have knowledge on our account. That is the appropriate result in such a case. Suppose, on the contrary, that she is not connected with the external world though she accepts that she is. Then she loses in the ultra justifi-cation game, and she will not turn out to have knowledge on our account. That is the proper result in such an instance since she is truly ignorant. Whether she is isolated or not, our coherence theory of justification yields the appropriate result concerning knowledge. (144)

How does any of this bear on skepticism? It depends on the kind of skepticism in question. If the skeptic is a dogmatic skeptic and claims that nothing about external reality *can* be known, then on Lehrer's account of knowledge, skepticism (in this form) is false. To the claim that nothing can be known about the external world, Lehrer can claim that, *for all we know*, something is known about the external world. None of this has force against the Pyrrhonian skeptic, who suggests that, *for all we know*, we know nothing about the external world. Lehrer's theory of justification provides no response to Pyrrhonian skepticism. Indeed, it seems to entail Pyrrhonian skepticism: for all we know, we know something about the external world, and for all we know, we do not.

Lehrer does not seem to recognize that his theory leads to this strong skeptical conclusion. Indeed, he seems to hide this from his reader (and perhaps from himself) by pretending that the ultra justi-

fication game, rather than being a stand-in for the disquotational account of truth, is actually a game that someone could play and win. Consider the following curious passage:

> Thus, on our theory of knowledge, whether we win or the skeptic wins the day depends on whether what we accept is correct, and especially on what we accept about when we are trustworthy and when we are not. We cannot refute the skeptic by appeal to demonstration. We argue against her from our acceptance system, which is precisely what she calls into question. We may, nonetheless, know that she is wrong. Assuming our complete justification for some of the things we accept is sustained within the members of our ultra system, we know those things to be true, and, indeed, we know that we know. If we do know that we know, then, of course, we know that the skeptic is mistaken in denying we know. (180)

Given our capacity to form nested epistemic claims (claims that we know that we know something), together with an externalist account of the truth relation, it certainly could be the case that we not only know things but also know that we know them. If we do know that we know certain things, then we will also know that skepticism is false. But on Lehrer's account of knowledge, this says no more than that, for all we know, we know skepticism is false. The Pyrrhonian skeptic will grant this, merely noting, for his part, that for all we know skepticism is not false. Lehrer's antiskeptical argument, if targeting a Pyrrhonian skeptic, rather than a dull dogmatic skeptic, is spinning its wheels.[19]

Notes

1. For a characteristic statement of the traditional position on these matters, see, for example, Bernard Bosanquet's *Implication and Linear Inference* (1920).
2. Blanshard, 1939.
3. BonJour, 1985. All page references are to this source.
4. In recent literature, epistemologists have often drawn a distinction between coherence theories of justification and coherence theories of truth, and then have gone on to defend a coherence theory of justification but not a coherence theory of truth. BonJour, as we shall see, adopts just this strategy. In contrast, Blanshard defends what he calls a coherence theory of truth. This may suggest that Blanshard's coherence theory is fundamentally different in intention from that defended by BonJour. That, however, is not right. Blanshard drew a distinction between taking coherence to be the *criterion* of truth and taking coherence to be the *definition* of truth. His theory concerning the criterion of truth corresponds to what is now called a theory

of justification. Though he held that coherence provides both the criterion
and the definition of truth, Blanshard made it his central task in the second
volume of *The Nature of Thought* to defend coherence as a criterion of truth.
Later Blanshard abandoned the view that coherence also provides the defin-
ition of truth.

5. See p. 115.

6. The qualification that the accessibility hold "in principle at least" is
more than a little suspicious.

7. If, unlike BonJour, the coherence theorist extends his theory to
encompass a priori knowledge, this argument will show that, on a coheren-
tist approach, we have no a priori knowledge either.

8. Again, for a vigorous attack on both normative and descriptive
accounts of human rationality, see Stich and Nisbett, 1980; Stich, 1985; and
Cherniak, 1986.

9. BonJour makes a similar claim on p. 152.

10. See 240 n. 7.

11. Morton White has developed a position along these lines under the
title of corporatism, which he contrasts with the holism found in the writ-
ings of Quine and others. He does not, however, develop the position in suf-
ficient detail to show how it would deal with conflicts between different
subsystems. See chapter 2 of White, 1981.

12. Here BonJour cites Wilfrid Sellars's important essay, "Empiricism
and the Philosophy of Mind," which is reprinted in Sellars, 1963.

13. BonJour spends some time discussing the status of the observational
requirement, calling it at one place an a priori requirement (142) and at
another a regulative principle (143). Since BonJour holds a noncoherentist
view of the a priori, the observational requirement cannot be established on
coherentist grounds. It thus occupies as anomalous a place in his theory as
does the doxastic presumption. BonJour describes the observational require-
ment as a regulative principle as opposed to a "first-level epistemic princi-
ple" to distinguish his position from various weak versions of foundational-
ism. Contemporary epistemologists spend a great deal of time worrying
about the dialectical relationships among various theories of epistemic jus-
tification. Here, for example, BonJour is eager to show that his position—
though departing from strict coherentism—is not sliding into the enemy
camp of the weak foundationalists. In this work I will not be much con-
cerned with such intramural controversies, for, to my mind, all theories of
justification go off the track well before such subtle issues are reached.

14. In chapter 9 I will examine Donald Davidson's attack on the notion
that there can be incompatible conceptual schemes. Even if Davidson's argu-
ments on this score are correct, they will not serve BonJour's purposes,
because Davidson employs externalist notions that BonJour rejects.

15. This is all crudely and inaccurately stated, but there is not space to
get things right. I will not, however, trade on my own inaccuracies by impos-
ing them on others.

16. Davidson's position will be examined in chapter 9.

17. Lehrer, 1990. All references, unless otherwise indicated, are to this source.

18. Earlier, Lehrer expressed the fourth clause of his analysis more perspicuously as follows: "(iv) S is completely justified in accepting p in some way that does not depend on any false statement" (18).

19. This criticism parallels the criticism I made earlier, in chapter 5, to Nozick's response to skepticism.

9

External Coherentism

I have two fundamental objections to standard positive coherence theories of the kind developed by Blanshard and BonJour. First, their methodological internalism leads to the skeptical conclusion that we know almost none of the things we commonly think we know, or, to avoid this result, they appeal to an ideal system of beliefs not accessible to the believer and slide, inconsistently, into a form of methodological externalism. Second, their commitment to ontological internalism raises the Cartesian problem of the external world. The long history of unsuccessful efforts to solve this problem suggests that any theory that allows it to arise will be defeated by it. My conjecture is that any coherentist theory combined with either methodological internalism or ontological internalism will be unable to answer skepticism.

If this conjecture is correct, it suggests that coherentism, to have any reasonable chance of success, must be radically externalized. We have already examined a number of externalist theories of knowledge in chapters 3 and 4. I argued there that none of these theories gives a fully adequate account of knowledge, and furthermore, none is capable of responding to the challenges of skepticism they address. None of these theories was, however, presented from a coherentist standpoint. A theory that does combine externalism with coherentism represents a departure in philosophy—one that restructures problems of epistemology in profound ways. Such a theory has been developed in a series of articles by Donald Davidson. Its distinctive feature is to combine a coherentist theory of knowledge and (in a way) a coherentist theory of truth with an externalist account of

meaning.[1] This important innovation is the subject of this chapter. I will structure the discussion around Davidson's programmatic essay "A Coherence Theory of Truth and Knowledge" (CTTK),[2] supplementing it with material drawn from "Epistemology Externalized" (EE),[3] and "Subjective, Intersubjective, Objective" (SIO).[4]

Davidson's Version of Coherentism

The following somewhat obscure passage sketches the problem of epistemology as Davidson sees it:

> If meanings are given by objective truth conditions [which they are] there is a question how we can know that the conditions are satisfied, for this would appear to require [which it does not] a confrontation between what we believe and reality; and the idea of such a confrontation is absurd. (CTTK, 307)[5]

That is, if we combine the doctrine that "meanings are given by objective truth conditions" with the further doctrine that correspondence is the test or criterion of truth, then, in testing for truth, we would have to compare "what we believe with reality." The idea of such a comparison is, according to Davidson, absurd—though he does not say why.[6] To avoid this absurdity, something has to give. Davidson will not abandon the doctrine that "meanings are given by objective truth conditions," so the doctrine that must go is that correspondence is the *test* of truth. In its place, Davidson accepts coherence as the test of truth: "But if coherence is a test of truth, then coherence is a test for judging that objective truth conditions are satisfied, and we no longer need to explain meaning on the basis of possible confrontation" (307). So the connection between truth and meaning looks something like this: "meaning is given by objective truth conditions"; objective truth conditions are tested by coherence; therefore, meaning is not explained by confrontation, but is explained by coherence instead.

If this program can be made good, Davidson thinks that the following consequences will follow: "Given a correct epistemology, we can be realists in all departments. We can accept objective truth conditions as the key to meaning, a realist view of truth, and we can insist that knowledge is of an objective world independent of our language" (307). The significance of these large claims will emerge only as the theory is elaborated in detail.

The Nature of Coherence

For Davidson, as for most coherentists, coherence is a property of a system of beliefs. There are, however, various ways in which we might understand a belief. A belief might be taken abstractly in a wide sense to include anything that someone *might* accept as true. A belief taken this way is rather like the traditional notion of a proposition. In contrast, we might limit our attention to actual beliefs. For reasons that will soon emerge, Davidson's coherentist theory is concerned with actual beliefs of actual human beings: "Beliefs for me are states of people with intentions, desires, sense organs; they are states that are caused by, and cause, events inside and outside the bodies of their entertainers" (308). For Davidson, a system of beliefs is not a set of abstract entities standing, perhaps, in certain logical relationships to one another. A belief is a state of a human being that can be caused (directly or indirectly) by events in the surrounding world and may itself cause events in the world.

A second feature of Davidson's position is that he does not have a highly idealized and restrictive sense of coherence of the kind examined in chapter 8. Instead of demanding that a coherent system must be made up of mutually supporting beliefs where each one of them is true, he tells us, more modestly, "All that a coherence theory can maintain is that *most* of the beliefs in a coherent set of beliefs are true" (308, emphasis added).[7] To avoid giving the impression that beliefs are somehow countable, the same point can be made this way: "This way of stating the position can at best be taken as a hint, since there is probably no useful way to count beliefs, and so no clear meaning to the idea that most of a person's beliefs are true. A somewhat better way to put the point is to say there is a presumption in favor of the truth of a belief that coheres with a significant mass of belief" (308).[8] From this, according to Davidson, it follows that "every belief in a coherent total set of beliefs is justified in the light of this presumption," and "if knowledge is justified true belief, then it would seem that all true beliefs of a consistent believer constitute knowledge" (308).[9]

These last remarks have to be read carefully. Notice that Davidson does not say that all the beliefs of a consistent believer constitute knowledge. Davidson cannot say this because he acknowledges that a belief can be a member of a coherent system of beliefs and thus be justified, but still be false. For Davidson, justification does not have the force of my adequate-grounds condition or Dretske's conclusive-reason condition. Davidson, like Wittgenstein and Chis-

holm, thinks we need epistemic grace to know things, since being justified, by itself, need not guarantee truth.

Setting aside needed elaborations and refinements, as a first approximation, Davidson's account of knowledge comes to this:

(i) If a belief (true or false) is a member of a set of coherent beliefs, then it is a justified belief.

(ii) If, furthermore, this belief is true (and if knowledge is justified true belief), then this belief counts as knowledge.[10]

Truth

If knowledge is treated as justified true belief, three things need explaining: justification, truth, and belief. For Davidson, justification is explained by coherence. Truth comes next.

> It should be clear that I do not hope to define truth in terms of coherence and belief. Truth is beautifully transparent compared to belief and coherence, and I take it to be primitive. Truth, as applied to utterances, shows the disquotational feature enshrined in Tarski's Convention T, and that is enough to fix its domain of application. Relative to a language or a speaker, of course, so there is more to truth than Convention T: there is whatever carries over from language to language or speaker to speaker. What Convention T, and the trite sentences it declares true, like "'Grass is green' spoken by an English speaker, is true if and only if grass is green," reveal is that the truth of an utterance depends on just two things: what the words as spoken mean, and how the world is arranged. There is no further relativism to a conceptual scheme, a way of viewing things, a perspective. (CTTK, 308–9)

This passage says a great deal, and some of it can be fully appreciated only with reference to Davidson's other writings.[11] At this stage I will only note a crucial difference between Davidson's and Lehrer's use of the disquotational theory. Lehrer, as we have seen, uses the disquotational theory to provide a quick deflationary solution to the isolation problem. Davidson does not use the disquotational theory this way, for he holds—and this is a central feature of his externalism—that the isolation problem is itself the product of conceptual confusion.

There is a deflationary component in Davidson's position that comes out in his insistence that disquotational theory helps us see that "the truth of an utterance depends on just two things: what the words as spoken mean, and how the world is arranged," and thus no

further relativization to a conceptual scheme is needed. Disquotation theory plays a key role in showing how there can be correspondence without confrontation, but, by itself, it does not provide a quick and easy (or quick and dirty) answer to the challenge of skepticism.

The Skeptical Challenge

As matters now stand, the following seems to remain an open possibility: a person might have an adequately coherent set of beliefs that are "comprehensively false about the actual world" (309). The skeptic's challenge is to ask what rules out this possibility of massive error. Davidson accepts the challenge head-on: "Even a mild coherence theory like mine must provide a skeptic with a reason for supposing coherent beliefs are true" (309–10). One way of dealing with this skeptical challenge is to declare that it is in some way illegitimate or unintelligible, because, as Wittgenstein put it, "All testing, all confirmation and disconfirmation of a hypothesis takes place already within a system" (Wittgenstein, 1969b, 105). If that is so, it makes no sense to ask for the justification of an entire system, and so such challenges may simply be dismissed as unintelligible. Rorty, as Davidson (correctly, I think) reads him, seems to adopt this dismissive approach. On this point Davidson explicitly parts company with Rorty.

> As Rorty puts it, "Nothing counts as justification unless by reference to what we already accept, and there is no way to get outside our beliefs and our language so as to find some test other than coherence."[12] About this I am, as you see, in agreement with Rorty. Where we differ, if we do, is on whether there remains a question how, given that we cannot "get outside our beliefs and our language so as to find some test other than coherence," we nevertheless can have knowledge of, and talk about, an objective public world not of our making. I think this question does remain, while I expect that Rorty doesn't think so. If this is his view, then he must think I am making a mistake in trying to answer the question. Nevertheless, here goes. (310)

To take stock: Davidson does not respond to skepticism simply on the basis of the disquotational theory of truth. He does not, in the style of Rorty, rest his response simply on holism. He also rejects all internalist attempts to prove the existence of the external world through marshaling adequate *evidence* from within the system of beliefs in the style of BonJour. What he attempts to do instead is to

bring truth and knowledge together through an externalist account of meaning.

Sensation and Belief

One way, and perhaps the only way, that human beings are connected to the world around them is through sensation. This suggests that the right response to the skeptic's challenge is to give an adequate account of beliefs that come from sensation. It is those beliefs, it seems, that connect us to the world. But the relationship between sensation and belief is a matter of dispute. For some (e.g., Chisholm), a sensation is a kind of belief—having a sensation is having a belief. For others (e.g., C. I. Lewis), sensations are not beliefs, although certain beliefs are directly about them and true in virtue of them. But whatever form such views take, it has commonly been assumed that sensations play an evidentiary role for our beliefs about objective reality. Under this assumption, the philosophical program is to show how this evidential base can support a suitably broad range of empirical beliefs. I have already discussed the difficulties such views have in sustaining this burden, and I will not go into this again.

Davidson suggests that, in contrast to these earlier theories, the relationship between sensations and belief is not inferential, but causal: "Sensations cause some beliefs and in *this* sense are the basis or ground of those beliefs" (311). This suggests, in turn, a causal theory of justification or knowledge in the style, say, of Alvin Goldman, but Davidson resists this temptation: "But a causal explanation of a belief does not show how or why the belief is justified" (311). Justification, for Davidson, is a matter of coherence. The task, as Davidson sees it, is to find some way of "transmuting a cause into a reason" (311), and at this point it is far from obvious how this might be done.

At least this much seems clear to Davidson: This problem cannot be solved by trying to link beliefs to the world through causal intermediaries. "If the intermediaries are merely causes, they do not justify, while if they deliver information, they may be lying" (312). Our problem, put simply, is to find some way to link our beliefs about the world to the world in a manner that shows how such beliefs are justified. Causal relations do not solve the problem because they are not justificatory. Epistemic intermediaries generate problems rather than solve them, for we are bound to ask how *they* are related to the world and whether they are reporting it correctly. What is needed is some mode of connection that does not encounter any of these difficulties. Davidson finds it in the notion of meaning.

Meaning and Justification

For Davidson, meaning and justification are linked by the following principle: "Giving the meaning of a sentence will demand that we specify what would justify asserting it" (312). Davidson maintains that we give a bad account of meaning—and bring troubles on ourselves—if we hold that "whatever there is to meaning must be traced back somehow to experience, the given, or patterns of stimulation, something intermediate between belief and the usual objects our beliefs are about" (313).[13] We might call this a foundationalist conception of meaning: meaning, at least in its basic forms, comes in immediately accessible chunks. Now, however, a certain picture forces itself on us. On one side, there are beliefs, with their meanings largely, if not completely, accessible to us. For the most part we simply know, straight off, the contents of our beliefs. On the other side— quite separate—there is the matter of truth: what beliefs mean is one thing, whether they are true or false is a wholly separate matter. "Once we take this step, we open the door to skepticism, for we must then allow that a very great many—perhaps most—of the sentences we hold to be true may in fact be false. It is ironical. Trying to make meaning accessible has made truth inaccessible. When meaning goes epistemological in this way, truth and meaning are necessarily divorced" (313). We seem to have reached a dilemma: "The search for an empirical foundation for meaning or knowledge leads to skepticism, while a coherence theory seems at a loss to provide any reason for a believer to believe his beliefs, if coherent, are true. We are caught between a false answer to the skeptic, and no answer" (314).

For the skeptic, this is a delightful result, for it puts a heavy burden on the antiskeptic. It also suggests that there is only one way to avoid this skeptical dilemma. Davidson puts it this way: "What is needed to answer the skeptics is to show that someone with a (more or less) coherent set of beliefs has a reason to suppose his beliefs are not mistaken in the main" (314). Furthermore, and this is a crucial point: "The answer to [the] problem must be to find a *reason* for supposing that most of our beliefs are true that is not a form of *evidence*" (314). Citing evidence, as Davidson sees, will merely produce a new target for skeptical attack.

The Fundamental Argument

Davidson's argument in behalf of the claim that most of our beliefs are true has two parts, which I will label the third-person argument and the first-person argument:

The Third-Person Argument: "A correct understanding of the speech, beliefs, desires, intentions and other propositional attitudes of a person leads to the conclusion that most of a person's beliefs must be true" (314).

The First-Person Argument: "Anyone with thoughts . . . must know what a belief is, and how in general beliefs are detected and interpreted. These being perfectly general facts we cannot fail to use when we communicate with others . . . there is a pretty strong sense in which we can be said to know that there is a presumption in favor of the overall truthfulness of anyone's beliefs, including our own" (314).

The Third-Person Argument

In the third-person argument, we are asked to imagine ourselves attempting to interpret or translate the utterances of a speaker whose language we do not know. We will further imagine that no translation manual exists and that the language we encounter does not bear any detectable historical relationships to languages we already know. Our informant utters certain sentences on certain occasions and our task is to figure out what this person means by these utterances. On the assumption that the person is using these sentences to express beliefs, that also tells us what the person believes. Interpreting these utterances and discovering the person's beliefs expressed in uttering the sentences is something we do simultaneously.

How should we go about accomplishing this two-sided task? Adapting a strategy from Quine, Davidson suggests we begin by examining those occasions when our speaker *assents* to a sentence. "This is a fair place to start the project of identifying beliefs and meanings, since a speaker's assent to a sentence depends both on what he means by the sentence and what he believes about the world" (315). It is possible that later investigation could reveal that we had the bad luck to begin our efforts at interpretation with a priest bent on keeping the language of his culture secret. Our working assumption is, and must be, that for the most part we are not being hindered by difficulties of this kind. We do this, first, because this assumption seems entirely natural given the role that language usually plays in human interactions: "A speaker who wishes his words to be understood cannot systematically deceive his would-be interpreter about when he assents to sentences" (315). Second, we do not seem to have any other choice but to make this assumption if we are going to attempt to interpret the speaker's sentences. We might call this the sincerity assumption.

But sincerity is not enough. If our informant is sincere but dim-witted (or an aphasic), his sincere assent will provide little help in interpreting the meaning of his utterances. We will have to suppose further that our informant is (largely) cognitively competent. We interpret, that is, using what has come to be known as the principle of charity: "The principle directs the interpreter to translate or inter-pret so as to read some of his own standards of truth into the pattern of sentences held to be true by the speaker" (316). I am not entirely sure what Davidson has in mind in speaking of *standards* of truth, but it at least includes "reading our own logic into the thoughts of the speaker" (316). Here, expanding on ideas introduced by Quine, Davidson suggests that "this leads directly to the identification of logical constants, as well as to assigning a logical form to all sen-tences" (316).

However, it is a further extension of the principle of charity that is most important for our present purposes: "Something like charity operates in the interpretation of those sentences whose *causes* of assent come and go with time and place: when the interpreter finds a sentence of the speaker the speaker assents to regularly under con-ditions he [the interpreter] recognizes, he takes those conditions to be the truth conditions of the speaker's sentence" (316, emphasis added). We were told right at the beginning that the meaning of a sen-tence is given by its (objective) truth conditions. We are now being told—at least for occasion sentences—that the conditions that regu-larly cause assent to a sentence constitute the truth conditions for that sentence. It follows then, for occasion sentences at least, that the conditions that regularly cause the assent to the sentence consti-tute the meaning or content of that sentence. From this it follows that: "We can't in general first identify beliefs and meanings and then ask what caused them. The causality plays an indispensable role in determining the content of what we say and believe" (317). In "Epis-temology Externalized" Davidson refers to this doctrine as "the basic intuition that in the simplest cases words and thoughts refer to what causes them" (EE, 195). It is this basic intuition—this causal, hence externalist, account of meaning—that drives Davidson's argument.[14]

Davidson's causal semantics—as we might call it—works in a simple straightforward way only for occasion sentences. A more elaborate theory will be needed to explain the truth conditions for other kinds of sentences, and here Davidson relies on his project of extending a Tarski-like semantics to natural languages. Without going into that, it is important to note that Davidson extends the principle of charity beyond the interpretation of logical constants

and occasion sentences. As he puts it: "Here I would extend the principle of charity to favor interpretations that as far as possible preserve truth" (316). To speak impressionistically, I'd say Davidson holds that meaning has both an intrasystematic and an extrasystematic dimension. The key idea about the extrasystematic dimension of meaning—the dimension that links the system to the world—is that it is causal. There are, however, many sentences and beliefs that do not stand in direct causal relation to their truth conditions, but are connected only indirectly through other beliefs. Davidson holds that we must apply the principle of charity to the interpretation of these intrasystematic connections as well. If that's correct, then we arrive at the following strong conclusion: "Once we agree to the general method of interpretation I have sketched, it becomes impossible correctly to hold that anyone could be mostly wrong about how things are" (317). Notice that in this passage Davidson says no more than that "it becomes impossible correctly *to hold* that anyone could be mostly wrong about how things are." This raises a question posed by Davidson himself: "Why couldn't it happen that speaker and interpreter understand one another on the basis of shared but erroneous beliefs?" (317). If that could happen, aren't we back in the skeptical soup?

Davidson suggests that anyone who insists on this objection has probably failed to understand the central idea in the position he is developing. "What I take to be the important aspect of this approach is apt to be missed because the approach reverses our natural way of thinking of communication derived from situations in which understanding has already been secured. Once understanding has been secured, we are able, often, to learn what a person believes quite independently of what caused him to believe it" (317). I think this is correct, but Davidson, if anything, understates the case. Not only often, but *typically* we learn what a person believes knowing little or nothing about the circumstances that caused the belief. People say things, and we usually believe that they believe what they are saying. Beyond this, we normally believe what they are saying is true—though sometimes to our peril. We sometimes challenge expressions of belief with respect to either their sincerity or their truth, but this happens relatively rarely. Furthermore, when we do make such challenges, we normally do not expect the person to recall the specific occasion on which this belief was formed.[15] The situation is hardly different with respect to a person's own beliefs. I know, for example, that Salem is the capital of Oregon, yet I have no recollection of the occasion when I first learned this. A great many

of my beliefs have this standing. In this way, the rudimentary causal connections that establish the semantic link with the world lie deeply in the background. It is thus easy, to return to Davidson's point, to reverse priorities and suppose that we start out with a wholly independent understanding of our system of beliefs, then ask how they are connected with the world. This, he tells us, "may lead us to the crucial, indeed, fatal, conclusion that we can in general fix what someone means independently of what he believes and what causes the belief. But if I am right, we can't in general first identify beliefs and meanings and then ask what caused them. The causality plays an indispensable role in determining the content of what we say and believe" (317). From the perspective of an internalist semantics, Davidson's position will seem unpersuasive, but the central point of Davidson's argument is to reject internalist semantics.

But even if we do not accept an internalist semantics, we still want Davidson to answer the question he posed for himself: "Why couldn't it happen that speaker and interpreter understand one another on the basis of shared but erroneous beliefs?" (317). Davidson tells us this might sometimes happen but "cannot happen as a rule" (317). To show this, Davidson uses an intellectual conceit that obviously pleases him. He asks us to imagine an omniscient—or maybe semiomniscient—being playing the role of a radical interpreter.[16] This being is semiomniscient—or so it seems—because She does not know directly (straight off) what beliefs a person has. The heavenly interpreter in Davidson's story assigns beliefs the same way that finite, fallible interpreters do. Constrained in this way, the omniscient interpreter must operate on the assumption that most of the speaker's beliefs are true—by Her standards. She must do this in order to identify the meaning of the speaker's utterances and the content of his beliefs. But what is thought to be true for a heavenly interpreter is true. So a heavenly interpreter will assign to our speaker mostly true beliefs and true utterances. Furthermore, if the heavenly interpreter turns Her attention to a fallible interpreter at work, She will conclude that most of this interpretation is correct— that is, She will interpret the interpretive beliefs of the fallible interpreter as correct.

What are we supposed to conclude from this exotic example? I confess that I am not completely sure. Perhaps it just comes to this:

> An omniscient interpreter constrained to play the radical interpretation game could attribute massive error neither to a person's system of beliefs nor to a finite interpreter's interpretation of another person's system of belief.

Davidson's actual conclusion is this:

> Once we agree to the general method of interpretation I have sketched, it becomes impossible correctly to hold that *anyone* could be mostly wrong about how things are. (317, emphasis added)

The First-Person Argument

If I have followed the argument correctly, so far Davidson has attempted only to make a transition from the claim:

> It is impossible correctly to hold that *another* could be mostly wrong about how things are.

to:

> It is impossible correctly to hold that *anyone* could be mostly wrong about how things are.

If Davidson ended by saying it is impossible to *hold* that anyone (including oneself) could be mostly wrong about how things are, we would have reason to feel dissatisfied. Davidson sees this, for after completing the first stage of the argument, he explicitly asks: "How does this help the person himself who wonders what reason he has to think his beliefs are mostly true?" (318). We will hardly have made headway against the skeptic if we leave open the possibility that a charitable interpreter is attributing a pack of falsehoods to the speaker and to himself. Thanks, as they say, a lot.

In responding to this possibility, Davidson's argument moves very quickly.

> In order to doubt or wonder about the provenance of his beliefs an agent must know what a belief is. . . . *But beliefs are . . . identified, directly and indirectly, by their cause.* What an omniscient interpreter [Her again] knows a fallible interpreter gets right enough if he understands a speaker, and this is just the complicated causal truth that makes us the believers we are, and fixes the contents of our beliefs. The agent has only to reflect on what a belief is to appreciate that most of his basic beliefs are true. (318–19, emphasis added)

The key idea in this passage occurs in a number of places in Davidson's writings, perhaps most clearly in "Epistemology Externalized," where he puts the matter this way: "As long as we adhere to the basic intuition that in the simplest cases words and thoughts refer to what causes them, it is clear that it cannot happen that most of our plainest beliefs about what exists in the world are false. . . .

The situations which normally cause a belief determine the condi-
tions in which it is true" (EE, 195).

The Golden Triangle

The discussion thus far has been limited in an important way. We
have imagined a speaker giving prompted assents in various con-
texts, and we as interpreters trying to discover the conditions that
uniformly elicit these responses. We have, however, been tacitly tak-
ing it for granted that these sentences produced by the speaker form
part of a language. Furthermore, when we suppose that a speaker is
using a language, we have in mind a language that the speaker shares
with others and uses to communicate with them. It could turn out
that our speaker is a linguistic self-starter and has somehow con-
cocted his language on his own. The speaker's language would then
be private in that sense. I don't see any a priori reason why such a
linguistic self-starter might not exist,[17] but, in fact, they don't. Lan-
guage as it appears in the world is largely a community affair, a social
product.

Recognizing the social dimension of language acquisition and
language use allows Davidson to deal with a seemingly difficult
problem and in the process enrich his theory in an important way.
The problem is this: Suppose, again, we are trying to correlate a
speaker's prompted assent to sentences with the circumstances that
elicit them. For example, when a rabbit runs by, the person says,
"That's a rabbit." It is natural to say that it was the rabbit passing by
that prompted the assent. There were, however, many other regular-
ities that are candidates for the prompter of assent. The speaker's
sensory organs were affected in a certain way; so were his inner brain
states. Why pick the passing of the rabbit as the assent prompter in
preference to any of these other closer-in stimuli? If we do pick one
of these closer-in stimuli, then we get the peculiar result that the
speaker's utterance is about, say, his retinal image.

To avoid this difficulty, we need some mechanism that fixes the
appropriate level of content. Davidson finds this in the procedures in
which one person *acquires* a language from another:

> If we are teaching someone a language, the situation becomes more
> complex, but more interpersonal. What seems basic is this: an
> observer (or teacher) finds (or instills) a regularity in the verbal
> behavior of the informant (or learner) which he can correlate with
> events and objects in the environment. This much can take place
> without developed thought on the part of the observed, of course,

but it is a necessary condition for attributing thoughts and mean-
ings to the person observed. For until the triangle is completed con-
necting two creatures, and each creature with common features of
the world, there can be no answer to the question whether a crea-
ture, in discriminating between stimuli, is discriminating between
stimuli at sensory surfaces or somewhere further out, or further in.
Without this sharing of reactions to common stimuli, thought and
speech would have no particular content—that is, no content at all.
(SIO, 7)

This argument has the important consequence that if we are to
ascribe a meaning to a speaker's sentences, this must be done in
terms of a *publicly ascertainable* stimulus. It is for this reason that
the sentence "That's a rabbit" is about a rabbit rather than some
other uniform causal intervener between the speaker(s) and the rab-
bit. We might call this Davidson's public language argument.

Davidson's Externalist Semantics

Davidson holds that his theory, if correct, refutes one form of skep-
ticism—a skepticism that holds that most or all of our (basic, first,
simplest, plainest)[18] beliefs about the world may be false. In order to
see the force of his argument, it is important to recognize that it is
not driven simply by the principle of charity. And a good thing, too.
The principle of charity by itself has no tendency to refute skepti-
cism of the kind under consideration. The point is perfectly simple,
and not denied by Davidson. In interpreting S's utterances, the radi-
cal interpreter, R, must suppose that most of S's basic beliefs are true
by his (R's) standards. That is a methodological requirement to get
the enterprise of interpretation started. R can satisfy that demand
even if most of his beliefs, including those he attributes to S, are
false. There is no way, given the principle of charity alone, to make
the transition from:

(i) In interpreting S's basic beliefs, R must hold that most of
them are true by R's standards.

to:

(ii) Most of S's utterances are true.

In Davidson's theory, the central ingredient in his antiskeptical
argument is not the principle of charity, but, instead, his causal
account of sentence meaning and, relatedly, belief content.
Abstractly we might put the matter this way. Let us suppose that ϕ

specifies the conditions for a belief having the content it does, that is, for being the belief it is. We can further suppose that X specifies the conditions under which the belief is true. We can then ask how the belief content conditions ϕ are related to the truth conditions X. Let us suppose that there is an important set of beliefs such that these two conditions are identical. In that case, these beliefs could not be the beliefs they are without being true.

Plunging ahead, what kind of account of belief content-conditions and belief truth-conditions could yield such a convergence? Various theories might have this result, but for Davidson the causal condition that produces assent does both these things; it makes the belief the belief it is and also lays down the conditions under which it is true. So if the skeptic claims that most or all of such sentences so assented to or most or all beliefs so caused are false, then the skeptic must be wrong. These beliefs cannot be the beliefs they are and also be false.

The Problem of Error

The previous discussion, I think, points to the core of Davidson's argument, but does not represent it accurately, for as stated, it leaves unanswered problems that he consciously seeks to avoid. The most obvious problem is this: In responding to the skeptic, the argument seems to overshoot the mark by making it conceptually impossible *ever* to make a mistake in assenting to an occasion sentence. Relatedly, there could be no such thing as a false occasion belief. That we could write down a false occasion sentence is quite beside the point. What we cannot do, given the theory as I have crudely stated it, is to give our sincere assent to a false occasion sentence. This is an unfortunate result.

Davidson turns his attention to the problem of error toward the close of "A Coherence Theory of Truth and Knowledge." First, we should recall that Davidson does not hold that for radical interpretation to be possible, the interpreter must hold that all of the speaker's occasion sentences are true. It is sufficient for the enterprise for the interpreter to hold that most of them are true. Second, and what is more important, Davidson appeals to the holistic side of his position to deal with the problem of error:

> The problem of error cannot be met sentence by sentence, even at the simplest level. The best we can do is cope with error holistically, that is, we interpret so as to make an agent as intelligible as possible, given his actions, his utterances and his place in the

world. About some things we will find him wrong, as the necessary
cost of finding him elsewhere right. As a rough approximation, find-
ing him right means identifying the causes with the objects of his
beliefs, giving special weight to the simplest cases, and counte-
nancing error where it can be best explained. (318)

Actually, Davidson's emphasis is wrong. It is not simply that the
problem of error arises *even* at the simplest level; it is precisely at
the simplest level that, given his theory, the problem of error is most
pressing. The picture now becomes more complex: It seems that the
causes for assenting to occasion sentences can be assigned in a vari-
ety of ways where decisions on particular assignments will be made
in relationship to the coherence of a total system of assignments. In
some cases overall consistency might best be preserved by holding
that the speaker gave his assent to an occasion sentence erroneously.

So with these necessary adjustments to take care of the problem
of error, we arrive at the following response to skepticism: For an
important class of beliefs, there is sufficient concurrence of truth con-
ditions and meaning conditions such that it is not possible for these
beliefs to be the beliefs they are and for most of them not to be true.
This is guaranteed because, for the most part, the causes that produce
the belief constitute the truth conditions for the belief as well.

I think we can now see how the various parts of Davidson's
attack on skepticism fit together. The disquotational theory of truth
is intended to show how we can have correspondence without con-
frontation, thereby avoiding the necessity of finding a legitimate way
of comparing two distinct realms: thought and reality. The principle
of charity is intended to place strong limits on the range of beliefs
that we can intelligibly attribute to other believers, thereby ruling
out the possibility of competing systems of belief radically different
from our own. The externalist semantics is intended to guarantee a
close fit between meaningfulness and truth for a centrally important
set of beliefs, namely occasion-beliefs, thereby showing that for
these beliefs it is not possible for all or most of them to be false. (The
principle of charity is then invoked to extend this claim beyond
occasion-beliefs to the entire system of beliefs.) Finally, the appeal to
coherence is intended to do at least two things: (1) to provide the
basis for the notion of justification needed for knowledge, and (2) to
provide a way of showing how, despite the strong presumption
against this, occasion-beliefs can sometimes be false. Taken
together, this elaborate argument is intended "to rescue us from a
standard form of skepticism by showing why it is impossible for all
our beliefs to be false together" (319).

The Cartesian Skeptic's Reply

Since Davidson's arguments are directed against Cartesian skepti-
cism, we can first ask how a Cartesian skeptic, or someone who
thinks that Cartesian skepticism demands an answer, would reply.
Davidson's antiskeptical argument is only as good as its component
parts, and each has been challenged. The disquotational theory of
truth has not won universal acceptance and has recently been seri-
ously challenged by John Etchemendy.[19] Not everyone agrees that
the principle of charity can be used, as Davidson attempts to use it,
to rule out the possibility of radically different conceptual
schemes.[20] We must also ask whether Davidson's causal theory of
belief-content is itself plausible—a question that can be answered
only after his position has been fully articulated and applied to a
wide range of linguistic settings. That has yet to be done.[21] These are
all important challenges, and in various places Davidson has
attempted to answer them.[22] But instead of trying to assess the par-
ticular merits on both sides of these complex issues, I shall, as far as
possible, take Davidson's position on its own terms and then ask
whether it is sufficient to refute Cartesian-style skepticism.

The most natural response for a Cartesian to Davidson's anti-
skeptical argument is to say that it is flagrantly question-begging.
Davidson, as we saw, pictures the interpretive situation, as well as
the teacher-learner situation, as triangular in structure:

<div align="center">
Interpreter
(Teacher)
</div>

<div align="center">
Speaker   Objective
(Learner)   Causes
</div>
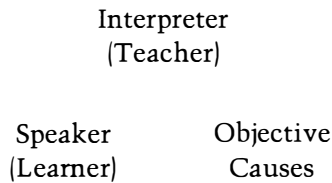
The interpreter, in Davidson's story, assigns meanings to the speak-
er's utterances (and, derivatively, contents to his beliefs) by identify-
ing them with the causes that elicit them. That, the Cartesian will
argue, takes for granted the very point at issue—reliable access (or
any access) to the external conditions that produce beliefs. More
simply, Davidson's whole program is written from a realist stand-
point, and that, in itself, is question-begging.

I don't find this anti-antiskeptical argument persuasive. If one
starts out as an ontological internalist—trapped, as it were, behind
the veil of ideas—then Davidson's realism will seem question-beg-
ging. But before any of this can embarrass Davidson, adequate rea-
sons have to be given for privileging ontological internalism as the
starting place for philosophical reflection. Until such an argument is

produced, it is quite unclear whether there is any question to beg. The so-called problem of the external world is itself the result of an extended piece of philosophizing and is not, as some seem to think, a primitive problem that every epistemological theory must address. If the philosophizing that leads to this problem is flawed, there may be no such problem to be solved.

We can, however, put more pressure on Davidson's antiskeptical argument by asking how it would handle problems that arise from skeptical scenarios. The brain-in-the-vat example discussed in chapter 4 will serve our purposes. To defeat Davidson's argument, we need only show that it is possible for us to be massively mistaken in our beliefs concerning the world around us. Well, it seems possible (thinkable, conceivable, imaginable) that I am nothing more than a brain suspended in a vat with my experiences controlled by a group of clever neurophysiologists. Does this exemplify the possibility that I might have all the beliefs I now have, yet most of them be false? Davidson's answer to this is no: Since the content of beliefs—or at least the content of basic beliefs—is determined by what causes them, our envatted beliefs would not constitute false beliefs about the external world, for they would not be about the external world at all. In short, the conclusion derived from this skeptical scenario actually begs the question against Davidson's position by assuming that the content of a belief is independent of its causes. If Davidson's theory is correct, then the skeptical scenario we have sketched is conceptually incoherent.

Suppose, however, that the brain in the vat is hooked to a speaker and can, in this way, "speak" to us. The physiologist sends some electrical impulses into the brain, and a voice from the speakers says, "I am walking through a bosky glen." The physiologist smiles and says, "We really fooled Vatman that time." But if Donald Davidson were present, he would point out that Vatman has been fooled only if we interpret the words "I am walking through a bosky glen" to mean what we do by these words, and we have very good reason to suppose that Vatman does not use them with this meaning. Vatman's response was elicited by a cause entirely different from the sensations caused by walking through a bosky glen. So we are back to the familiar point that Vatman is not deceived in believing that he is walking through a bosky glen, for his remarks do not refer to a bosky glen.

Things become more difficult when we ask what, on Davidson's theory, Vatman *is* speaking or thinking about, for here an appeal to the triangular relationship examined above seems wholly out of place. Davidson responds to this question in "Epistemology Externalized," where he is discussing a view similar to his own found in the writings of Tyler Burge:

Burge considers a case much like standard brain-in-the-vat cases and shows why such a brain cannot (for long) be radically deceived about its environment. I approve of this argument, having often used it myself. If anything is systematically causing certain experiences (or verbal responses), that is what the thoughts and utterances are about. This rules out systematic error. If nothing is systematically causing the experiences, there is no content to be mistaken about. (EE, 199)[23]

I take it to be the clear implication of this passage that if the physiologist is unsystematic in using the brain probes to induce experiences, then the thoughts have no content; if, on the other hand, he is systematic in using them, then Vatman's thoughts are about these brain probes. If the neurophysiologist is systematic, he cannot fool Vatman for long. If the neurophysiologist is unsystematic, then Vatman will not have a determinate belief, for no determinate content has been established. Thus, for Davidson, as for Burge, these skeptical scenarios are flawed through presupposing an internalist account of semantic content that they reject.

Suppose this argument is correct. The skeptic, as Davidson pictures him, holds that it is possible for most of our simplest, plainest beliefs about the world to be false. Davidson responds: Those beliefs would not be the beliefs they are, indeed, they might not be beliefs at all, unless most of them were true. But now a new and equally fearsome skeptical challenge arises: Couldn't I be mistaken in my belief that I have beliefs; and don't the skeptical scenarios show how this could be possible? On an internalist semantics that contains something like the doxastic principle, this is not possible since, on such a theory, the content of our beliefs is largely accessible to us. Davidson rejects such internalist theories, for, among other things, he thinks that they lead to an unresolvable skepticism. He put it this way in "A Coherence Theory of Truth and Knowledge": "It is ironical. Trying to make meaning accessible has made truth inaccessible" (313). Davidson seems to have fallen into the reverse difficulty. In arguing that most of our basic beliefs (if we have any) must be true, he seems to have foreclosed the possibility of establishing that we have any such beliefs at all. We can call this the semantic update of Cartesian skepticism. It is surely no less palatable than the original.[24]

The Pyrrhonian Skeptic's Reply

It seems that the Cartesian skeptic is not without resources in responding to Davidson's challenge. But instead of pursuing that

matter further, I wish to ask a different question: Even if Davidson's argument, or one of the same general type, is successful against Cartesian skepticism, what force will it have against Pyrrhonian skepticism of the kind examined in this work? The answer is none whatsoever. Davidson's argument is narrowly focused on skeptics who claim that it is possible for "all our beliefs to be false together" (319) and then from this go on to draw skeptical conclusions. Davidson's strategy is to block this skeptical argument by denying the premise from which it starts. Yet, despite the promise in the title "A Coherence Theory of Truth and *Knowledge*," Davidson does not present, even in outline, a fully developed theory of knowledge. In particular, he acknowledges that one central task remains to be done:

> We cannot, alas, draw the picturesque and pleasant conclusion that all true beliefs constitute knowledge. For though all of a believer's beliefs are to some extent justified to him, some may not be justified enough, or in the right way, to constitute knowledge. The general presumption in favor of the truth of belief serves to rescue us from a standard form of skepticism by showing why it is impossible for all our beliefs to be false together. This leaves almost untouched the task of specifying the conditions of knowledge. (319)

What has been left untouched, as the passage indicates, is the problem of justification—what I have called the Agrippa problem. Davidson sees clearly something often missed by others: Dissolving skeptical arguments that turn on a bad theory of meaning does not, straight off, eliminate all skeptical problems. In fact, it is hard to see how the Agrippa problem turns on questions of meaning at all, and it is thus equally hard to see how a correct theory of meaning could aid in its solution or dissolution.

Why, then, if Davidson's position has no bearing on the Agrippa problem, have I spent so much time considering it? First, if correct, Davidson's externalist version of the coherence theory of truth does provide a new response to a standard challenge to coherentism: the possibility that a coherent system of beliefs may still be massively in error. His suggestion that this possibility cannot be eliminated on evidential grounds, but can be blocked using a correct theory of meaning, is a new departure that demands canvassing. If sound, it would represent an important step for the resuscitation of the coherence theory of knowledge—the development of a coherence theory of justification would be the next step. It is important to see that the first step in this project remains in question. Second, many seem to think the externalist semantics championed by Davidson, Burge, and others provides a new and powerful response to the skeptic's challenge in

general. Those who make such claims usually have some (often unspecified) version of Cartesian skepticism in mind. Seeing that Pyrrhonian skepticism, in the form of the Agrippa problem, can survive even if headway is made against Cartesian skepticism shows that the task of refuting skepticism is much larger than has been generally supposed.

Notes

1. Davidson is not alone in combining a coherence theory with an externalist semantics, nor was he the first to do so. In his later writings Wittgenstein's dual commitment to holism and to the public character of language can also be viewed as a kind of coherentism combined with an externalist semantics. In fact, Davidson and Wittgenstein, though starting from different places, show remarkable similarities on a number of fundamental points. This will become clear in chapter 10 and appendix B, where Wittgenstein's views on these and related matters are examined and evaluated.

2. Davidson, 1986.

3. Davidson, 1991.

4. Forthcoming; citations to manuscript copy.

5. I have included the bracketed expressions in this passage because without them the passage is easily misread. Parenthetical page citations in this chapter are from "A Coherence Theory of Truth and Knowledge" unless otherwise indicated.

6. His reason, I suppose, is the transparent failure of theories of knowledge to deal with skepticism when such a demand for confrontation is made.

7. This modest version of coherentism may avoid the charge that I made against BonJour that coherentism yields a functional equivalent of skepticism.

8. Though he notes the difficulty in speaking about *most* of a person's beliefs, Davidson continues to use this phrase and variants on it throughout his essay.

9. It is not clear to me why Davidson here speaks simply of consistency rather than coherence. The appeal to *presumptive* justification is also suspicious, but I will not go into this issue again.

10. In this context Davidson raises an important problem that most coherentists have ignored: "Since no person has a completely consistent body of convictions, coherence with *which* beliefs creates a presumption of truth?" (308). In chapter 8 I argued that the possibility that all human belief systems are inconsistent raises serious problems for standard forms of coherentism. Though Davidson candidly notes this difficult and important problem, he does not present an answer to it.

11. Davidson cites his "True to the Facts" (Davidson, 1969), and "On the Very Idea of a Conceptual Scheme" (Davidson, 1973–74).

12. This passage is from Rorty, 1979, 178.

13. Davidson attributes this view of meaning to both Quine and Dummett. It is a broadly held view of meaning, one that Wittgenstein subjects to devastating attack in the first part of his *Philosophical Investigations*.

14. In his essay "Subjective, Intersubjective, Objective," Davidson offers an elegant account of the two ways in which the principle of charity governs translation by distinguishing what he calls the principle of coherence and the principle of correspondence.

> The Principle of Coherence prompts the interpreter to discover a degree of logical consistency in the thought of the speaker; the Principle of Correspondence prompts the interpreter to take the speaker to be responding to the same features of the world that he (the interpreter) would be responding to under similar circumstances. Both principles can be (and have been) called principles of charity: one principle endows him with a modicum of logical truth, the other endows him with a degree of true beliefs about the world. (Davidson, 1992, 6)

15. Sometimes we do this, for example, when we are evaluating eyewitness testimony.

16. This conceit strikes me as unfortunate, though not in seeming to make Davidson's theory rest on divine favor. Nobody, I think, is dim enough to read him in this way. The example is unfortunate because, as we shall see, it ties the heavenly interpreter's epistemic hands in a way that may seem question-begging.

17. See Fogelin, 1987b, chapters 11 and 12.

18. Davidson uses all of these qualifying terms in one place or another. Sometimes he speaks of "most of our beliefs" without qualification.

19. See Etchemendy, 1990.

20. See, for example, the sixth chapter of Cherniak, 1986.

21. I owe this point to Hilary Kornblith. In his opinion, the prospects for the successful development of such a causal theory of content are not promising.

22. For a critical evaluation of Davidson's views on these matters, see, in particular, LePore, 1986.

23. Davidson here specifically cites two works by Burge: "Cartesian Error and the Objectivity of Perception" (1986), and "Individualism and the Mental" (1979).

24. In conversation, Walter Sinnott-Armstrong made the point this way: Davidson's argument seems to leave open the radical skeptical possibility that most of those things that we take to be true beliefs are either false beliefs or not beliefs at all.

10

Pyrrhonism

Neo-Pyrrhonism

At the beginning of chapter 5—echoing ideas from the Introduction—I described the Pyrrhonian skeptic as

> going through the world claiming to know certain things, and sometimes claiming to be sure or even absolutely dead certain of them. The Pyrrhonian skeptic freely participates in common epistemic practices, drawing on all the practical distinctions embodied in them. These practices are often fallible. Often this fallibility doesn't matter, since the price of being wrong is not high. When the cost of error becomes excessive, the skeptic, like others, may seek ways to improve these practices so that the chances of error are reduced. Pictured this way, the skeptic is rather like Hume's moderate skeptic (whom he improperly contrasted with the Pyrrhonian skeptic): cautious, agreeable, and sane.

I argued, against Barnes and some others, that this way—essentially Frede's way—of treating Pyrrhonian skepticism is textually correct. I will not go into that again. I would, however, like to respond to another complaint that might be expressed in these words: "If *that's* what Pyrrhonian skepticism amounts to, it is hardly a specimen of skepticism at all. Such a version of skepticism is pretty thin soup."

At first glance, Pyrrhonian skepticism may, indeed, seem mild in comparison with various forms of Cartesian skepticism. There is something exhilarating, almost giddy, in the thought that all of our common beliefs about the world might just be false, and Cartesian skeptical scenarios seem to raise just this possibility in a vivid form. But, as I have said a number of times, Cartesian skepticism seems to rely on an antecedent philosophical commitment to the way of

192

ideas—a commitment that a Pyrrhonian skeptic would not make. To the Pyrrhonist, the Cartesian-style skeptic is not skeptical enough. More to the point, it does not take radical—globally dislocating—scenarios to introduce suspension of belief. It is quite sufficient to note—and dwell on—the fact that our empirical claims are made in the face of unchecked, though checkable, defeaters. This is an important point to make, because it might be possible to bring forth arguments showing that skepticism based on skeptical scenarios is conceptually incoherent. Davidson's arguments, which were examined earlier, if made good might show this. The skeptical problems raised by checkable but unchecked defeaters cannot be dealt with in a parallel fashion. Given any empirical assertion, it is always possible—indeed always easy—to point to some uneliminated (though eliminable) possibility that can defeat this claim. Nothing like brains in vats are needed to achieve this purpose. It doesn't even take a great deal of ingenuity to raise these skeptical doubts. A reliance on examples involving papier-mâché will usually be sufficient. Dwelling on uneliminated defeaters can produce skeptical doubts no less strong than those produced by skeptical scenarios. If anything, the situation is worse with uneliminated but eliminable defeaters. With respect to them, no transcendental style of argument will work; the only way to eliminate these defeaters is actually to eliminate them. The recognition that we make knowledge claims without doing so gives one as robust a skeptical challenge as one would like.

It is the fragility of our common epistemic practices that leads philosophers into justificationalist programs of the kind examined in the second part of this study. Justificationalism in both its foundationalist and nonfoundationalist modes is an attempt to secure a suitably wide range of knowledge claims against skeptical challenges. But, as we have seen, such justificatory programs inevitably raise the Agrippa problem. The central thesis of Part II of this study is that no justificatory program seems to show any prospect of solving the Agrippa problem. The strength of each position seems to be wholly exhausted in the weaknesses of its competitors. We have thus arrived at the following result: Reflection on our ordinary epistemic practices reveals their fragility, and when we turn to epistemologists for help, we are disappointed.

Again, Is There a Fact of the Matter in Knowing?

I initially raised the question whether there is a fact of the matter in knowing in chapter 5, then postponed answering, for at that stage of

the investigation I could make no more than the following negative
point:

> My analysis implies nothing one way or another on this matter, and
> this is precisely because nothing in this analysis either privileges or
> refuses to privilege particular justificatory frameworks. For all that
> has been said so far, there may be one ultimate justificatory frame-
> work that grounds all others. There may be a plurality of justifica-
> tory frameworks that ground various domains of knowledge. There
> may be no justificatory framework that stands up under the unlim-
> ited heightening of scrutiny. In fact, this third possibility strikes me
> as being correct, but that is not something that follows from the
> analysis of knowledge claims I have presented. It is something that
> has to be shown in detail by examining various attempts to produce
> philosophical theories of justification. That's the task of the second
> part of this study.

What conclusion can be drawn now that the survey of justifica-
tory theories has been brought to a close? I have not attempted to
survey every theory of empirical justification and to show that each
of them is unsatisfactory. I have not, in lieu of this, attempted to pro-
duce a systematic and exhaustive classification of every possible the-
ory of empirical justification and then, on the basis of this classifi-
cation, argued that every possible type of theory of justification must
be inadequate. The classificatory scheme I introduced in chapter 6
served only expository purposes. I have done nothing more than
examine a set of characteristic theories of justification, each of
which can be taken as representative of its own kind. My claim is
that the theories I have examined come *nowhere near* providing an
adequate response to the Agrippa problem. It does not seem likely
that any of the theories examined can avoid the kinds of criticisms I
have leveled through some *refinement*. It is possible that someone
will produce a wholly new sort of theory of empirical justification
that will provide a satisfactory solution to the Agrippa problem, or
perhaps someone will accomplish this through hitting upon an
utterly novel way of developing one of the traditional theories of
empirical justification. It would be an unseemly dogmatism to rule
these possibilities out in advance. What I have tried to show, using a
number of exemplary cases, is that Pyrrhonian skepticism, when
taken seriously and made a party to the debate, is much more
intractable than those who have produced theories of empirical jus-
tification have generally supposed. As far as I can see, the challenge
of Pyrrhonian skepticism, once accepted, is unanswerable.

Supposing, then, that an adequate theory of empirical justification—that is, an adequate response to the Agrippa problem—is not possible, how does this bear on the question whether there is a fact of the matter in knowing? Part of the answer simply repeats things I said in Chapter 5. At a particular level of scrutiny there will be a fact of the matter (or facts of the matter) that will settle the question whether something is known or not. The crucial question is whether any justificatory procedure is privileged in the sense that justification of that kind is final. If the Agrippa problem cannot be solved, then the answer to that question is no. From this it follows that, in the sense in which philosophers (in particular, justification theorists) have sought a fact of the matter for knowledge, there is none. If the Agrippa problem cannot be solved, there is no reason to suppose that knowledge of the kind sought by justificationalist philosophers exists.

The Pyrrhonist's Use of Epistemic Terms

This discussion began with the concern that Pyrrhonian skepticism is so mild—so anodyne—that it hardly merits the title of skepticism. It should be clear by now that this is not so. Here, however, a concern of a reverse sort presents itself: Having unleashed what amounts to an unmitigated skepticism with regard to empirical justification, how can the Pyrrhonists, in good faith, continue to employ—apparently without qualms—standard terms of epistemic appraisal? The answer is that the Pyrrhonist is under no constraint to conform his activities—including his linguistic activities—to philosophical standards. In daily life, levels of epistemic standards are fixed (often unreflectively) by the exigencies of the given context. The Pyrrhonist undogmatically accepts the everyday epistemic practices of his culture.

At this point, however, it is important not to turn Pyrrhonism into yet another version of justificationalism by reading it as a social theory of justification. The Pyrrhonist, we might say, falls in with others in their modest, undogmatic, epistemic practices. The Pyrrhonist, like others, simply enters into what Wittgenstein calls forms of life, and does so without believing that these forms of life are justified. Furthermore, if we press for justification—and here the ancient Pyrrhonists and Wittgenstein concur—we quickly become aware that none is forthcoming.

Why don't we feel this lack of justification? Well, philosophizing

can make one feel it; but why, if it is so ready at hand, don't we feel it in our everyday use of terms of epistemic evaluation? The answer—and here I follow Wittgenstein—lies in the way in which we acquire and then employ ordinary concepts. Only rarely are we actually told that the procedures we learn to employ are correct procedures, and rarer still is their correctness actually proven. In school, for example, we are not taught that reference books are reliable; we are simply taught to look things up in reference books. In general, we are taught *from* books with no doubt cast on their reliability. It is only later, if at all, that we are told that some books are not reliable, perhaps because they have been carelessly written or are out of date. At no point in our education are we presented evidence for the general reliability of reference books.[1] Again, we are not taught that our senses are reliable. Our parents and teachers treat our senses as reliable, unless, that is, they have reasons to suppose that they are not. Our parents and teachers rely on the reliability of our senses in order to teach us rudimentary truths about the world around us. Later we gain quite specific information concerning those occasions where our senses may "mislead" us. Accepting the senses as reliable is part of the background framework we acquire in the process of learning other things. Learning when the senses cannot be trusted constitutes a minor adjustment, a fine-tuning, within this background framework. I think we can now see why skeptical doubts do not have a purchase on our everyday cognitive activities: It is part of our training in the use of ordinary concepts, including ordinary epistemic concepts, to exclude or bracket such doubts. This, I think, is what Wittgenstein is stressing in the following passages:

> *OC*, 472. When a child learns language it learns at the same time what is to be investigated and what not. When it learns that there is a cupboard in a room, it isn't taught to doubt whether what it sees later on is still a cupboard or only a kind of stage set.

> *OC*, 473. Just as in writing we learn a particular basic form of letters and then vary it later, so we learn first the stability of things as the norm, which is then subject to alterations.

Is Skepticism Statable?

I think Wittgenstein's answer to this question is no. My answer to it is yes. Perhaps surprisingly, it seems to make little difference which answer we give, for, as I shall argue, the challenge of Pyrrhonian skepticism remains even if it cannot be put into words. Since the

supposed unstatability of skepticism is sometimes thought to be a reason for rejecting it, I'll look at this matter in some detail.

As I read him, Wittgenstein seems to be saying that the term "to know" and other terms of epistemic appraisal appear in our discourse in just those circumstances where routine, unreflective procedures are likely to prove unreliable. In my words, terms of epistemic appraisal find their typical employment in contexts where higher (or closer) methods of scrutiny have been triggered. In typical settings, the claim to know something is to assert that these (now) higher standards have been met. Here the claim to know amounts to the claim of having adequate reasons, of the sort now demanded.[2]

If Wittgenstein's reflections are correct, they explain why it seems odd for Moore to claim to know such things as that he possesses two hands or that the world has existed for many years before his birth. These are odd things to say since nothing has put them in question. It is for this reason that Wittgenstein thinks that Moore, in attacking the idealists by citing such uncontested facts, has actually misused the verb "to know."[3] Early in *On Certainty* he comments:

> *OC*, 11. We just do not see how very specialized the use of "I know" is.

Next, given Wittgenstein's at least near-identification of the meaning of an expression with its use in the language,[4] this further suggests that the meaning of "I know" is given by this very specialized use. This use appears in contexts where there is some practical reason for doubting, and therefore a practical point to supplying reasons that will resolve these doubts. It is in such contexts, Wittgenstein seems to be saying, that the expression "I know" finds its use, and hence its meaning. If all this is correct, then the meaning of "I know" is tied to these special sorts of contexts, and when used outside them the expression loses its meaning. This leads Wittgenstein to the following conclusion or, perhaps, recommendation:

> *OC*, 260. I would like to reserve the expression "I know" for the cases in which it is used in normal linguistic exchange.

Let's look at 260 carefully. If we could reserve the expression "I know" for those cases in which it is used in normal linguistic exchanges, it would seem that we would then have a new and powerful argument against skepticism. Everything will depend on how we take the word "reserve." We might, perhaps, just *tell* people to reserve the expression "know" for such cases. They might ask why they

should do this. The reply that this will help them avoid getting entangled in skeptical problems might be a correct, even a useful, thing to point out, but it would hardly count as a refutation of skepticism. *OC*, 260 will have force against the skeptic only if it is taken as saying that we *ought* to reserve the expression for these cases, and the reason that Wittgenstein seems to be giving for this is that the expression "I know" will lose its meaning if its use is not so restricted.

The following passage from *On Certainty* bears directly on these issues:

> *OC*, 392. What I need to shew is that a doubt is not necessary even when it is possible. That the possibility of the language-game does not depend on everything being doubted that can be doubted.

I think Wittgenstein wants to say—and all but says—that the exclusion of such remote possibilities is part of the meaning, part of the semantic content, of knowledge claims. If that is his position, then at this point I part company with him.

The basis for this disagreement is derived from Paul Grice's distinction between what the use of an utterance conversationally implicates, and what the statement made in using an utterance actually says or entails.[5] For Grice, conversational implicatures arise because our linguistic exchanges are governed by a system of conversational maxims. One such maxim he calls the rule of quality. It has two parts. The first tells us not to assert things we believe to be false; the second, which is more to our present purposes, tells us not to assert things for which we lack adequate evidence. Because normal conversational exchanges are governed by these maxims, when someone asserts something, that person conversationally implies, but does not assert, that she believes what she is asserting and, furthermore, has adequate grounds on behalf of the assertions she makes. To the best of my knowledge, Grice nowhere discusses the phenomenon I have called varying levels of scrutiny, but it is easy enough to sketch a conversational maxim applicable to it. In a given context, the level of scrutiny or the appropriate level of adequacy will be fixed by the purposes and goals of the conversational exchange: more specifically by the standardness or nonstandardness of the setting, by the benefits of being right, by the costs of being wrong, by professional norms, and the like. In common conversational exchanges, parties to it naturally adopt mutually recognized standards of adequacy, and hold each other to them. The conversational maxim for levels of adequacy, then, might look like this:

Make your conversational contributions such that they conform to
the standards of adequacy mutually adopted within the conversa-
tional exchange.

It is possible to attempt to raise these standards, but in doing so we are
usually expected to signal this ("I know this may seem farfetched—
finicky, nitpicking, etc.—but . . . "). This suggests that a seemingly
remote possibility is relevant. We can also say something, then cancel
its standard conversational implication ("It is possible that both
backup systems failed simultaneously, but that's hardly likely"). Here
a possibility is set aside as too remote. But if no signals are given to the
contrary, it is a standard implication of raising an objection that it is
congruent with the level of scrutiny governing the conversational
exchange.

This is all too quick, but perhaps it is sufficient to make the
point that Grice has made, that it is easy to confound semantic con-
tent with content conversationally implied. Going back to Dretske's
example of the painted mule discussed in chapter 4, it is natural to
say that, in identifying an animal as a zebra, one meant a zebra rather
than, say, an antelope, not a zebra rather than, say, a painted mule.
There is a perfectly good sense of "meant" where this is correct. In
identifying the animal as a zebra, it was no part of my *intention* to
rule out the possibility that the animal was a painted mule. Yet if my
assertion is true, then this remote possibility is ruled out. The same
point holds against Wittgenstein. If what he is saying is that ignor-
ing or bracketing remote possibilities is part of what is meant by
claims to know things, then he is right in a way and wrong in a way.
In making everyday knowledge claims, I do not pretend to have elim-
inated all defeaters. A great many such possible defeaters are simply
set aside, ignored, or bracketed. Still, if one of these bracketed
remote possibilities is realized or is likely to be realized, this will
bear *directly* on my knowledge claim.

Suppose, however, someone restated the argument this way: "If
remote possibilities are not bracketed, our empirical knowledge
claims will always fail. But if all of our empirical knowledge claims
fail—and that is something that we can see in advance—then we
arrive at the absurd conclusion that our language contains an elabo-
rate system of statements that in every instance of their use is false."
The answer is that, in daily life, we unreflectively bracket remote
defeating possibilities, and for this reason find it possible to make
epistemic claims. This, however, is a Pyrrhonian, not an anti-
Pyrrhonian, point. It is possible to *describe* those circumstances

under which we employ epistemic claims in a nontentative way. This description, however, does not show that our epistemic practices are legitimate. If anything, when their actual grounds are revealed for what they are, their legitimacy seems to be lost.

My basic point, which I have made a number of times, is that radical skeptical doubts can be raised without appeals to skeptical scenarios. Skeptical doubts are ready at hand to anyone who steps back and makes our ordinary justificatory procedures objects of scrutiny. This is not something we ordinarily do. It is even something we resist doing. Yet once we step back and take those procedures that govern our thought as objects of thought, we cannot help being struck by their fragility. Our cognitive activities depend on our taking for granted things that, at any time, can let us down or become genuine objects of doubt. As Wittgenstein notes on a number of occasions, it is not that our inquiries rely on things that we *cannot* doubt; instead, they rest on things that we *do not* doubt.[6] To this I will only add that, once these nondoubted commitments are surfaced—and taken seriously by being made objects of inquiry—*as a matter of fact*, doubt follows. I see no reason why this doubt, when invoked by reflections on the "groundlessness of our believing," cannot be put into words—such words as "When it comes right down to it, I really do not know that . . ." or "It is really surprising how little I have to say in support of my belief that . . ." It will be *misleading* to make such remarks in a context where others are operating (and assume that you are operating) at a normal or unheightened level of scrutiny. But this unwanted conversational implication can be canceled in the same way that other unwanted conversational implications can be canceled: simply by explicitly canceling it.

It seems to me, then, that there is no good reason for rejecting skeptical assertions on the grounds that they are meaningless, and no good reason for not expressing them on the grounds that they are inexpressible. Tutored by Grice, we now see that the use of an expression may be odd—even exceedingly odd—yet, for all that, be meaningful and perhaps true. That, broadly speaking, is the status I assign to the Pyrrhonist's skeptical pronouncements. It is for this reason that I find various linguistic refutations of skepticism unpersuasive. Having said that, I now want to make a second, perhaps more surprising, response to such linguistic attacks on skepticism: Even if it could be shown that skeptical doubts are inexpressible, this need not constitute a refutation of skepticism.

Returning to Wittgenstein, we find he tells us that it is difficult

"to realize the groundlessness of our believing" (*OC*, 160). Well, suppose we do realize this, what then? Is this something we are allowed to say? If we can say it, haven't we managed to put our radical, unrestricted skeptical doubts into words without destroying their sense? In back-to-back passages, Wittgenstein seems to take two quite different views on this matter.

> *OC*, 369. If I wanted to doubt whether this is a hand, how could I avoid doubting whether the word "hand" has any meaning? So that is something that I seem to *know* after all.

Except for the occurrence of the word "seem," this sounds like a standard linguistic/transcendental argument against skepticism. I cannot, except perhaps in extraordinary circumstances, doubt that I have a hand, for such a doubt would deprive the term "hand" of its meaning, and thus the doubt would lose its content. There are other passages in Wittgenstein's writings that seem to say much the same thing,[7] and it is easy to suppose that this is Wittgenstein's fundamental response to skepticism. Yet in the very next entry Wittgenstein adopts a different and, to my mind, a much deeper response.

> *OC*, 370. But more correctly: The fact that I use the word "hand" and all the other words in my sentence without a second thought, indeed that I should stand before the abyss if I wanted so much as to try doubting their meaning—shews that absence of doubt belongs to the essence of the language-game, that the question "How do I know . . ." drags out the language-game, or else does away with it.

Here Wittgenstein is plainly not arguing that our secure grasp of the meaning of the word "hand" guarantees that (at least for the most part) I cannot be mistaken in calling something a hand. Here he is making the deeper point that our grasp of the meaning of the word "hand" and our ability to be mostly right in identifying hands are interrelated, and *both* depend on the fact that we don't *drag out the language game* by pressing the question "How do I know . . . ?" So there is no argument here that invokes the meaningfulness of language as the guarantor of truth. Our knowledge that this is a hand and our understanding of the meaning of the term "hand" stand and fall together. If Wittgenstein is right, under the pressure of unrestricted challenge, they both fall. Skepticism, taken seriously, opens the abyss of meaninglessness.

I think we might be able to see why Wittgenstein, under the influence of such reasoning, would resist putting skeptical doubts into words: It would be an attempt to say something that cannot be

said, but only shown. This contrast between saying and showing goes back to the *Tractatus*, where it was one of Wittgenstein's leading ideas. It was also a distinction, it seems to me, that Wittgenstein never abandoned, even if his specific reasons for invoking it changed over time. Toward the end of *On Certainty*, we find this doctrine reemerging precisely at the point at which he is trying to talk directly about the underlying supports of our language games.

> *OC*, 501. Am I not getting closer and closer to saying that in the end logic cannot be described? You must look at the practice of language, then you will see it.

Among other things we are supposed to see if we look is the "groundlessness of our believing" (*OC*, 166). But, for Wittgenstein, seeing our beliefs *as* groundless not only undercuts their status as beliefs, but also undercuts their status as possible objects of belief—or, for that matter, possible objects of doubt. The abyss that Wittgenstein refers to in *OC*, 170 is indescribable, but makes itself manifest.

We can now, I think, make sense out of Wittgenstein's seemingly ambivalent attitude toward skepticism. Throughout his philosophical writings he rejects skeptical arguments of a Cartesian kind. That is, he rejects direct skeptical arguments that are based on what are now called skeptical scenarios. He holds that such doubts are pseudo-doubts because they admit, in principle, of no procedure for answering them. In this study I have not chanced my hand on the question whether Wittgenstein and others are right in thinking that global skepticism of this kind can be dismissed as meaningless. I am inclined to think it cannot, but nothing in this study turns on this matter. Though he clearly rejects Cartesian skepticism, it seems to me that Wittgenstein is clearly committed to the version of skepticism that I have labeled Pyrrhonian skepticism. That he holds, or seems to hold, that this commitment cannot be expressed directly but must, instead, make itself manifest, in no way takes back or mitigates this commitment. Viewed in this light, whether Pyrrhonian doubts are statable, as I think they are, or unstatable, as Wittgenstein seems to suggest, becomes a matter of relative indifference, since the skeptical challenge remains in either case.

A Temporary Stopping Point

This study falls into two parts that, on the surface at least, may seem to exhibit incompatible tendencies. The first seems to embody old-fashioned philosophizing in its attempt to give an analysis of knowl-

edge; the intent of the second part seems to be to undercut philosophy. This may seem inconsistent. In fact, the two parts of the study fit together in the following ways. First, it seems to me to be possible to give an account of the way in which terms of epistemic appraisal function without making any commitments to substantive claims about what is and what is not known. I realize that this is a controversial claim, and in chapter 5 I have tried to say something in support of it. In any case, I do not think the analysis in Part I relies on the kind of philosophical commitments challenged in Part II.

Second, there is an important systematic connection between the two parts of this study. If the analysis in Part I is correct, including the notion of shifting levels of scrutiny, then we can see how demands for philosophical modes of justification can spring quite naturally from our ordinary concept of knowledge. It takes nothing more than reflection on our ordinary modes of justification to feel the need for something more. Our knowledge claims have an objective component. Although knowledge claims are always made within restricted frameworks, they are not relativized to these frameworks. In the language of Part I, the demand for adequate grounds is not relativized to a particular framework with a fixed level of scrutiny, even though the assessment of a responsible epistemic performance is. It is this disparity between the objective demands of the adequate-grounds clause and the relativized demands of the epistemic-responsibility clause that, once perceived, generates the demand for philosophical theories of justification.[8] An attempt to transcend our actual modes of justification is an immediate and natural consequence of noting their fragility.

These reflections, then, seem to yield a dual conclusion. First, Pyrrhonian doubts are the natural and intelligible result of the unrestricted examination of our epistemic practices. Second, Pyrrhonian doubts, once raised, seem incapable of resolution.

Notes

1. Cf. *On Certainty* (Wittgenstein, 1969b):

   OC, 600. What kinds of grounds have I for trusting text books of experimental physics?
   I have no grounds for not trusting them. And I trust them. I know how such books are produced—or rather, I believe I know. I have some evidence, but it does not go very far and is of a very scattered kind. I have heard, seen and read various things.

2. See *OC*, 243.

3. See *OC*, 6.

4. At section 43 of the *Philosophical Investigations,* Wittgenstein says, "For a *large* class of cases—though not for all—in which we employ the word 'meaning' it can be defined thus: the meaning of a word is its use in the language." It is not clear from the text why Wittgenstein states this doctrine in a qualified form.

5. See Paul Grice's "Logic and Conversation," reprinted in Grice, 1989.

6. See, for example, *OC*, 232, 342, and 509.

7. See, for example, *OC*, 80, 81, and 114.

8. Thompson Clarke makes a parallel point concerning Cartesian skepticism in Clarke, 1972.

Appendix B: Two Wittgensteins

Throughout this work I have relied on ideas derived from Wittgenstein's later writings, where he rejects all attempts to use philosophy to support, ground, or vindicate our common cognitive activities. On this reading, Wittgenstein is the paradigmatic neo-Pyrrhonist. There are, however, other strains in Wittgenstein's later philosophy that suggest that he was not an opponent of philosophical justificationalism in all its forms, but simply an opponent of one version of justificationalism, namely, foundationalism. More specifically, there are motifs in Wittgenstein's later writings that suggest that he simply abandoned one version of justificationalism, namely foundationalism, in favor of another version, namely holism or coherentism. To the extent that this is true, Wittgenstein is not a neo-Pyrrhonian rejecting all forms of philosophical justification but is, instead, a partisan of a particular version of justificationalism, one with its roots in German idealism. For want of a better name, I shall call this second strain the non-Pyrrhonian side of Wittgenstein's later philosophy.

It seems clear to me now, as it has not before, that both neo-Pyrrhonian and non-Pyrrhonian commitments play important roles in Wittgenstein's later philosophy. I am now inclined to read Wittgenstein's later writings as a constant battle between these two aspects of his thought. In these writings we find a strong drive to *replace* his earlier foundationalist theory with a nonfoundationalist theory—a drive constantly being curbed, sometimes successfully, sometimes not, by an opposing drive to eliminate philosophy altogether. This is a conflict between doing philosophy and doing away with it.

I shall examine the non-Pyrrhonian side of Wittgenstein's later philosophy working on the assumption that in a philosophical position certain concepts are assigned a privileged status. This privileged status has at least two sides. First, these concepts are assigned a foundational role within the philosophical position. Here we might speak of concept foundationalism. Many philosophers who are

antifoundationalist with respect to basic propositions are foundationalist with respect to what they take to be basic concepts. Second, these concepts are typically *exempted from criticism*—they are treated with respect. A philosophical position is characterized, in part at least, by the set (system, complex) of the concepts it assigns this privileged status.

For the neo-Pyrrhonian there are no privileged concepts. The fundamental idea I am pursuing is found in Wittgenstein's later writings in remarks of the following kind:

> *PI,* 97. We are under the illusion that what is peculiar, profound, essential, in our investigation, resides in its trying to grasp the incomparable essence of language. That is, the order existing between the concepts of proposition, word, proof, truth, experience, and so on. The order is a *super-order* between—so to speak—*super-concepts*. Whereas, of course, if the words "language," "experience," "world," have a use, it must be as humble a one as that of the words "table," "lamp," "door."

My notion of a privileged concept is a counterpart of Wittgenstein's notion of a superconcept (*ein Überbegrift*). This, then, is another aspect of the neo-Pyrrhonian position: Philosophy depends for its existence on privileging certain concepts, but concepts will not bear this philosophical burden.

But even if Wittgenstein was correct in targeting superconcepts for special criticism, it seems he was not always immune to their charms. In particular, in rejecting the constellation of superconcepts underlying the *Tractatus,* he sometimes seems to fall into the error of replacing them with a reverse system of superconcepts. That move—to replace the superconcepts of the *Tractatus* with their polar opposites—constitutes the non-Pyrrhonian side of Wittgenstein's philosophy. For many admirers of Wittgenstein, it is precisely this non-Pyrrhonian side of his philosophy that they find attractive. From the point of view of this study, this non-Pyrrhonian element in Wittgenstein's late philosophy represents a return to darkness.

Turning Things Around

Sketched in broad terms, Wittgenstein's *Tractatus* is a typically Cartesian text in favoring:

atomism over holism
privacy over publicity
thinking over doing

With some historical justification, I will say that a position that embodies these preferences is committed to the Cartesian complex.

Wittgenstein's atomism comes out with respect to language in his claim that elementary propositions—the basic units of representation in the *Tractatus*—are independent of one another (5.134–5.135). It comes out on the side of the world with the claim that the causal nexus does not exist (5.136–5.1361). Although this is more difficult to get correct, it also seems clear that, in the *Tractatus*, language is essentially a private affair (5.62). Finally, the Tractarian agent does nothing *in* the world, for the world, as Wittgenstein tells us, is "independent of my will" (6.373). There is no way of altering the world, though I may be able to alter my attitude toward the world in various ways. The world of the happy man, we are told, is different from the world of the unhappy man, though not because happiness or unhappiness has any effect on the content—the facts— of the world (6.43).

Given this broad sketch, one is immediately struck with the disparity between the Tractarian standpoint and our common ways of viewing the world. The *Tractatus* gains whatever plausibility it commands only through being unfolded from seemingly reasonable assumptions about what language and the world must be like in order for language to represent (or picture) the world. Once those seemingly reasonable assumptions are rejected, the position loses its force.

In his later writings, Wittgenstein did not reject the Tractarian position in toto. For example, he never gave up the idea that logical constants—as well as numerals—do not stand for (or represent) objects. The constructivist views of mathematics adumbrated in the *Tractatus* became a central feature of his later philosophy of mathematics. I do not think he ever fully gave up the idea that there are certain things that cannot be said but only shown. There are other important carry-overs as well. What he did reject, root and branch, were the three doctrines that make up what I have called the Cartesian complex. Our question here is whether in doing so he simply reversed priorities and thereby accepted a new system of super (or privileged) concepts. This is a complex question, but the answer to it, I think, is yes, at least sometimes. To the extent that this is correct, Wittgenstein's later philosophy is not adequate to its own standards, since these standards apply equally to superconcepts of every kind. Thus, in attacking Wittgenstein's later use of a new set of superconcepts, I will be using one side of his position, the neo-Pyrrhonian, against the second side, the non-Pyrrhonian.

Holism

In the *Philosophical Investigations*, Wittgenstein says the following things:

> *PI*, 47. "Simple" means: not composite. And here the point is: in what sense "composite"? It makes no sense at all to speak absolutely of the "simple parts of a chair." . . .
>
> If I tell someone without any further explanation: "What I see before me is now composite," he will have the right to ask: "What do you mean by 'composite'? For there are all sorts of things that can mean!"—The question "Is what you see composite?" makes good sense if it is already established what kind of complexity—that is, which particular use of the word—is in question.

Notice that Wittgenstein treats simplicity and complexity on a par. It makes no more sense to speak of something being absolutely complex than it does to speak of something being absolutely simple. I wish to suggest that the same sort of correlative relationship exists between parts and wholes. Understanding what a whole φ is involves understanding what will count as a part—and this will vary with context. If that is correct, the same error occurs in speaking about absolute wholes and absolute parts as in speaking about absolute simplicity and absolute complexity. The error of the *Tractatus*, as Wittgenstein came to see it, was to introduce, and then to rely on, the notion of something being absolutely simple. Our question now is whether Wittgenstein, in his later writings, traded in the superconcepts of atomism for the superconcepts of holism.

The charge is not easily made good against the *Philosophical Investigations*, where Wittgenstein usually restricts himself to the modest contextualist claim that we can understand the meaning of a term only by seeing its use within the language game where it finds employment. A healthy—even radical—pluralism dominates the *Philosophical Investigations.* It is a central theme of the opening sections of the *Philosophical Investigations* that language consists of a variety of techniques, and philosophers (including the author of the *Tractatus*) have befuddled themselves by trying to find some common essence running through all these various uses of language. Here we might speak of a pluralistic contextualism, that is, a contextualism that does not involve a holistic commitment to the existence of a single overarching context. The *Philosophical Investigations*, as I read this work, presents a pluralistic rather than a holistic contextualism.

I also think pluralistic contextualism is a more plausible view than holistic contextualism because pluralism squares with the undoubted fact that human beings can radically disagree with each other in certain areas, yet still understand and agree with each other over a very wide range of other topics. It is even possible for such mutual understanding to exist between people with fundamentally different worldviews. That my tax accountant subscribes to a religion that I find wholly irrational does not lead me to question his ability to prepare my tax return. Indeed, our disagreement on religion, though total, usually touches little else. The claim "If he believes *that*, he will believe anything" is almost always false. So is its twin, "If he doubted that, there would be nothing left that he would not doubt." There are occasions when we are inclined to engage in such hyperbole for rhetorical effect. Philosophical holism takes these rhetorical excesses literally and attempts to base a philosophy on them.

When we turn from the *Philosophical Investigations* to Wittgenstein's very late reflections found in *On Certainty*, we do seem to find an appeal to a kind of holism that Wittgenstein, by his own principles, should have rejected. In fact, many passages in *On Certainty* are at least compatible with the view that language consists of a plurality of more or less autonomous language games. Consider the claim:

> *OC*, 225. What I hold fast to is not *one* proposition but a nest of propositions.

Nothing here suggests a commitment to a supernest (*ein Übernest*). The same modest reading is possible for the following important passage:

> *OC*, 105. All testing, all confirmation and disconfirmation of a hypothesis takes place already within a system.

From the fact that all testing takes place within a system, we cannot infer that there is some system in which all testing takes place.[1]

At the same time, *On Certainty* contains many passages expressing a holistic version of contextualism that is as robust as one can imagine.

> *OC*, 410. Our knowledge forms an enormous system. And only within this system has a particular bit the value we give it.

> *OC*, 411. If I say "*we assume* that the earth has existed for many years past" (or something similar), then of course it sounds strange that we should *assume* such a thing. But in the entire system of our language-games it belongs to the foundations.

The reference in 411 to "the entire system of our language-games" represents, to my mind, a radical departure from the pluralistic views of the *Philosophical Investigations*. It also embodies the move that converts ordinary concepts (i.e., a system or a whole) into super or privileged concepts. In *On Certainty*, systems (totalities, wholes) become super or privileged concepts in the two senses introduced above: They are assigned a philosophical burden, and they are exempted from examination.

Wittgenstein is hardly alone in making such a move to holistic versions of contextualism. The move, indeed, has reached epidemic proportions. I'll cite just a few examples from influential philosophers:

> In an obvious way this structure of interconnected sentences is a single connected fabric including all sciences, and indeed every-thing we say about the world; for the logical truths at least, and no doubt many more commonplace sentences too, are germane to all topics and thus provide connections. (Quine, 1960, 15–16)

That's Quine; here's Rorty:

> The holist line of argument says that we shall never be able to avoid the "hermeneutic circle"—the fact that we cannot understand the parts of a strange culture, practice, theory, language, or whatever, unless we know something about how the whole thing works, whereas we cannot get a grasp on how the whole works until we have some understanding of its parts. (Rorty, 1979, 319)[2]

Similar holistic passages occur in Hilary Putnam's description of the *internalist perspective* that he champions in *Reason, Truth, and History*. Those who disagree with him are said to hold an externalist, or "God's eye view of the world" (49).[3]

Against these united voices, I wish to suggest that these non-contextualized appeals to wholes or totalities make no sense. I have no idea what would count as a whole culture or a whole language. I have just the same troubles with Davidson's seemingly more mod-est talk about the totality of a person's belief. This strikes me as a pseudo-totality as well, since I have no idea how beliefs are to be individuated, then toted up. Davidson comes close to saying the same thing:

> There is probably no useful way to count beliefs; and so no clear meaning to the idea that most of a person's beliefs are true. (David-son, 1986, 308)

But reference to no *useful* way to count beliefs understates the diffi-culty. No one is asking that a number be assigned to a person's beliefs before we countenance talk about *all* that person's beliefs.

The difficulty is that this totality is so underspecified that it is hard to see how even the notions of *more* and *less* can gain purchase on it.

Here we might turn dialectical and say that holists are not sufficiently holistic about the concept of a whole. But this is the wrong tack to take, since it invites a response of the following kind: "We can understand what a whole language is only after we understand language as a whole, and since we have barely begun to understand our language as a whole, our concept of a whole language is itself fragmentary and incomplete." Holistic theories are characteristically sealed from criticism by promissory notes of this kind.

But philosophers who hearken to Wittgenstein's call will not argue in this dialectical fashion; instead, they will attempt "to bring words back from their metaphysical to their everyday use." This will involve looking at a variety of natural contexts in which words like "all," "whole," and "total" are actually used. It is useful, for example, to reflect on the differences between eating a whole pie, a whole meal, and a whole chicken. We can also compare a scientist's talk about the physical universe with philosophers' talk about their favorite totalities. Scientists ask quite definite questions about the total universe: how big is it, how old is it, how much matter does it contain? And so on. Through reflecting on these and other examples we gain an understanding of the constraints that allow us to talk usefully and intelligibly about wholes, totalities, parts, and so forth. If we do this, we will be struck by the absence of these standard constraints in the writings of holistic philosophers—including Wittgenstein in his holistic moods.

Publicity

The second feature of the Cartesian complex noted above is that it assigns a privileged status to the individual's private self-consciousness. Indeed, the doctrine that philosophizing should begin or, perhaps, has no other choice but to begin, with the deliverances of self-consciousness is virtually definitive of what we call modern philosophy. It is precisely this move to the subject that gives Descartes the distinction of being the father of modern philosophy. Though there are some difficult contrary texts (perhaps 5.64), in this regard, the *Tractatus* seems to stand squarely in the Cartesian tradition. This comes out in passages of the following kind:

5.62. The world is *my* world.

5.63. I am my world. (The microcosm.)

6.631. At death the world does not alter, but comes to an end.

It is hard not to take these passages as solipsistic, and, indeed, Wittgenstein describes them this way himself, though he goes on to say that, properly understood, solipsism and realism amount to the same thing (5.64).

It is not to my present purposes to try to unravel the Tractarian identification of solipsism with pure realism. In the end it seems little more than a bluff. Certainly Wittgenstein did not return to this claim in his later writings. This is not to say that Wittgenstein forgot about solipsism; indeed, it is not an exaggeration to say that in his later writings Wittgenstein was *obsessed* with this problem. For Wittgenstein, the only way to avoid solipsism is to reject the conceptual confusions that generate it. This is to be done, Wittgenstein thought, by showing that "subjective utterances" or "statements about the mental" do not function in the way that philosophers have commonly supposed they do. In particular, they are not reports about private mental objects or happenings accessible only to the person who experiences such objects or happenings.

Here I do not wish to examine Wittgenstein's various positive attempts to give a correct account of the way in which first-person psychological utterances function (e.g., the no-ownership account in the *Blue Book*, or the expression theory of the *Philosophical Investigations* and *Zettel*). Nor do I wish to examine *all* the moves that occur in that stretch of text in the *Philosophical Investigations* that is said to contain the private-language argument. I have argued elsewhere that the fundamental move in the private-language argument involves the use of a skeptical paradox concerning rule-following—an interpretation later taken up by Saul Kripke.[4] Briefly: However a rule is applied, it is always possible to find an interpretation of the rule showing that the application is correct, and equally possible to find an interpretation showing that the application is incorrect. If that is so, then rule-following seems a sham. Given this as the background, Wittgenstein then rejects the possibility of there being a private language on the grounds that such a language could not avoid this paradoxical result. In *Wittgenstein*, my own assessment of this argument is that it can be used to show the *contingent* impossibility for creatures such as ourselves to possess a private language. I further concluded—against Wittgenstein and many of his followers—that Wittgenstein's argument could sustain no stronger conclusion.[5] Here I am in fundamental disagreement with Wittgenstein, for beyond what I have called the contingent impossibility of a private language, he seemed to hold that a private language is a conceptual—one might say, transcendental—impossibility. In the second edition of

Wittgenstein, I called the argument intended to show this the public check argument. It is of interest here because it seems to present a clear case of Wittgenstein privileging the notion of publicity over privacy and, in the process, treating publicity as a superconcept.

The public check argument can be stated simply. We are to consider the possibility that a person (*NN*) commands a language that is private in the sense that no other person can tell, even under the best circumstances, whether the terms in this language are being applied correctly or not. If anyone could know that they are being applied correctly, only *NN* could. Setting aside the serious problem of how *NN* could acquire such a language, what rules out its possibility? Wittgenstein's argument runs as follows: If *NN* is speaking a language, then the use of the expressions in his private language must be rule-governed. If the use of these expressions is rule-governed, there exists a distinction between applying the rule correctly and applying it incorrectly. But *NN's* language will not support this distinction, for his only standard for the correct application of an expression is that it seems to him to be correct. That, however, empties the notion of following a rule of content. Such rules are nonrules and a language governed by them is a nonlanguage. So a language that is private in the specified way turns out to be no language at all.

We can next notice how *NN's* situation is improved when he employs a language where the correctness or incorrectness of the use of the expressions in his language *is* capable of a public check. It then becomes possible to draw the distinction between something seeming to *NN* to be correct and its actually being correct. However correct an application may seem to *NN*, it will be actually correct only if it is in conformity with public rules. It is in this way that content is given to the distinction between *NN's* thinking himself to be following a rule and *NN's* actually following it. For Wittgenstein, society not only gives *NN* his language, but also makes it possible for him to have something that counts as a language.

Laid out in this bare-boned way, the public check argument invites an obvious question: How is the distinction between following a rule and only seeming to follow one sustained at the public level? What serves as an independent check there? Wittgenstein's answer seems to be that nothing does (or could) serve this function. He further holds that nothing has to. Our public language and our public practices in general *take care of themselves.* It seems, then, that our public language has been exempted from the task of meeting the success condition that led Wittgenstein to declare a private language impossible. This move appears in passages of the following kind:

For what would *this* mean: "Even though everyone believed that
twice two was five it would still be four"?—For what would it be
like for everybody to believe that?—Well, I could imagine, for
instance, that people had a different calculus, or a technique which
we should not call "calculating." But would it be wrong? (Is a coro-
nation wrong? To beings different from ourselves it might look
extremely odd.)

   Of course, in one sense mathematics is a branch of knowl-
edge,—but still it is also an *activity*. And "false moves" can only
exist as the exception. For if what we now call by that name
became the rule, the game in which there were false moves would
have been abrogated. (*PI*, pp. 226–27)

Richard Rorty captures this—to my mind unfortunate—side of Witt-
genstein's later philosophy in these words. "Explaining rationality
and epistemic authority by reference to what society lets us say,
rather than the latter by the former, is the essence of what I shall call
'epistemological behaviorism,' an attitude common to Dewey and
Wittgenstein" (Rorty, 1979, 174). Rorty continues by relating this
appeal to society (the public) to holism: "This sort of behaviorism
can best be seen as a species of holism—but one which requires no
idealist metaphysical underpinning. It claims that if we understand
the rules of a language-game, we understand all that there is to
understand about why moves in the language-game are made" (174).
The telltale expressions in this passage are "rationality" and "epis-
temic authority"—both terms of epistemic approval. We are here
being shown the basis—the ground—for such epistemic approval. It
does not lie (ultimately) in private self-certainty but lies, instead, in
the constraints of public practice. It is in this way that a set of super-
concepts can be generated out of a zealous attack on a polar opposite
set of superconcepts.

   In *Wittgenstein* I argued that the public check argument fails
because it relies on an arbitrarily selective use of a general skeptical
argument. To this I shall now add that the *deference* shown to pub-
licity in this regard, together with the role assigned it to block skep-
ticism, are clear indications that the notion of publicity has been ele-
vated to the status of a superconcept.

   But even if the text does show a strong non-Pyrrhonian moment
with respect to publicity, the neo-Pyrrhonian voice is not completely
silent, even on the primacy of the public. The following passage
occurs near the end of Wittgenstein's *Remarks on the Foundations
of Mathematics:*

On the other hand it is not clear that the general agreement of people doing calculations is a characteristic mark of all that is called "calculating." I could imagine that people who had learned to calculate might in particular circumstances, say under the influence of opium, begin to calculate differently from one another, and might make use of these calculations; and they were said not to be calculating at all and to be deranged—but that their calculations were accepted as a reasonable procedure.

But must they not at least be trained to do the same calculations? Doesn't *this* belong essentially to the concept of calculating? I believe that we could imagine deviations here too. (p. 187)

Here Wittgenstein invokes his anti-essentialism to temper his own insistence that the command of a language, for example of the language of calculating, must be grounded in training in public practices. In fact, that is the kind of creatures we are, but it is imaginable that we could have been otherwise.

My claim, then, is not that Wittgenstein, in his later writings, simply traded in the superconcept of privacy for the superconcept of publicity. In some places he does seem to do just this, but in others, for example in the passage just cited, he catches himself doing this and corrects himself. This, I think, is the clearest example of the struggle between non-Pyrrhonian themes and neo-Pyrrhonian themes in Wittgenstein's later writings.

Action

John Dewey complained that Cartesian philosophers cut themselves off from the rise of modern experimental science by sharply separating thought from action and then giving the nod to thought. Some passages in the *Tractatus* seem to place Wittgenstein in this Cartesian tradition:

6.43. If good and bad willing changes the world, it can only change the limits of the world, not the facts.

If willing—whether good or bad—cannot change the facts of the world, then we never *do* anything *in* the world either. Plausible or not, this is one of the central doctrines of the *Tractatus*. The rejection of this standpoint is one of the primary doctrines in Wittgenstein's later philosophy.

What we might call the primacy of acting (or doing) appears in section 1 of the *Philosophical Investigations*. Wittgenstein imagines

someone handing a shopkeeper a slip saying "five red apples" and the shopkeeper, after going through various procedures (counting, matching samples, etc.), handing over five red apples. Wittgenstein then has an interlocutor ask:

> "But how does he know where and how he is to look up the word 'red' and what he is to do with the word 'five'?"

Wittgenstein replies:

> Well, I assume that he *acts* as I have described. Explanation comes to an end somewhere.

Similar passages occur in *On Certainty:*

> *OC*, 109. "An empirical proposition can be *tested*" (we say). But how? and through what?

> *OC*, 110. What *counts* as its test?—"But is this an adequate test: And, if so, must it not be recognizable as such in logic?"—As if giving grounds did not come to an end sometime. But the end is not an ungrounded presupposition: it is an ungrounded way of acting.

Then later in *On Certainty:*

> *OC*, 204. Giving grounds, however, justifying the evidence, comes to an end; but the end is not certain propositions' striking us immediately as true, i.e., it is not a kind of *seeing* on our part; it is our *acting*, which lies at the bottom [*Grund*] of the language-game.

What are we to make of this commitment to *acting* over *seeing?* I think it can be taken in two ways, reflecting the two competing forces in Wittgenstein's later thought: as descriptive in content and philosophically deflationary in intent, or as normative in content and philosophically constructive in intent. Some passages suggest one reading, some the other; many can be taken either way.

We can look first at passages that exemplify the purely descriptive motif:

> *OC*, 148. Why do I not satisfy myself that I have two feet when I want to get up from a chair? There is no why. I simply don't. This is how I act.

What is the force of the claim "There is no why"? I take it to be wholly negative. To see this, we can examine Wittgenstein's response to the argument that we are, after all, perfectly *justified* in not checking whether we have two feet before we get up from a chair. The supposed justification could take various forms. One might be empirical: We have, it could be argued, overwhelmingly strong obser-

vational evidence for the existence of our feet in the past (we have observed them daily), and we also have strong inductive understanding of the conditions under which feet get detached from the body (they do not, for example, fall off unnoticed). The reason, then, that we do not check to see if we have two feet when we want to get up from a chair is that we have unassailable empirical evidence that we do have two feet. A second justification would be more pragmatic: Checking on whether one has feet when there is no call to do so would be a waste of our intellectual resources. It doesn't pay to check on such things. Though different, both the empirical and the pragmatic justification have this in common: They both hold that we believe, for example, that we have feet on the basis of good reasons.

It is a central theme of the neo-Pyrrhonian side of Wittgenstein's later philosophy to reject such justificationalist commitments. The following passages bear on the empirical style of justification:

> *PI,* 472. The character of the belief in the uniformity of nature can perhaps be seen most clearly in the case in which we fear what we expect. Nothing could induce me to put my hand into a flame— although after all it is *only in the past* that I have burnt myself.

> *PI,* 473. The belief that fire will burn me is of the same kind as the fear that it will burn me.

In this primitive case, neither the fear nor the belief is the product of ratiocination.

On the other hand, both the belief and the fear can be made objects of ratiocination, and then something can be said in their behalf. This can lead us to the natural but mistaken assumption that the reasons we now give in behalf of these beliefs are their grounds. Wittgenstein makes the point this way:

> *PI,* 479. The question "On what grounds do you believe this?" might mean: "From what are you now deducing it (have you just deduced it)?" But it might also mean: "What grounds can you produce for this assumption on thinking it over?"

If asked, I could probably produce a very strong empirical argument in behalf of the existence of my (unobserved) feet or in behalf of the claim that fire will burn me. As Wittgenstein puts it, "A hundred reasons present themselves, each drowning the voice of the others" (*PI,* 478). That, however, does not show that my beliefs were *based* on any such reasons. Perhaps some of these beliefs are based on reasoning of this kind. The supposition that they *must be* is, for Wittgenstein, merely a rationalist prejudice.

Turning next to pragmatic justification, in a number of places Wittgenstein attempts to distance himself from pragmatism.[6] In one passage in *On Certainty*, he pairs his criticism of empirical foundationalism with an aside attacking pragmatism:

> *OC*, 131. No, experience is not the ground for our game of judging. Nor is its outstanding success.

In a similar vein, the *Philosophical Investigations* contains the following remarkable passage:

> *PI*, 467. Does man think, then, because he has found that thinking pays?—Because he thinks it advantageous to think?
> (Does he bring his children up because he has found it pays?)

Wittgenstein then gives the discussion an unexpected turn by noting that "*sometimes* [we do] think because it has been found to pay" (*PI*, 470). For example, "There are fewer boiler explosions than formerly, now that we no longer go by feeling in deciding the thickness of the walls, but make such-and-such calculations instead" (*PI*, 469). In particular contexts it makes good sense to say that it pays to think (calculate, double-check, etc.). The mistake is to take these particular contexts as representative rather than special.

The references to acting (doing, etc.) so far examined seem to be in line with the neo-Pyrrhonian side of Wittgenstein's later philosophy. The difficult question I wish to raise now is whether Wittgenstein, sometimes at least, departs from neo-Pyrrhonism by treating acting (doing) as a new kind of foundation, or as an alternative foundation, to the rational foundations favored by philosophers. Is the reference to action descriptive of how language games are employed, or is the appeal to action in some way supposed to provide a ground in the sense of a vindication for these language games? To the extent that Wittgenstein has adopted the second course, it seems that he has simply reversed the polarities of thinking and doing and thus turned doing into a superconcept.

Unfortunately, there are passages that suggest that Wittgenstein has made this move—or succumbed to this temptation. I have already cited one such passage from *On Certainty*, for example:

> *OC*, 204. It is not a kind of *seeing* on our part; it is our *acting*, which lies at the bottom [*Grund*] of the language-game.

Similar remarks occur elsewhere in *On Certainty*.

> *OC*, 342. That is to say, it belongs to the logic of our scientific investigations that certain things are *in deed* [*in der Tat*] not doubted.

This passage is echoed at *OC*, 402, where Wittgenstein cites Goethe's Faust:

> . . . und schreib getrost
> "Im Anfang war die Tat."

"In the beginning was the *deed*"—not, that is, the *word*. All this suggests that, where philosophers have been looking for ultimate propositions to ground (i.e., vindicate) beliefs, they should, instead, have turned to ungrounded acting. This suggests that our language games *do* have a foundation (a bedrock) that secures them against skeptical assault; the foundation, however, is not intellectual, but practical.

The action-passages cited so far all come from *On Certainty*. There is, I think, nothing quite like them in the *Philosophical Investigations*. But even if the *Philosophical Investigations*, unlike *On Certainty*, is free of blatant instances of turning acting and doing into superconcepts, there are still instances where Wittgenstein's emphasis on action over thought leads him to distort the meaning of common words for philosophical purposes. "Acting," "doing," and so forth, are, first, perfectly ordinary words doing ordinary jobs. Second, in their ordinary use they do not stand in the contrast to thinking that Wittgenstein is attempting to exploit. If I inquire into someone's actions or doings, I will be told about pieces of his *intentional* behavior. It hasn't proved easy to give an adequate account of *action*, but in the standard cases, at least, an action is a kind of behavior guided by thought. Thus *acting* does not stand in contrast to *seeing* (comprehending, thinking) in the way that Wittgenstein's discussion demands; action typically is an expression of thought.

One mark that an ordinary concept has been elevated to a superconcept is that the person using that concept is led to make claims that, on the common understanding, are blatantly false. This occurs, for example, in the context in which Wittgenstein is attempting to deal with the skeptical doubts he has raised concerning rule-following. Under the pressure of this problem he produces the following philosophical hyperbole:

> *PI*, 219. When I obey a rule, I do not choose.
>   I obey the rule *blindly*.

Here Wittgenstein seems to have hatched a gratuitous paradox, for to say that someone does something blindly, in the ordinary way of understanding this metaphor, is to say that the action is performed without rule—rhyme or reason. And Wittgenstein's contrast between following a rule as a matter of *choice* and doing this *blindly* is similarly peculiar. Sometimes when I follow a rule, I am very careful to

watch my step. It seems, then, that even in the *Philosophical Investigations* Wittgenstein shows some tendencies to turn "doing" and "acting" into philosophical superlatives—into superconcepts.

In the opening, say, 137 sections of the *Philosophical Investigations*, Wittgenstein exemplifies a profoundly new way of doing philosophy through a sustained critique of the underlying viewpoint of the *Tractatus*. It is important to see that this critique of the *Tractatus* is not narrowly aimed at its particular shortcomings. The critique is intended to exemplify a method for dealing with any attempt at philosophical justification. It is part of a *general* critique of philosophizing. The aim, then, is not to replace the Tractarian set of concepts with another set that will do the job better. This becomes clear in any number of passages.

> *PI*, 118. Where does our investigation get its importance from, since it seems only to destroy everything interesting, that is, all that is great and important? . . . What we are destroying is nothing but houses of cards and we are clearing up the ground on which they stand.

> *PI*, 124. Philosophy may in no way interfere with the actual use of language; it can in the end only describe it.
>   For it cannot give it any foundation either.
>   It leaves everything as it is.

> *PI*, 133. The clarity that we are aiming at is indeed *complete* clarity. But that simply means that the philosophical problem should *completely* disappear.

In these passages and many others, we hear the voice of the neo-Pyrrhonian Wittgenstein.

On the other side, if the textual analysis given above is correct, there is a second voice in Wittgenstein's later writings that speaks in opposition to the first. The situation is not like that found in the *Tractatus*, where a fully coordinated system of superconcepts is presented in an effort to solve a set of philosophical problems. In the later writings there are what we might call outbreaks or eruptions of (traditional, transcendental) philosophizing that evade the critical eye that should have detected them. This is the Wittgenstein who, in complex and indirect ways, attempted to *replace* the package of atomism, privacy, and thought with the package of holism, publicity, and action. What we learn from the neo-Pyrrhonian Wittgenstein, over against the non-Pyrrhonian Wittgenstein, is that the value of a coin is not increased by turning it over.[7]

Notes

1. It would not help to suggest that all testing takes place within the system of all systems, unless it is shown that the totality of systems is itself a system in the relevant sense. It is far from obvious that this is so.

2. Since Rorty here equally insists that the whole must be understood through its parts, it may seem unfair to treat him as a simple holist. But his reference to whole cultures and languages shows that he is willing to speak about and argue from indeterminate wholes. The indeterminacy of these wholes is not mitigated by references to parts that are themselves indeterminate.

Incidentally, I find the problem of the hermeneutic circle—if it is a problem at all—completely unbeatable. Imagine playing the following game. One at a time you are supposed to place two checkers (one red, one black) on the board following the rule that you may only place a red checker to the right of a black checker, and a black checker to the left of a red checker. It would be easy enough to see if a particular arrangement satisfied these conditions, but there would be no way of producing an arrangement satisfying them, since there is no way of getting the game started.

3. Putnam, 1981.

4. For the relationship between my reading of the private-language argument and Kripke's, see Fogelin, 1987b, 241–46 n. 10.

5. See Fogelin, 1987b, chapter 12.

6. For similar reasons he attempts to distance himself from behaviorism in psychology and finitism in mathematics.

> Finitism and behaviorism are quite similar trends. Both say, but surely, all we have here is . . . Both deny the existence of something, both with a view of escaping from a confusion. (*RFM*, II, 18)

7. Crispin Wright also detects two voices in Wittgenstein's later philosophy. First there is the Wittgenstein who makes pronouncements about philosophy, telling us that, in the end, philosophical problems should simply disappear. This is the neo-Pyrrhonian Wittgenstein. But then, as Wright shrewdly notes:

> It is difficult to reconcile Wittgenstein's pronouncements about the kind of thing which he thinks he ought to be doing with what he actually seems to do. Not that his actual treatment of the particular issues seems flatly inconsistent with his general methodological ideas. Rather, we can put the would-be interpreter's difficulty like this: it is doubtful how anyone who read only a bowdlerised edition of the *Investigations*, from which all reference to philosophical method and the nature and place of philosophy had been removed, would be able to arrive at the conclusion that the author viewed these matters in just the way in which Wittgenstein professes to do. (Wright, 1980, 262)

I don't agree with this. I think the opening part of the *Philosophical Investigations* exemplifies the philosophical method that Wittgenstein pronouncements on philosophy describe. What is right in Wright's reading is that there is also much in Wittgenstein's later writings that is at least alien to, if not incompatible with, this neo-Pyrrhonian standpoint. At bottom, I do not think Wright and I disagree on this point; we simply admire different, seemingly competing, aspects of Wittgenstein's philosophy.

References

Almeder, Robert. 1974. "Truth and Evidence." *Philosophical Quarterly* 24:365–68.

_____. 1976. "On Seeing the Truth: A Reply." *Philosophical Quarterly* 26:163–65.

Annas, Julia, and Jonathan Barnes. 1985. *The Modes of Scepticism: Ancient Texts and Modern Interpretations*. Cambridge: Cambridge University Press.

Armstrong, D. M. 1961. *Perception and the Physical World*. London: Routledge and Kegan Paul.

Audi, Robert. 1988. *Belief, Justification, and Knowledge*. Belmont, Calif.: Wadsworth.

Austin, J. L. 1979. *Philosophical Papers*. 3d ed. Oxford: Oxford University Press.

Barnes, Jonathan. 1982. "The Beliefs of a Pyrrhonist." *Proceedings of the Cambridge Philological Society*, eds. E. J. Kenny and M. M. MacKenzie, 2–29.

Bayle, Pierre. 1734. *A general dictionary, historical and critical: in which a new and accurate translation of that of the celebrated Mr. Bayle, with the corrections and observations printed in the late edition at Paris, is included: and interspersed with several thousand lives never before published: the whole containing the history of the most illustrious persons of all ages and nations particularly those of Great Britain and Ireland, distinguished by their rank, actions, learning and other accomplishments: with reflections on such passages of Mr. Bayle, as seem to savour scepticism and the Manichee system*. London: Printed by James Bettenham.

Beiser, Frederick C. 1987. *The Fate of Reason: German Philosophy from Kant to Fichte*. Cambridge, Mass.: Harvard University Press.

Blanshard, Brand. 1939. *The Nature of Thought*. 2 vols. New York: Macmillan.

BonJour, Laurence. 1976. "The Coherence Theory of Empirical Knowledge." *Philosophical Studies* 30:281–312.

_____. 1985. *The Structure of Empirical Knowledge*. Cambridge, Mass.: Harvard University Press.

Bosanquet, Bernard. 1920. *Implication and Linear Inference*. London: Macmillan.

Burge, Tyler. 1979. "Individualism and the Mental." *Midwest Studies in Philosophy*, eds. Theodore Uehling, Howard Wettstein, and Peter French. Minneapolis: University of Minnesota Press.

_____. 1986. "Cartesian Error and the Objectivity of Perception." *Subject, Thought, and Context*, eds. Philip Pettit and John MacDowell. Oxford: Clarendon Press.

Burnyeat, Miles. 1980. "Can the Sceptic Live His Scepticism?" *Doubt and Dogmatism*, eds. M. Schofield, M. F. Burnyeat, and J. Barnes. Oxford: Clarendon Press.

_____, ed. 1983. *The Skeptical Tradition*. Berkeley: University of California Press.

_____. 1984. "The Sceptic in His Place and Time." *Philosophy in History*, eds. Richard Rorty, J. B. Schneewind, and Quentin Skinner. Cambridge: Cambridge University Press.

Cherniak, Christopher. 1986. *Minimal Rationality*. Cambridge, Mass.: MIT Press.

Chisholm, Roderick. 1957. *Perceiving: A Philosophical Study*. Ithaca: Cornell University Press.

_____. 1977. *Theory of Knowledge*. 2d ed. Englewood Cliffs, N.J.: Prentice Hall.

_____. 1978. "Comments and Replies." *Philosophia* 7, nos. 3–4: 597–636.

_____. 1982a. *First Person: A Study in Reference and Intentionality*. Minneapolis: University of Minnesota Press.

_____. 1982b. *The Foundations of Knowing*. Minneapolis: University of Minnesota Press.

_____. 1986. "Knowledge as Justified True Belief." *Empirical Knowledge*, ed. Paul K. Moser, 246–51. Totowa, N.J.: Littlefield.

_____. 1989. *Theory of Knowledge*. 3d ed. Englewood Cliffs, N.J.: Prentice Hall.

Clarke, Thompson. 1972. "The Legacy of Skepticism." *Journal of Philosophy* 69:764–69.

Clifford, W. K. 1879. "The Ethics of Belief." *Lectures and Essays*. London: Macmillan.

Davidson, Donald. 1969. "True to the Facts." *Journal of Philosophy* 66:748–64.

_____. 1973–74. "On the Very Idea of a Conceptual Scheme." *Proceedings and Addresses of the American Philosophical Association* 67:5–20.

_____. 1980. *Essays on Actions and Events*. New York: Oxford University Press.

_____. 1986. "A Coherence Theory of Truth and Knowledge." *Truth and Interpretation*, ed. Ernest LePore, 307–19. Oxford: Blackwell.

_____. 1991. "Epistemology Externalized." *Dialectica* 45, nos. 2–3: 191–202.

_____. 1992. "Subjective, Intersubjective, Objective." Manuscript.

Dretske, Fred. 1970. "Epistemic Operators." *Journal of Philosophy* 67, no. 24: 1007–23.

_____. 1971a. "Conclusive Reasons." *Australasian Journal of Philosophy* 49:1–22.

_____. 1971b. *Seeing and Knowing.* Chicago: University of Chicago Press.

_____. 1981. *Knowledge and the Flow of Information.* Cambridge, Mass.: MIT Press.

Ducasse, C. J. 1951. *Nature, Mind, and Death.* LaSalle, Ill.: Open Court.

Dummett, Michael. 1959. "Truth." *Proceedings of the Aristotelian Society* 59:141–62.

Empiricus, Sextus. 1961–71. *Sextus Empiricus.* Trans. R. G. Bury. Loeb Classical Library. 4 vols. London: William Heinemann.

Etchemendy, John. 1990. *The Concept of Logical Consequence.* Cambridge, Mass.: Harvard University Press.

Feldman, Richard. 1974. "An Alleged Defect in Gettier Counterexamples." *Australasian Journal of Philosophy* 52, no. 1: 68–69.

Firth, Roderick. 1967. "The Anatomy of Certainty." *Philosophical Review* 76:3–27.

_____. 1978. "Are Epistemic Concepts Reducible to Ethical Concepts?" *Values and Morals: Essays in Honor of William Frankena, Charles Stevenson, and Richard Brandt,* eds. Alvin I. Goldman and Jaegwon Kim, 215–29. Dordrecht: Reidel.

Fogelin, Robert J. 1967. *Evidence and Meaning.* International Library of Philosophy and Scientific Method. London: Routledge and Kegan Paul.

_____. 1968. "Wittgenstein and Intuitionism." *American Philosophical Quarterly* 5, no. 4: 267–74.

_____. 1976. *Wittgenstein.* 1st ed. London: Routledge and Kegan Paul.

_____. 1981. "Wittgenstein and Classical Skepticism." *International Philosophical Quarterly* 21, no. 1: 3–15.

_____. 1983. "The Tendency of Hume's Scepticism." *The Skeptical Tradition,* ed. Miles Burnyeat, 397–412. Berkeley: University of California Press.

_____. 1985. *Hume's Skepticism in the Treatise of Human Nature.* London: Routledge and Kegan Paul.

_____. 1987a. "Review: *The Modes of Skepticism,* by Julia Annas and Jonathan Barnes." *Canadian Philosophical Review* 7, no. 2: 50–52.

_____. 1987b. *Wittgenstein.* 2d ed. London: Routledge and Kegan Paul.

_____. 1991. "Review: *The Way of Words,* by Paul Grice." *Journal of Philosophy* 88, no. 4: 213–19.

_____. 1992. *Philosophical Interpretations.* New York: Oxford University Press.

_____. 1993. "Review: *Skepticism,* by Christopher Hookway." *Philosophy and Phenomenological Research* 53, no. 1: 215–20.

Fogelin, Robert J., and James H. Moor. 1991. "Lehrer on Incompatible Though Equally Coherent Systems." *Philosophical Studies* 64:229–32.

Frede, Michael. 1984. "The Skeptic's Two Kinds of Assent and the Question of the Possibility of Knowledge." *Philosophy in History,* eds. Richard Rorty, J. B. Schneewind, and Quentin Skinner, 255–78. Cambridge: Cambridge University Press.

_____. 1987. "The Skeptic's Beliefs." *Essays in Ancient Philosophy,* 179–200. Minneapolis: University of Minnesota Press.

Gettier, Edmund L. 1963. "Is Justified True Belief Knowledge?" *Analysis* 26, no. 3: 144–46.

Goldman, Alvin I. 1967. "A Causal Theory of Knowing." *Journal of Philosophy* 64, no. 12: 355–72.

_____. 1976. "Discrimination and Perceptual Knowledge." *Journal of Philosophy* 73, no. 20: 771–91.

_____. 1986. *Epistemology and Cognition.* Cambridge, Mass.: Harvard University Press.

Goodman, Nelson. 1972. *Problems and Projects.* Indianapolis: Bobbs-Merrill.

Grice, H. P. 1961. "The Causal Theory of Perception." *Proceedings of the Aristotelian Society,* supp. vol. 35:121–52.

_____. 1989. *Studies in the Way of Words.* Cambridge, Mass.: Harvard University Press.

Harman, Gilbert. 1965. "Inference to the Best Explanation." *Philosophical Review* 64:88–95.

_____. 1970. "Induction." *Induction, Acceptance, and Rational Belief,* ed. Marshall Swain, 83–99. Dordrecht: Reidel.

_____. 1973. *Thought.* Princeton, N.J.: Princeton University Press.

Hume, David, 1938. *An abstract of a treatise of human nature, 1740, a pamphlet hitherto unknown.* Cambridge: The University Press.

_____. 1975. *Enquiries Concerning the Human Understanding and Concerning the Principles of Morals.* 3d ed. Oxford: Clarendon Press.

_____. 1978. *A Treatise of Human Nature.* 2d ed. Oxford: Oxford University Press.

Klein, Peter D. 1971. "A Proposed Definition of Propositional Knowledge." *Journal of Philosophy* 68, no. 16: 471–82.

_____. 1981. *Certainty: A Refutation of Scepticism.* Minneapolis: University of Minnesota Press.

Kripke, Saul A. 1980. *Naming and Necessity.* Cambridge, Mass.: Harvard University Press.

Kyburg, Henry. 1961. *Probability and the Logic of Rational Belief.* Evanston, Ind.: Northwestern University Press.

Lehrer, Keith. 1965. "Knowledge, Truth, and Evidence." *Analysis* 25, no. 5: 168–75.

_____. 1970. "The Fourth Condition of Knowledge: A Defense." *Review of Metaphysics* 14:122–28.

_____. 1971. "Why Not Scepticism?" *The Philosophical Forum* 2, no. 3: 283–98.

_____. 1974. *Knowledge.* New York: Oxford University Press.

_____. 1990. *Theory of Knowledge*. Dimensions of Philosophy Series. Boulder, Col.: Westview Press.

Lehrer, Keith, and Thomas D. Paxson, Jr. 1969. "Knowledge: Undefeated Justified True Belief." *Journal of Philosophy* 66, no. 8: 225–37.

LePore, Ernest, ed. 1986. *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson*. Oxford: Blackwell.

Lewis, C. I. 1946. *An Analysis of Knowledge and Valuation*. LaSalle, Ill.: Open Court.

Lewis, David K. 1973. *Counterfactuals*. Cambridge, Mass.: Harvard University Press.

Lowy, Catherine. 1978. "Gettier's Notion of Justification." *Mind* 87:105–8.

Luper-Foy, Steven, ed. 1987. *The Possibility of Knowledge*. Totowa, N.J.: Rowan and Littlefield.

Martin, Raymond. 1975. "Empirically Conclusive Reasons." *Philosophical Studies* 28:215–16.

_____. 1983. "Tracking Nozick's Sceptic: A Better Method." *Analysis* 43:3–33.

Mates, Benson. 1981. *Skeptical Essays*. Chicago: University of Chicago Press.

Moore, G. E. 1925. "A Defence of Common Sense." *Contemporary British Philosophy*, 2d series, ed. J. H. Muirhead, 193–223. London: Allen and Unwin.

_____. 1939. "Proof of an External World." *Proceedings of the British Academy* 25.

_____. 1959. *Philosophical Papers*. London: Allen and Unwin.

Moser, Paul K. 1985. *Empirical Justification*. Dordrecht: Reidel.

Naess, Arne. 1969. *Scepticism*. London: Routledge and Kegan Paul.

Nielsen, Kai. 1973. *Scepticism*. London: Macmillan.

Nozick, Robert. 1981. *Philosophical Explanations*. Oxford: Oxford University Press.

Pappas, George, ed. 1979. *Justification and Knowledge*. Dordrecht: Reidel.

Pappas, George, and Marshall Swain, eds. 1978. *Essays on Knowledge and Justification*. Ithaca: Cornell University Press.

Pollock, John L. 1974. *Knowledge and Justification*. Princeton, N.J.: Princeton University Press.

_____. 1986. *Contemporary Theories of Knowledge*. Totowa, N.J.: Rowan and Littlefield.

Popkin, Richard Henry. 1960. *The History of Scepticism from Erasmus to Descartes*. Assen: Van Gorcum.

_____. 1979. *The History of Scepticism from Erasmus to Spinoza*. Berkeley: University of California Press.

_____. 1980. *The High Road to Pyrrhonism*. San Diego: Austin Hill Press.

Price, H. H. 1935. *Perception*. New York: Robert M. McBride.

Putnam, Hilary. 1981. *Reason, Truth, and History*. Cambridge: Cambridge University Press.

Quine, Willard Van Orman. 1960. *Word and Object.* Cambridge, Mass.: MIT Press.

———. 1982. *Methods of Logic.* 4th ed. Cambridge, Mass.: Harvard University Press.

Rescher, Nicholas. 1973. *The Coherence Theory of Truth.* Oxford: Clarendon Press.

Rorty, Richard. 1979. *Philosophy and the Mirror of Nature.* Princeton, N.J.: Princeton University Press.

Sedley, David. 1983. "The Motivation of Greek Skepticism." *The Skeptical Tradition,* ed. Miles Burnyeat, 9–29. Berkeley: University of California Press.

Sellars, Wilfrid. 1963. *Science, Perception, and Reality.* New York: Humanities Press.

———. 1968. *Science and Metaphysics.* London: Routledge and Kegan Paul.

———. 1975a. "The Adverbial Theory of Objects of Sensation." *Metaphilosophy* 6:144–60.

———. 1975b. "Epistemic Principles." *Action, Knowledge, and Reality: Critical Studies in Honor of Wilfrid Sellars,* ed. Hector-Neri Castaneda, 332–47. Indianapolis: Bobbs-Merrill.

Shatz, David. 1987. "Nozick's Conception of Skepticism." *The Possibility of Knowledge,* ed. Steven Luper-Foy, 242–66. Totowa, N.J.: Rowan and Littlefield.

Shope, Robert K. 1983. *The Analysis of Knowing.* Princeton, N.J.: Princeton University Press.

Skyrms, Brian. 1967. "The Explication of 'X Knows that p.'" *Journal of Philosophy* 64:373–89.

Sosa, Ernest. 1979. "Epistemic Presuppositions." *Justification and Knowledge,* ed. George S. Pappas, 79–92. Dordrecht: Reidel.

———. 1980. "The Raft and the Pyramid." *Midwest Studies in Philosophy* 5:3–25.

Stalnaker, Robert. 1984. *Inquiry.* Cambridge, Mass.: MIT Press.

Stich, Stephen P. 1985. "Could Man Be an Irrational Animal?" *Synthese* 64:115–35.

Stich, Stephen P., and Richard E. Nisbett. 1980. "Justification and the Psychology of Human Reasoning." *Philosophy of Science* 47, no. 2: 188–202.

Stough, Charlotte L. 1969. *Greek Skepticism: A Study in Epistemology.* Berkeley: University of California Press.

Strawson, P. F. 1985. *Skepticism and Naturalism: Some Varieties.* New York: Columbia University Press.

Stroud, Barry. 1968. "Transcendental Arguments." *Journal of Philosophy* 65, no. 9: 241–56.

———. 1977. *Hume.* London: Routledge and Kegan Paul.

———. 1984. *The Significance of Philosophical Scepticism.* Oxford: Oxford University Press.

Swain, Marshall. 1974. "Epistemic Defeasibility." *American Philosophical Quarterly* 11, no. 1: 15–25.

Tarski, Alfred. 1944. "The Semantic Conception of Truth." *Philosophy and Phenomenological Research* 4:341–75.

Unger, Peter. 1971. "A Defense of Skepticism." *Philosophical Review* 80:198–218.

_____. 1975. *Ignorance: A Case for Scepticism.* Oxford: Clarendon Press.

White, Morton. 1981. *What Is and What Ought to Be Done.* Oxford: Oxford University Press.

Wittgenstein, Ludwig. 1958. *Philosophical Investigations.* 3d ed. Trans. G. E. M. Anscombe. Oxford: Blackwell.

_____. 1961. *Tractatus Logico-Philosophicus.* Trans. D. F. Pears and B. F. McGuinness. London: Routledge and Kegan Paul.

_____. 1967a. *Zettel.* Trans. G. E. M. Anscombe. Oxford: Blackwell.

_____. 1967b. *Remarks on the Foundations of Mathematics.* 2d ed. Trans. G. E. M. Anscombe. Oxford: Blackwell.

_____. 1969a. *The Blue and Brown Books.* 2d ed. Oxford: Blackwell.

_____. 1969b. *On Certainty.* Trans. G. E. M. Anscombe. Oxford: Blackwell.

Wright, Crispin. 1980. *Wittgenstein on the Foundations of Mathematics.* Cambridge, Mass.: Harvard University Press.

This page intentionally left blank

Index