

INFORMATION TO USERS

This reproduction was made from a copy of a document sent to us for microfilming. While the most advanced technology has been used to photograph and reproduce this document, the quality of the reproduction is heavily dependent upon the quality of the material submitted.

The following explanation of techniques is provided to help clarify markings or notations which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting through an image and duplicating adjacent pages to assure complete continuity.
2. When an image on the film is obliterated with a round black mark, it is an indication of either blurred copy because of movement during exposure, duplicate copy, or copyrighted materials that should not have been filmed. For blurred pages, a good image of the page can be found in the adjacent frame. If copyrighted materials were deleted, a target note will appear listing the pages in the adjacent frame.
3. When a map, drawing or chart, etc., is part of the material being photographed, a definite method of "sectioning" the material has been followed. It is customary to begin filming at the upper left hand corner of a large sheet and to continue from left to right in equal sections with small overlaps. If necessary, sectioning is continued again—beginning below the first row and continuing on until complete.
4. For illustrations that cannot be satisfactorily reproduced by xerographic means, photographic prints can be purchased at additional cost and inserted into your xerographic copy. These prints are available upon request from the Dissertations Customer Services Department.
5. Some pages in any document may have indistinct print. In all cases the best available copy has been filmed.

**University
Microfilms
International**

300 N. Zeeb Road
Ann Arbor, MI 48106

8322639

Cohen, Stewart Mark

JUSTIFICATION AND TRUTH

The University of Arizona

PH.D. 1983

University
Microfilms
International 300 N. Zeeb Road, Ann Arbor, MI 48106

JUSTIFICATION AND TRUTH

by

Stewart Mark Cohen

A Dissertation Submitted to the Faculty of the

DEPARTMENT OF PHILOSOPHY

In Partial Fulfillment of the Requirements
For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

1 9 8 3

THE UNIVERSITY OF ARIZONA
GRADUATE COLLEGE

As members of the Final Examination Committee, we certify that we have read
the dissertation prepared by Stewart Mark Cohen
entitled Justification and Truth

and recommend that it be accepted as fulfilling the dissertation requirement
for the Degree of Doctor of Philosophy.

Kurt J. Gier

5/23/83
Date

John L. S. J. J.

5/23/83
Date

Charles R. R.

5/23/83
Date

Date

Date

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

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

Kurt J. Gier
Dissertation Director

5/23/83
Date

STATEMENT BY AUTHOR

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

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

SIGNED:

Stewart Coker

TABLE OF CONTENTS

	Page
CHAPTER 1	
OBJECTIVE CONNECTION AS RELIABLE PROCESS	1
Introduction	1
Intuitive Considerations Regarding Necessity	3
Intuitive Considerations Regarding Sufficiency	14
Reliable Process and Undermining Evidence..	20
Logical Considerations	25
CHAPTER 2	
OBJECTIVE CONNECTION AS RELIABLE INDICATION....	36
Intuitive Considerations Regarding Necessity	36
Intuitive Considerations Regarding Sufficiency	45
Logical Considerations	49
CHAPTER 3	
THE DOXASTIC CONNECTION	59
CHAPTER 4	
EPISTEMIC JUSTIFICATION WITHOUT A TRUTH CONNECTION	68
CHAPTER 5	
KNOWLEDGE AND INFORMATION	79
REFERENCES	116

ABSTRACT

The dissertation is a study of the connection between justification and truth. It presents and critically discusses various ways of construing the connection. A dilemma is argued for to the effect that any construal of the connection is defective, while any theory of epistemic justification that ignores the connection incurs an explanatory deficit.

The objective construal of the connection between epistemic justification and truth views such justification as probabilistic. A currently popular version of this view is a theory called Reliabilism. This theory is discussed in terms of both intuitive and purely logical considerations.

Another way to cast out the connection between justification and truth is subjectively. It might be claimed that a subject must have beliefs about the connection between his evidence and the truth of the proposition he believes. This approach is characteristic of coherence theories. These theories are assessed with respect to their psychological reality.

Since objective and doxastic construals of the connection between epistemic justification and truth fail, theories which eschew a truth connection altogether are

discussed. Such an approach is characteristic of foundations theories. It is argued that these theories fail to achieve a level of generality that provides very much insight into the nature of epistemic justification.

The final section of the dissertation is a detailed discussion of naturalized epistemology. The stalking horse is Fred Dretske's information-theoretic approach which relies on a very strong truth connection. The lessons of the previous chapters are applied to Dretske's theory demonstrating its inability to account for the normative aspects of epistemic justification.

CHAPTER 1

OBJECTIVE CONNECTION AS RELIABLE PROCESS

Introduction

Epistemologists generally agree that two components of knowledge are justification and truth. If S believes that P, then S knows that P only if S is justified in believing that P and it is true that P. A central issue in epistemology concerns the connection between justification and truth. While one could view them as conceptually distinct components of knowledge, a variety of philosophers from Descartes to the present have presupposed the view that justification and truth are conceptually related—that there is an internal connection between a belief being justified and being true. The appeal of this view is no doubt rooted in the conviction that knowledge does not arise when a belief merely happens to be true. The motivation for requiring that a true belief be justified in order for it to count as an instance of knowledge just is, in some sense, to provide a connection to truth.

A further consideration that supports this view is the fact that justification is a generic notion. In addition to being epistemically justified, a belief can

be pragmatically or morally justified. The belief of a defense attorney who, in order to provide a better defense, convinces himself that his client is innocent may be said to be justified in one of these latter senses. One might contend that the connection to truth is what distinguishes epistemic justification from these other senses of justification.

The burden of this view is to characterize the nature of the connection between epistemic justification and truth. One connection that is immediately apparent is that S is epistemically justified in believing that P only if S is epistemically justified in believing that P is true. But this trivial connection does not distinguish epistemic justification from moral or pragmatic justification. For one is morally or pragmatically justified in believing that P is true. This is not surprising since this trivial connection seems to be more a point about belief than justification. To believe that P, just is to believe that P is true. The connection between epistemic justification and truth must amount to something beyond this trivial connection.

One might classify theories of epistemic justification according to how they construe the truth connection. In what follows, I propose to do this, and to consider some of the problems that might arise for the various

classifications. I will argue that the issues concerning the truth connection present epistemology with a dilemma.

Intuitive Considerations Regarding Necessity

The strongest view one could take regarding the truth connection is that taken by Descartes. The Cartesian view is that justification logically entails truth. To put it schematically: It is a conceptual truth that, if conditions C justify belief B for subject S, then C¹ logically entails that B is true.

The legacy of the Cartesian view is skepticism. Descartes demonstrated in the first meditation that no such connection is forthcoming (the arguments of the later meditations notwithstanding). Given any plausible specification of C for any S, it will always be logically consistent to suppose that not B. This is what the evil demon argument shows. Where, e.g., C comprises facts about sensory data, and where B is a belief about the truth of some empirical proposition, it is always logically possible that the evil demon has arranged for C to obtain

1

Probably, Descartes is most reasonably interpreted as holding this view vis a vis some notion of philosophical certainty rather than ordinary justification. I will sidestep the scholarly issues by simply referring to this view as the Cartesian view whether or not Descartes actually held it.

where B is false. Not wishing to be saddled with this skeptical result, most contemporary philosophers have rejected the Cartesian view and have opted instead for a fallibilist theory of justification. A fallibilist theory allows that where C makes B justified for S, it is still possible that B is false.

While fallibilism does seem to avoid skepticism, one might still inquire as to what the truth connection comes to on a fallibilist view. A natural proposal would be to construe the connection as probabilistic. This tack has been taken by several philosophers (Goldman, Swain, et al.) whose theories can be grouped under the general heading of Reliabilism. For purposes of exposition,² I focus on Goldman's view.

Goldman's brand of Reliabilism explicates the connection between justification and truth in terms of the truth-frequency of the justified belief. There is no logical guarantee that a justified belief is true on this

2

Alvin I. Goldman, "What is Justified Belief," in George S. Pappas Justification and Knowledge (Dordrecht: 1979); Marshall Swain, Reasons and Knowledge (Ithaca: 1981). Two philosophers have proposed reliability analyses of knowledge (London: 1973) and Fred I. Detske, Knowledge and the Flow of Information (Oxford: 1981). Because these analyses are aimed at knowledge rather than justification, it is not clear whether what I say applies to them.

view. What is guaranteed is that a justified belief is produced by a cognitive process that is reliable, i.e., by a cognitive process that tends to produce true beliefs. Goldman leaves it open just how reliable a cognitive process must be in order for it to confer justification on the belief³ it produces. However, he is explicitly a fallibilist. Since reliability is a probabilistic notion, for our purposes Goldman's position amounts to: It is a conceptual truth that, if C justifies B for S, then C makes it probable that B is true.⁴ So, for example, if the belief "There is something red before me" is justified in virtue of its being produced by a perceptual process, then the fact that the belief was so produced makes it probable that it is true.

It is clear that Reliabilism avoids the skepticism of the Cartesian conception of the truth connection. For the evil demon hypothesis only demonstrates the possibility that there are belief forming processes which are not reliable. The fact that this is consistent with their

3

Ibid., (p. 11).

4

Unlike Goldman's frequency probability analysis, Swain analyzes reliability in terms of inductive probability. On this interpretation, C makes B more probable than ~B only if C constitutes better evidence for B than ~B op. cit. (p. 100). This epistemic interpretation would seem to render the analysis circular.

actual reliability obviates any skeptical conclusion. But skepticism aside, I think the evil demon hypothesis (or its contemporary neurophysiologist version) uncovers a defect in the Reliabilist position. We can see this by supposing the hypothesis to be true. Imagine that unbeknown to us, our cognitive processes (e.g., perception memory, inference) are not reliable owing to the machinations of the malevolent demon. It follows on a Reliabilist view that the beliefs generated by those processes are never justified.

Is this a tenable result? I maintain that it is not. Of course, we are not here supposing that we know that the demon hypothesis is true. Certainly if we were to know that our cognitive processes are unreliable then the beliefs they generate would not be justified. What we want to suppose is the mere truth of the demon hypothesis. Now part of what the hypothesis entails is that our experience is just as it would be if our cognitive processes were reliable. Thus, on the demon hypothesis, we would have every reason for holding our beliefs, that we have in the actual world. Moreover since we actually have reason to believe that our cognitive processes are reliable, it follows that in the demon world we would have every reason to believe that our cognitive processes were in fact reliable. We might even imagine that a brilliant philosopher had seemingly demonstrated (a la Descartes of the

later meditations) the falsity of the demon hypotheses, to the extent that anyone who could follow the reasoning was (intuitively) justified in accepting the conclusion.

It strikes me as clearly false to deny that under these circumstances our beliefs could be justified. If we have every reason to believe e.g., perception, is a reliable process, the mere fact that unbeknown to us it is not reliable should not affect it's justification - conferring status (a fortiori if we have good reason to believe that the conditions which in fact make perception unreliable do not obtain.)^{5,6}

5

Goldman might reply that in the evil demon case, the cognitive processes themselves are reliable. It is only because of the intervention of the evil demon that the beliefs they produce turn out false. Equivalently, Goldman might hold that reliability in a non-manipulated environment is a necessary condition of justification. I don't believe anything crucial hangs on this. Certainly we can imagine a world where the cognitive processes are unreliable in a non-manipulated environment. We might imagine a world where, e.g., the perceptual processes naturally produce consistent sets of hallucinations.

6

In response to this objection, Goldman has suggested (in conversation) that his theory be interpreted in the following way. S's belief at t is justified only if s's belief at t results from a cognitive belief-forming process that is reliable in the actual world. (He makes the same point in a different context in "What is Justified Belief?" op. cit. p. 17). On this interpretation, the e.g., perceptual beliefs of the inhabitants of the Demon world would be justified, since perception is a reliable belief-forming process in the actual world.

However, the theory so construed would be too strong. We can imagine a world with beings whose perceptual faculties operate according to different

My argument hinges on viewing justification as a normative notion. Intuitively, if S's belief is appropriate to the available evidence, he is not to be held responsible for circumstances beyond his ken. Goldman can be viewed as having illustrated this when he discusses a possible counter example to his analysis. He considers whether wishful-thinking would be a justification - conferring process in world W, if it was a reliable process in W. In a footnote he points out that "if people in world W learn inductively that wishful thinking is reliable, and regularly base their beliefs on this inductive inference, it is quite unproblematic and straightforward that their beliefs are justified."⁷ Goldman's point is that in this case, the inductively inferred reliability of wishful thinking can be sufficient for beliefs produced by wishful-thinking to be justified in W. But if this is true, the picture should not change if wishful-thinking turns out, contrary to the inductive evidence, to be unreliable in W.

Now Goldman might contend that inductive inference is itself a reliable belief-forming process and that this

natural laws than the ones that hold in the actual world. Since the perceptual faculties of such beings would be unreliable in the actual world, they would be precluded from having justified perceptual beliefs.

⁷

Op. cit. (p. 23).

fact explains the justifiedness of the beliefs in question. But again, I fail to see how the picture changes epistemically if as a result of infelicitious circumstances, beliefs produced in accordance with the canons of inductive inference turn out to be inscrutably false most of the time.⁸

An entrenched Reliabilist may be unmoved by this appeal to the normative character of epistemic justification. One might insist that a world where the evil demon hypothesis is true is a world where there are no justified beliefs. I will argue that a theory with this consequence fails to capture a central, perhaps the central distinction in epistemology. Moreover, I think it will be clear that the distinction is most plausibly construed as marking the difference between justified and unjustified belief.

When Goldman is marshalling data for his theory, he lists certain belief-forming processes whose outputs we would consider to be unjustified. He gives these examples: confused reasoning, reliance on emotional attachment, mere hunch or guesswork and hasty generalization. Goldman notes that these processes all share the

8

If the evil demon should occasionally arrange for a belief to be true, then we would have a Gettier case. This indicates that reliability is better suited to being a separate component of knowledge rather than a constituent of justification.

feature of unreliability. He contrasts these with processes that intuitively issue in justified belief, e.g., standard perceptual processes, remembering, good reasoning, and introspection. Noting that these latter process all have reliability as a common feature, Goldman goes on to suggest, "The justificational status of a belief is a function of the reliability of the process or processes that cause it." (p. 10)

I have argued that the evil demon case shows that reliability is not necessary for justification. This claim can be bolstered by considering the contrast between the two categories of belief-forming processes Goldman lists.

Since Goldman cites reliability as the relevant feature that distinguishes the class of justification-conferring processes from the class of non-justification conferring processes, he is committed to the view that in the evil demon world, the cognitive processes he lists are indistinguishable from the perspective of epistemic justification. Thus we can imagine two inhabitants of this world, A who is a good reasoner, i.e., reasons in accordance with the canons of inductive inference, and B who engages in confused reasoning, wishful thinking, reliance on emotional attachments, guesswork etc. Since the beliefs of A & B are both produced by unreliable processes (the evil demon sees to this), a reliabilist

theory of justification must render identical epistemic appraisals of both sets of beliefs.

Plainly, this cannot be correct. A's beliefs are conditioned by the evidence whereas B's beliefs are not. A is a good reasoner whereas B is not. A's beliefs are reasonable whereas B's beliefs are not. There is a fundamental epistemic difference between the beliefs of A and the beliefs of B. But the Reliabilist does not have the theoretical means to display this difference.

I would claim that the distinction between the beliefs of A and B is marked precisely by the concept of justified belief. Beliefs produced by good reasoning are paradigm cases of justified belief and beliefs arrived at through fallacious or arbitrary reasoning are paradigm cases of unjustified belief. Whether or not reasoning results in false belief, even if this happens more often than not, is irrelevant to the question of whether the reasoning is good. To maintain otherwise would be on a par with confusing truth and validity.

9

Strictly speaking, Goldman holds only that reasoning processes must be conditionally reliable in order for them to yield justified beliefs. "A process is conditionally reliable when a sufficient proportion of it's output-beliefs are true given that it's input-beliefs are true" ("What is Justified Belief." p. 13). However we can easily suppose that the Demon world is a counter-inductive world. Of course this fact is concealed from its inhabitants by the Demon and so does not affect the justificational status of their beliefs. In such a world,

The Reliabilist might acknowledge that there is a clear epistemic distinction between the beliefs of A and B, but that the distinction has nothing to do with their justificational status. Rather the difference is that A's beliefs are reasonable or rational whereas B's are unreasonable or irrational.

This maneuver would be of no help. First of all, "reasonable" and "rational" are virtual synonyms for "justified." But we need not quibble over semantics. If the Reliabilist wants to distinguish "justified" from "reasonable" or "rational" he may do so. But clearly the important epistemic concept, the one epistemologists have been concerned with is what the Reliabilist would call "reasonability" or "rationality." The difference between the beliefs of A in the evil demon world and the beliefs of B in the evil demon world seems to capture the concept that epistemologists have been worried about rather than the difference between the beliefs of A in the actual world and the beliefs of A in the evil demon world. In effect, the Reliabilist would be changing the subject.

Another move open to the Reliabilist is to limit his theory to non-discursive belief. The above objection hinges on the role of reasoning in justification. We

induction would not even be conditionally reliable.

noted that one can be justified in holding a belief in virtue of having arrived at the belief through good reasoning even if the reasoning is unreliable. However this standard of assessment would not apply to non-discursive processes like perception. (We can assume for purposes of this discussion that certain processes like perception do not involve reasoning). Here one might claim that the only criterion available for assessing such processes is reliability. Thus the only way a perceptual belief can be appraised with respect to its justificational status is in terms of its reliability.

Suppose perception is a non-discursive process. It does not follow that perceptual beliefs can only be assessed in terms of their reliability. Again the evil demon hypothesis will help to clarify this. Consider again two inhabitants of the evil demon world, A and B. Both have unreliable perceptual processes. Suppose both A and B believe there is something ϕ before them on the basis of being appeared to ϕ -ly. While A has no evidence to the contrary, B is presented with strong evidence that owing to a clever deception there is nothing ϕ before him.

I think it's clear that there is a fundamental epistemic difference between A's perceptual belief and B's perceptual belief -- a difference which again underscores the normative character of epistemic justification.

Notice that we need not assume that B disregards the evidence as a result of any discursive process. He may just arbitrarily ignore it. But from an epistemic point of view, B ought not to have proceeded in the way he did. We might say that contrary to A, B has been epistemically irresponsible in accepting that there is something before him. As a result, while A is justified in his perceptual belief, B is not.

Intuitive Considerations Regarding Sufficiency

An intuitive case can also be made for the non-sufficiency for epistemic justification of Goldman's theory. The recipe for finding a counter-example to reliability as a sufficient condition for justification is to take an intuitively unjustified process (that is a process that intuitively does not produce justified beliefs) and suppose that it were reliable. Goldman himself¹⁰ discusses such a case. He imagines a possible world S where a benevolent demon arranges things so that whenever (or at least the majority of times) the inhabitants of W arrive at a belief by the process of wishful thinking, the belief is true. It follows that in W wishful thinking is a reliable process. Thus on Goldman's view, it turns out that such beliefs are justified.

¹⁰ Alvin J. Goldman, "What is Justified Belief?" in George S. Pappas, Knowledge and Justification (Dordrecht:

This result is strongly counterintuitive. Of course, the inhabitants of W could amass inductive evidence for the reliability of wishful thinking and thereby acquire a straightforward justification for their beliefs. But Goldman is committed to the view that prior to the marshaling of inductive evidence, the wishful beliefs are justified solely in virtue of the reliability of wishful thinking.

Goldman acknowledges that this case poses a *prima facie* difficulty for his theory and discusses several candidate responses.

One possibility, according to Goldman, is to require that the reliability of the process occur in a non-manipulated environment, i.e. an environment in which there is no purposeful arrangement of events and states of affairs to accord with beliefs. On this criterion, the wishful beliefs of the inhabitants of W would not be justified because of the intervention of the benevolent demon.

I do not believe this revision rules out anything essential in the counterexample. We can imagine that the inhabitants of W actually have a clairvoyant power that enables them to reliably predict the future. However the inhabitants of W are unaware that they have this

power. Furthermore, it is a fact about their psychology that they only wish that P if they can (unbeknown to them) predict that P. As far as they know (or believe) they are merely wishing that p and inferring P on that basis. It seems to me not to matter whether the reliability of wishful thinking in W is explained by the clandestine intervention of a benevolent demon or the unconscious operation of a clairvoyant faculty. The crucial factor, what by my lights makes the beliefs unjustified in these cases, is the fact that the reliability of the belief forming process is due to facts that are completely outside the ken of the subject. If as far as the subject knows, the state of affairs expressed by P is merely something he wishes for, then he is being epistemically irresponsible in accepting that P. This, again, is a consequence of the normative character of epistemic justification.

The second response Goldman discusses is to claim that a belief is justified if it is produced by a process that is reliable in our world. Thus the cognitive processes in W are not to be assessed in terms of their reliability in W, but rather in terms of their reliability in the actual world. Since wishful thinking is not a reliable process in the actual world, it does not follow that it issues in justified belief in W, despite its

reliability in W. This would save the theory from the counter-intuitive result.

To revise the theory in this way seems to me to rob it of any intuitive appeal it might have had. Surely if reliability theories have any intuitive force, its due to the presumed close connection between epistemic justification and truth. But if that connection is going to serve as the basis for an analysis of epistemic justification, it must be conceptual. This would entail that beliefs that are justified are produced by reliable processes in any possible world. If it is allowed that a belief can be justified in some possible world by an unreliable process, then the conceptual basis for a reliabilist theory is undermined. The connection between the truth of the belief and its justificational status would be lacking. The mere fact that the process which generates the belief is reliable in the actual world would seem to be insufficient to motivate an intuition even in those who have reliabilist intuitions.

A further problem with this amended version of the theory is pointed out by Goldman himself. Goldman imagines the possibility that a benevolent demon indolent up to the present actually exists. If he began ensuring the reliability of wishful thinking, it could turn out to be reliable (in the long run) in the actual world.

Thus beliefs produced by wishful thinking would be justified in any possible world including the actual world. Goldman concedes that this contravenes our intuitive judgments.

His response to this objection is somewhat puzzling:

[In a conceptual analysis] ... What we really want is an explanation of why we count, or would count, certain beliefs as justified and others as unjustified. Such an explanation must refer to our beliefs about reliability, not the actual facts. The reason we count beliefs as justified is that they are formed by what we believe to be reliable belief-forming processes. Our beliefs about which belief-forming processes are reliable may be erroneous, but that does not affect the adequacy of the explanation. Since we believe that wishful-thinking is an unreliable belief-forming process, we regard beliefs formed by wishful thinking as unjustified. What matters then is what we believe about wishful thinking not what is true (in the long run) about wishful thinking.¹¹

There are several things Goldman might mean by these remarks. If he is claiming that our intuitions on the wishful thinking case might be perverted by a misconstrual of the facts of the case, then he is surely correct, but uninterestingly so. It may be that some people implicitly carry over their beliefs about the reliability of wishful thinking in the actual world to the imagined world W. If Goldman is claiming that this is always the case then his claim is doubtful. In my own case, I have no trouble (or so it seems to me) in imagining

¹¹ Ibid. p. 18.

the possibility that wishful thinking is reliable in W. When my intuitions tell me that the wishful beliefs of the inhabitants of W are unjustified, it is not because I am unable to suppose that it is really the case in W that wishful thinking is reliable. It is because of a strong conviction that wishful thinking reliable or not, is not the sort of cognitive process that generates justified beliefs. Of course it might be true that I am unable to sort out accurately the psychological factors that influence my intuitions. But it is surely an inadequate response to a counterexample simply to claim that those who have intuitions in support of the example are introspectively deceived. There must be some independent reason for supposing that this is so.

Goldman's final response to the wishful thinking case is to bite the bullet. Claiming that his own intuitions and those of others he has consulted are "not entirely clear", Goldman countenances the possibility that the right thing to say is that the beliefs in W produced by wishful thinking are justified. I don't presume to doubt the sincerity of Goldman or those he has consulted. No doubt intuitions can diverge on these bizarre cases, especially given that one's intuitions can be affected by one's theory.

However as we saw in the intuitive case against reliability as necessary for epistemic justification,

certain intuitions can commit one to untoward theoretical results. This can constitute grounds for abandoning those intuitions. Now suppose that the cognitive processes reliable in our world are also reliable in W. Again imagine two inhabitants of W, A and B. A reasons according to the canons of inductive inference, assumes that *prima facie*, things are the way they appear to be etc. B acquires his beliefs by wishful thinking. If one adheres to the position that reliability is a sufficient condition of justification, then one must give the same epistemic appraisal to the beliefs of A and B. But there is obviously a great epistemic difference between the beliefs of A and the beliefs of B. Surely any epistemological theory that lacks the resources to display this difference is deficient.

Again I would claim that the disparity between the beliefs of A and B is precisely in their justificational status. If the reliabilist wants to characterize the difference in terms of the rationality or reasonability of the beliefs, then I would claim that these concepts constitute the traditional domain of epistemic theories. But this is old ground and I will not belabor the point.

Reliable Process and Undermining Evidence

A *prima facie* problem for reliabilist theories is their inability to account for the role of evidence in epistemic justification. We saw this problem emerge

in the intuitive argument against the necessity of reliability in justification. We noted that a subject can have overwhelming evidence for a proposition and yet still fail to have a reliably produced belief in that proposition. This issue arises again in connection with the sufficiency of reliably produced belief for justification. The problem stems from the fact that a belief can be produced by a cognitive process that is in fact reliable even when one has strong evidence that the process is unreliable. Under such conditions, we would not want to say that the belief is justified. Goldman recognizes this problem and proposes a solution. I will argue that the solution is ineffective.

The case Goldman discusses is as follows: S has a reliable memory. He is told mendaciously by generally reliable sources that his memory is defective. Nonetheless, S persists in trusting his memory beliefs. Goldman says that intuitively these memory beliefs are not justified, the reason being that S has strong evidence that his memory beliefs are suspect. However, since S's memory is in fact reliable, Goldman's theory must sanction those beliefs as justified. Again it looks as if reliabilism overlooks crucial epistemic concepts--in this case, undermining evidence.

Goldman believes he can account for the role of undermining evidence within a reliabilist framework. He notes that although S has strong evidence against the truth of his memory beliefs, he fails to use it. If S were to use this evidence properly, he would give up his memory beliefs. Goldman goes on to claim that the proper use of evidence is a (conditionally reliable) process. If S had employed this process, he would not have acquired his memory beliefs. As such Goldman revises his theory to make justifiedness not only a function of the cognitive processes that are actually used, but a function of the processes that could and should be used:

If S's belief in p at t results from a reliable cognitive process, and there is no reliable (or conditionally reliable) process available to S, which, had it been used by S in addition to the process actually used, would have resulted in S's not believing p at t, then S's belief in p at t is justified.¹²

Goldman acknowledges that the principle has a technical problem, viz., one cannot use an additional process as well as the original process if the two processes result in different beliefs. Also he admits that the notion of an available process is somewhat vague. The first problem Goldman considers to be merely technical and does

12

Alvin I. Goldman, "What is Justified Belief?" in George S. Pappas Knowledge and Justification (Dordrecht: 1979) p. 10.

not attempt to solve it. The second problem Goldman claims reflects the vagueness of our ordinary concept. Since this is the concept he is trying to analyze, Goldman does not see this as an unacceptable result.

I will not quarrel with Goldman about these two points. I believe there is a much more basic and fundamental problem with Goldman's approach. Suppose "the proper use of evidence" turns out to be unreliable (for whatever reasons, e.g., the evil demon etc.). Then it will turn out that S has a belief that results from a reliable cognitive process and there will be no reliable process available to S which, had it been used by S in addition to the process actually used, would have resulted in S's not believing p at t. Thus Goldman's theory would entail that S has a reliable belief in this case.

Now Goldman can always claim that if the proper use of evidence turns out to be unreliable in this case, then S's memory would be justified. But if we examine the language he uses to set up the problem, we can see just how heroic this response would be. He tells us, "S has reason to believe that his belief is caused by an unreliable process....¹³ Jones has strong evidence¹⁴ against certain propositions concerning his past." But

¹³ Ibid. p. 19.

¹⁴ Ibid. p. 19.

certainly, if evidence *e* or reason *r* (unbeknown to *s*) leads *S* astray more often than not, this does not affect *e*'s status as evidence or *r*'s status as a reason. So even if the proper use of evidence turns out to be unreliable, it would still be true that *S* has reason to believe that his belief is caused by an unreliable process and that *S* has strong evidence against certain propositions concerning his past. Here we are able to see clearly that the concept of reliability has little to do with the relevant concept of evidence for epistemic justification.

Goldman might reply that evidence that leads one astray more often than not is not really evidence. It only appears to be evidence. One could certainly talk this way. But the problem can be recast in these terms. For if *S* is unable to determine that that what appears to be evidence is not really evidence, then *S* cannot dismiss what seems to be evidence arbitrarily and not suffer any epistemic penalty. This is a consequence of the normativity of epistemic justification.

Interestingly enough, Goldman virtually concedes the point. He tells us, "...Jones...fails to use a certain (conditionally) reliable process that he could and should have used. Admittedly, had he used this process he would have 'worsened' his doxastic states: he would have replaced some true beliefs with suspension of judgment. Still he couldn't have known this in the case

in question. So he failed to do something which
epistemically he should have done." [my emphasis].¹⁵

This passage is replete with normative language. The central point is that the subject is presented with 'evidence' that in some sense he has an epistemic duty to consider. If he fails to discharge this duty, the subject cannot acquire a justified belief.

Goldman acknowledges that the epistemic duty is not nullified simply because in this case the evidence is misleading. For S could not have known that in this case. But surely the situation does not change simply because the evidence is misleading in general, not just in this case. As I have set up the example, S still could not have known that in this case the evidence is misleading. For that matter S could not have known that the evidence is misleading in general. Thus S is remiss in his epistemic duty if he fails to make proper use of that 'evidence.' As such S cannot acquire a justified belief. Yet as we previously noted Goldman's theory yields the result that S does acquire a justified belief.

Logical Considerations

I consider the intuitive objections to reliabilism to be decisive. Once we recognize the normative character

¹⁵

Ibid. p. 20.

of epistemic justification, it is clear that reliabilist theories must fail. However, the entrenched reliabilist may reject this appeal to the normative character of justification. Thus we may consider some objections to reliabilism that are of a purely logical nature.

The major problem results from the vagueness of the requirement that justification-conferring cognitive mechanisms be reliable. Goldman recognizes this problem and proposes a tentative solution.

A critical problem concerning our analysis is the degree of generality of the process-types in question. Input-output relations can be specified very broadly or very narrowly, and the degree of generality will partly determine the degree of reliability. A process-type might be selected so narrowly that only one instance of it ever occurs, and hence the type is either completely reliable or completely unreliable ...If such process-types were selected, beliefs that are intuitively unreliable might be said to result from perfectly unreliable processes.¹⁶

The issue of how thinly to slice process-types can be equivalently reformulated as the issue of what circumstances are to serve as the context for the reliability of belief forming mechanisms. It is not open to the reliabilist to use a context-free requirement of reliability. Since some processes are reliable in some circumstances and unreliable in others, a context-free reliability theory would be untenable. For suppose a process M were unreliable

16

Alvin I Goldman, "What is Justified Belief?" in George S. Pappas Justification and Knowledge (Dordrecht: 1979) p. 12.

in general but reliable in the present circumstances. We can imagine a planet with sunlight that is generally insufficient for the general reliability of vision. Occasionally the clouds clear and the lighting improves to allow for reliable vision. Surely we would not want to preclude the possibility of justified visual beliefs in the periods of good lighting simply because of the general unreliability of vision resulting from the generally inadequate lighting. This relativization would seem to be necessary as well in cases where a process is generally reliable, yet unreliable in specific circumstances. Certainly the reliability of vision in periods of adequate lighting cannot be sufficient for the justification of visual beliefs in periods of darkness, even if there is adequate lighting most of time. (What seems relevant is what one is justified in believing about the adequacy of the lighting in the present circumstances, not what merely happens to be true.) Thus the reliabilist is forced to make reference to contexts.

The preceding example suggests that the reliabilist should formulate his theory in this way:

"S' belief that P at t is justified" iff "S' belief at t results from a cognitive process that is reliable in the circumstances at t."

But now the problem becomes how to decide at what level of generality the circumstances at t are to be specified.

(This is just another way of stating Goldman's problem of deciding at what level of generality process-types are to be sliced.) For the circumstances at t may be of some type c such that the process is reliable in c , but also be of some narrower type c' such that process is unreliable in c' . Suppose the circumstances at t were such that the lighting was adequate for reliable vision, but the atmosphere contained a gas that caused frequent hallucinations. Presumably the reliabilist would not want to sanction as justified the visual beliefs produced in these circumstances.

Since it would seem that the circumstances at t can be specified more and more narrowly without limit, perhaps the reliabilist could rely on the notion of a maximal specification of the circumstances at t relative to a process P . This notion can be defined as follows:

" M is a maximal specification of the circumstances at t relative to process P iff any further specification of the circumstances at t beyond M would not alter an assessment of the reliability of P from what it is in M ."

Then the reliabilist would reformulate his theory as:

" S ' belief that P at t is justified" iff " S ' belief at t results from a cognitive belief forming process that is reliable under a maximal specification of the circumstances at t ."

This theory is both too strong and too weak, since a maximal specification of the circumstances at t would

include the truth-value of P at t. But any process will be reliable in circumstances in which the belief it produces is true and any process will be unreliable in circumstances in which the belief it produces is false. This theory would yield the result that the truth of a belief P is both a necessary and a sufficient condition for p being justified. But this contravenes the truism that one can be justified in holding false beliefs and that one can accept true propositions without justification.

As I noted earlier, Goldman discusses this problem in terms of how thinly the process-types should be sliced. In terms of the present discussion, the issue is how narrowly the circumstances at t should be specified. One restriction Goldman tentatively suggests is that the specification of the circumstances at t should be content-¹⁷neutral with respect to the belief P. This avoids the result that e.g., an inferential process is reliable in certain circumstances simply because those circumstances include the fact that the inferred proposition is asserted by the Pope. As Goldman points out, if the Pope is infallible, this would guarantee the reliability of the inferential process. The Content-neutral restriction would preclude this since such a specification of the

17

Ibid., p. 12.

circumstances would limit the beliefs produced by the process to those whose propositional content were uttered by the Pope.

This restriction would also seem to handle the truth-value problem I have raised. For a specification of the circumstances that included the truth-value of the belief would limit the values of P to propositions with that same truth-value. Nonetheless, the content neutrality restriction is too strong. For sometimes a specification of the circumstances that limits the content of the beliefs explains the justifiedness of those beliefs. Consider the cognitive process of inferring that the testimony of others is correct. Surely such a process results in many instances of justified belief. The task of the reliabilist is to account for this fact in terms of the reliability of the inferential process. But the process will be reliable only when the circumstances are such that the testimony it takes as input is reliable. But this entails that the reliability of the process can only be assured if it is specified that the process operates on beliefs that are reliably held. This would involve a content restriction on the beliefs generated by the process, viz, beliefs that are reliably held. Since this is disallowed, the theory would be unable to explain the justification of beliefs inferred on the basis of reliable

testimony. (I would hold that the relevant consideration is not the actual reliability of the testimony but rather what the subject is justified in believing regarding the reliability of the testimony.)

Actually, it appears now that the content-neutrality restriction is obviously too strong. Any specification of circumstances is going to violate it since any such specification will restrict the content to the content of beliefs produced in those circumstances.

On the other hand, if it is maintained that certain circumstances admit of beliefs with any contents (that is the contents are not to be restricted to the actual contents of beliefs in those circumstances but rather the possible contents) I would hold that the same is true of Goldman's own example, viz., propositions inferred by the pope. Surely it is possible that the Pope assert virtually any proposition. Thus, construing the restriction in this way would render it ineffective in the very cases it was designed to handle.

There are further considerations that militate against reliability as a sufficient condition of justification. If we assume that reasoning often plays a central role in epistemic justification, then reliabilism (or at least Goldman's brand) founders on the lottery paradox. On separate occasions, a (less than perfectly) reliable

process could generate a set of beliefs that are jointly inconsistent. We can imagine a lottery with a large number of tickets. S reasons via a reliable inferential process that a particular ticket will lose. S uses this same reasoning for each ticket and concludes in each case that the ticket will lose. (We may also suppose that S knows that one ticket will win.) On Goldman's theory, it follows that s has a justified belief in the case of each ticket that it will lose. Now if s reasons correctly to a conclusion from a set of justified beliefs, then either S is justified in believing the conclusion or else S loses his justification for one of the premises. To reason from the premise p and the premise q to the conclusion p & q is valid reasoning. But then in this case S would be justified in concluding either that $t_1 t_2 \dots t_n$ will lose and one ticket will win; or S would lose his justification for one of his beliefs that a particular ticket will lose. Clearly the first alternative is unacceptable, (We may suppose that S recognizes the inconsistency of his conclusion). And since the fact remains that each of S' beliefs that a particular t_1 will lose has been reliably produced, the second alternative is equally unacceptable for the reliabilist. It follows that reliability cannot be a sufficient condition for epistemic justification.

In a footnote, Goldman acknowledges that his theory is susceptible to the lottery paradox and tries to defuse the objection:

In reply to this objection, we might simply indicate that the theory is intended to capture our ordinary notion of justifiedness, and this ordinary notion has been formed without recognition of this kind of problem. The theory is not wrong as a theory of the ordinary (naive) conception of justifiedness.¹⁸

I find this reply puzzling. Certainly it is part of the ordinary conception of justifiedness that a person cannot be justified in accepting a proposition he knows to be false. And we have seen that if we take reliability as a sufficient condition of justification, then our ordinary conception of justifiedness is violated in just this way. (Since all the premises of the reasoner in the lottery paradox example would be justified, he would be justified in accepting that no ticket will win.) So in one sense Goldman is just wrong to claim that his theory captures the ordinary conception of justifiedness. But Goldman may simply be saying that the lottery paradox is inherent in our ordinary conception of justification. On the one hand, that conception sanctions the particular beliefs that for each t_i , t_i will lose. On the other hand, the ordinary conception will disallow the belief that no

18

Ibid., p. 22.

ticket will win. It should be expected that the paradox will emerge in a theory of our ordinary conception of justification.

It is not at all clear that the lottery paradox is part of our ordinary conception of justification. I would agree that it is part of that conception that for each ticket t_i , the proposition that t_i will lose is warranted to some degree. I would deny that it is part of that conception that these propositions are warranted to the degree necessary for knowledge. This is evidenced in the fact that although we are inclined to say that the proposition that t_i will lose is justified for s , we would deny that s knows that t_i will lose even if it is true that t_i will lose. And this does not seem to be a consequence of any Gettier effect. If it were true that people knew that they were going to lose the lottery, then it would be hard to explain the rationality of buying a ticket even if the prize money were sufficiently great to make the expected utility positive. As such, Goldman is wrong in claiming that the lottery paradox is inherent in our ordinary conception of justification-adequate-for-knowledge. And it is just this conception that Goldman purports to analyze: "Many epistemologists have been interested in justification because of its presumed close relationship to knowledge. This

relationship is intended to be preserved in the conception
of justified belief presented here."¹⁹ Contrary to
Goldman then, I would view the susceptibility of his theory
to the lottery paradox to be a defect in that theory.

19

Ibid., p. 1.

CHAPTER 2

OBJECTIVE CONNECTION AS RELIABLE INDICATION

Intuitive Considerations Regarding Necessity

Marshall Swain has proposed a Reliabilist theory of justification. Swain holds that the justification of a belief is a function of the reliability with which a subjects holding the belief indicates the truth of the belief.

In this chapter, I will argue that Swain has not succeeded in presenting necessary conditions for justification. We can begin by examining Swain's theory. The precise statement of the theory is:

(DEJR) S's believing that h on the basis of R is epistemically justified at t iff: S's believing that h on the basis of R is a reliable indication that h at t.¹

Swain defines the concept of reliable indication as:

(4.1) S's believing that h on the basis of set of reasons r is a reliable indication that h at t iff: there is some set of relevant characteristics, C, such that

(1) S has C at t (that is, each member of C is a characteristic of S at t); and

1

Marshall Swain, Reasons and Knowledge (Cornell: 1981), p. 99.

- (2) the probability that h , given that S has C and that S believes that h on the basis of R , is greater than the probability that not- h , given those same facts about S .²

Swain illustrates his theory with this example:

Consider ... the case of Alfred who is sitting in his study watching the snow fall outside. Alfred has come to be in certain reason states R , as a causal result of the event of the snow falling. Moreover, Alfred comes to believe that it is snowing, and this belief is based upon his set of reason states, R . This belief is analogous to the state of the barometer that consists in the pointing of the needle to the mark P on its dial. Moreover, in constructing this case, we have supposed that Alfred is in good working order; he is wide awake, sober, has good vision and so forth. There is, in other words, some set of relevant characteristics, C , that Alfred has, and by virtue of his having these characteristics it can be said that he is in good working order. Just as we say of the barometer, given its condition, that its being in the state of registering P is a reliable indication of the atmospheric pressure we can also say of Alfred that his believing that it is snowing outside on the basis of his reasons R , is, given the condition C that he is in a reliable indication that it is snowing outside.³

Although Swain is not explicit about this, there is a clear need to include characteristics of the environment as well as characteristics of the subject in the reliability assessment. Alfred could be wide awake, sober, have good vision, etc. but the lighting could be inadequate or his vision could be obstructed. If these

2

Ibid., p. 100.

3

Ibid., pp. 97-98.

latter conditions obtained, Alfred could not be said to have a justified visual belief. Since Swain's analysis does not include a provision for environmental factors, we may broadly construe C (the characteristics of the subject) to range over the state of the subjects environment.

Swain's definition of 'reliable belief' is amended to deal with a variety of technical objections. At present we are concerned with the intuitive issues and the initial formulation will be adequate for this purpose. Definition (4.1) cashes out the notion of reliability in terms of the notion of probability. This is as it should be since reliability is a probabilistic notion. Swain makes several comments regarding the connection between probability and reliability and his interpretation of probability;

What are we saying of Alfred, his belief, and his characteristics when we say that these things provide a reliable indication that it is snowing? My answer is that ascriptions of reliability are evidential claims, that is, claims that these facts about the subject constitute good evidence for the truth of the proposition believed. The information that Alfred has the set of characteristics C, (where the characteristics would be specified) and that Alfred believes that it is snowing on the basis of the set of reasons R, (where the reasons would also be specified) constitutes good evidence for the claim that it is snowing outside; that is, we would have better evidence for the claim that it is snowing outside than for the claim that it is not snowing (assuming that we did not also have negative evidence).

Claims to the effect that a certain body of information constitutes better evidence for h than for the denial of h can be represented using conditional inductive probability expressions. Of Alfred we might say, 'The probability that it is snowing outside, given that Alfred believes that it is snowing

outside on the basis of the set of reasons R and given that Alfred has the set of characteristics C, is greater than the probability that it is not snowing outside, given those same facts about Alfred.'⁴

In order to understand Swain's theory, we need to understand the notion of evidential support, i.e., we need to know what conditional inductive probability measures. Unfortunately, Swain is of little help here. "Because ... I do not know how to define the notion of evidential support--I take it as undefined."⁵ The first thing to notice is that Swain is in danger of trivializing his theory. He is analyzing justification in terms of reliability and he is analyzing reliability in terms of inductive probability. But according to Swain, this sort of probability is a measure of the strength of the relation of evidential support. And certainly it is not very illuminating to view epistemic justification in terms of evidential support, wherein this latter notion is left undefined. To say that a certain proposition e constitutes good evidence for S to believe h just is to say that e justifies (to some degree) h for S.

In fact I believe that Swain's theory is not trivial. We can glean this from some clarificatory remarks that

⁴
Ibid., pp. 99-100.

⁵
Ibid., p. 101.

he makes about evidential support;

Another respect in which my strategy may be misunderstood emerges from the earlier remark that, 'ascriptions of reliability are evidential claims, claims to the effect that these facts constitute good evidence for the truth of the proposition believed.' One might ask, for whom do these facts constitute good evidence? For us? For God? For the subject in question? The answer is, perhaps these facts do not constitute good evidence for anyone. In saying that Alfred's belief that it is snowing on the basis of his reasons is, given his relevant characteristics, a reliable indication that it is snowing, we do not imply that these facts are a reliable indication for anyone. It may be that no one is aware of these facts, not even Alfred himself. To this degree ascriptions of reliability are objective and independent of whether anyone knows or believes them to obtain. We may note that these things are also true of barometers. A deserted barometer's indication of barometric pressure may be every bit as reliable as those of a barometer in daily use.⁶

We can now see that it is not straightforwardly circular to define reliability in terms of inductive probability. Inductive probability measures the degree of evidential support that certain facts/propositions confer on a hypothesis *h*. This is a logical relation that is not indexed to a particular person. Presumably Swain's account takes its impetus from the logical interpretation of probability developed by Carnap.

Swain's theory would be straightforwardly circular if he were to analyze reliability in terms of epistemic probability. This latter notion is always indexed to a subject. Where *h* on *e* has high inductive probability,

6

Ibid., pp. 101-102.

(i.e., where e evidentially supports h) h is epistemically probable for S just in case e belongs to S's stock of knowledge. Epistemic probability is closely akin to epistemic justification and an analysis of the latter in terms of the former would be straightforwardly circular. Swain's inductive probability analysis would turn out to be circular if the locution 'e evidentially supports h' were analyzed as 'e would justify h for some S if S were to believe h on the basis of e.' Seeking to avoid this result Swain leaves the notion of evidential support undefined. He must intend that we rely on our intuitive notion of evidential support presupposing that this notion is separable from an intuitive notion of 'facts/propositions that would provide a subject with justification.'

Whether or not this presupposition is correct, I think we can see that Swain's theory begins to look rather peculiar. He is analyzing epistemic justification in terms of the reliability with which a subject with certain characteristics holding a belief indicates that truth of that belief. Reliable indication is analyzed in terms of inductive probability, i.e., in terms of the evidential support that a subject S with certain characteristics holding a belief h provides for h. Clearly h would be epistemically probable for a subject S' who was aware of the evidential support provided by his beliefs and characteristics, i.e., S would be epistemically justified in

believing *h* on the basis of this evidence. But Swain holds that the mere inductive probability of *h* on this evidence is sufficient for *S* to be justified in believing *h*. The mere fact that *S* believing *h* with characteristics *C* evidentially supports *h* makes *S*'s belief justified, even if *S* is unaware of this evidence. I suggest that this is strongly counterintuitive.

So far, the objection to Swain is the standard objection to reliabilist or externalist theories of knowledge. As one would expect, externalists like Swain report externalist intuitions. Presumably Swain would find nothing untoward in the result we just noticed. The peculiar feature of Swain's theory is that the standard foundation for reliabilist intuitions is missing. Generally, a theory that made the justification of *h* for *S* a function of evidential support provided by *e* would require that *e* be part of *S*'s knowledge. Reliabilist theories can be seen as eschewing altogether this notion of evidential support. For example, Goldman's theory makes the justification of *h* for *S* a function of the statistical probability of beliefs generated by the cognitive process that generate *h*. The theory is sustained on an intuitive level by the objective connection between justification and truth that the theory ensures. Justified beliefs are more often true than false. One may object to such a

theory on the grounds that a subject's belief can be reliable in this sense without the belief being supported by evidence. But Goldman can always fall back on the truth connection ensured by his theory as a foundation for his intuition. He can hold that considerations of evidence are not as important for epistemic justification as considerations of truth frequency. I would disagree with Goldman but the disagreement would be systematic.

The situation with Swain is considerably different. He endorses the view that evidential support is constitutive of epistemic justification. However he denies that the evidence which makes S justified in believing h must be possessed by S himself. If Swain wants to embrace (what seems to me to be) the counterintuitive consequences of his theory, he cannot appeal to some extra-evidential notion like truth frequency to support his view. There is no guarantee that evidential support is connected to truth frequency. Thus Swain is in the peculiar position of accepting the framework of evidential support while rejecting one of its central presuppositions viz., that evidence for h justifies S in accepting h only if the subject in some sense possesses that evidence. Of course there is nothing inconsistent in Swain's position. It is just that his definition of 'reliability' robs this theory of any of the intuitive support generally delivered to

such theories. It places his theory in a framework that is hostile to the consequences of his theory.

Of course intuitions only go so far. Counterexamples to a theory are only effective in the context of shared intuitions. Nonetheless, I will present what I take to be a valid counterexample to Swain's theory as a necessary condition for justification.

Suppose S has perceptual evidence that there is a tree before him. The lighting is good and S has no reason to suspect that his vision is in any way defective. However, unbeknown to S, a hallucinogenic drug was slipped into his morning coffee. It seems clear that S believing there is a tree before him on the basis of being appeared to 'treely' does not evidentially support the proposition that there is a tree before him when it is conjoined with the proposition that he has ingested a hallucinogenic drug. Thus on Swain's theory, S is not justified in believing that there is a tree before him.

Plainly this cannot be correct. If S has evidence sufficient to justify the proposition that there is a tree before him, the mere fact that unbeknown to him there is evidence against this proposition should not affect his justification. We can even suppose that S had a chemical analysis of his coffee done by a local

laboratory and it was mistakenly certified as non-hallucinogenic. Now, S has evidence against the fact that he took an hallucinogenic drug -- a fortiori this fact should not affect his justification.

Intuitive Considerations Regarding Sufficiency

We have seen that Goldman's reliable process view is too weak. It is easy to see that Swain's reliable indication theory suffers from the same defect. As we noted in connection with Goldman, if we suppose that intuitively non-justification conferring processes are reliable, it does not seem to follow that they become justification conferring. Consider again the world W in which a benevolent demon insures the reliability of wishful thinking. If S believes that h on the basis of his reason, "I wish that h", and S's characteristics include being under the influence of the benevolent demon, it would seem that S's belief that h on the basis of his reasons and given his characteristics is a reliable indication of the truth of h. Thus, on Swain's view, S would have a justified belief that h. And we saw in connection with reliable process theories that there is no plausible revision to a reliabilist account that handles the problem of this counterintuitive result.

Also in connection with reliable process views, we noted that a difficulty arises regarding the role of undermining evidence. The reliability of a cognitive process is not diminished by the presence of undermining evidence. However, the normativity of epistemic justification dictates that one who fails to take into account such evidence cannot thereby acquire a justified belief.

According to Swain, the epistemically relevant probabilities are conditional on the subjects other beliefs. Thus it may seem that Swain can account for the role of undermining evidence. Swain suggests this when he claims:

...the PR model of justification is fully sensitive to the role a subjects evidence plays in questions of justifications. To say that a person has evidence for some proposition is at least to say that this person has some beliefs such that the propositions believed (the evidence) bear some appropriate evidential relation to the proposition for which they are evidence. . . the PR model of justification requires that evidential beliefs of this sort be epistemically justified themselves if they are to contribute to the justification of other beliefs.⁷

Contrary to Swain, I believe that his model cannot fully account for the role of a subject's evidence. The first problem is that Swain's model can explain the role of an evidential belief *b* only if *b* is part of the basis for the belief that *h* (where *h* is the belief to be justified). Swain correctly maintains that *be* can contribute to

7

Marshall Swain, Reasons and Knowledge (Cornell: 1981) p. 102.

the justification of h only if h is based on b . But in addition to contributing to the justification of h , b might undermine the justification for h , and b need not form part of the basis for h in order to have this effect. To recall an earlier example, if S believes his vision is defective, then he cannot acquire justified visual beliefs. But his belief about his vision need not form part of the basis for a visual belief. Thus Swain's model cannot account for the role the belief has in undermining the justification for the visual belief.

As I suggested earlier, Swain can probably handle this difficulty by including the background beliefs as part of the relevant characteristics of the subject that condition the probability of the belief h . So it would be a relevant characteristic of S that he believes that his vision is defective. Nonetheless, a further difficulty emerges. S 's holding a particular belief b will only affect the probability of h in the requisite way, if b itself is reliable. The mere fact that S holds the belief that his vision is defective will not make it improbable that his visual belief is false. S must reliably believe his vision is defective. But suppose S does not believe b reliably. This would not prevent b from upsetting S 's justification for h . Surely an unreliable background

belief can make S unjustified where S would otherwise be justified. But Swain's model is unable to explain this fact.

Finally Swain's reliable indication theory is faced with the problem that a person's evidence need not always take the form of an evidential belief. Thus a theory that takes the view that the justificational status of a belief is purely a function of the reliability with which that person's beliefs indicates the truth of the proposition believed cannot be correct. But this is precisely what a reliable indication view presupposes. Swain claims, "To say that a person has evidence for some proposition is at least to say that this person has some beliefs such that the propositions believed (the evidence) bear some appropriate evidential relation to the proposition for which they are evidence." Even if it is true that a person has evidence for a proposition h only if he believes some proposition which bears an evidential relation to h, it is clearly not the case that in order for a person to have evidence against a proposition h he must have some corresponding evidential belief. A person can be presented with evidence against a proposition which he arbitrarily ignores and thereby fail to acquire the appropriate evidential belief. This does not change the fact that the person has evidence against h. For example,

an optometrist may misleadingly tell S that his vision is defective. If S does not want to believe this for emotional or capricious reasons he may simply ignore what the optometrist said. (S could believe that the optometrist said what he did but also believe that the optometrist was lying; or S could simply block out the entire incident from his memory.) Surely S would not thereby escape the epistemic consequences of the evidence presented by the optometrists testimony, viz., his visual beliefs are not justified. But as I have set up the case, S has no beliefs that explain why he is deprived of his justification. And since Swain's reliable indication view can only account for the role of evidence if there is a corresponding evidential belief, it would yield the result that S can preserve his justification for a belief simply by ignoring relevant counter-evidence.

Logical Considerations

As I noted earlier, intuitions only go so far. Thus I will present some logical objections to Swain's theory that I believe are problematic even for one with reliabilist intuitions.

Swain's theory conditionalizes inductive probability on the relevant characteristics of S. The problem becomes how to restrict what counts as a

characteristic of S. Broadly construed, S's characteristics include such things as "truly believing h" "falsely believing h", "having a friend who truly believes h" etc. Since Swain's theory conditionalizes the probability of h on S's characteristics, these sorts of characteristics must be excluded if the theory is to avoid the consequence that all and only true beliefs are justified.

Swain anticipates the problem and incorporates a specific provision in his theory to handle it:

(4.2) A characteristic, c, of a person is relevant to determining whether S's belief that h is epistemically reliable only if S's having c entails neither that h is true nor that h is false, unless h is such that the probability that h, given any evidence, e, is 0 or 1.⁸

As Swain notes, (4.2) is inadequate. Certain characteristics of S, while not entailing h or not-h, might nonetheless guarantee the probabilistic facts that determine whether h is justified for S. These characteristics include, "reliably believing that h", "unreliably believing that h", "being next to Jones who reliably believes that h" etc. Thus Swain replaces (4.2) with:

(4.3) A characteristic, c, of a person, S, is relevant to determining whether S's belief that h is epistemically reliable only if S's having c is probabilistically neutral with respect to h when taken alone, unless h is such that the probability that h, given any evidence, e, is 0 or 1.⁹

8

Marshall Swain, Reasons and Knowledge (Cornell: 1981) p. 9¹⁰⁶.

Ibid., p. 109.

The crucial notion in (4.3) is "probabilistic neutrality". Swain explicates it thus:

...a characteristic, *c*, of a person, *S*, is probabilistically neutral with respect to whether *h* if and only if the probability that *h* given tautologous evidence, *T*, is equal to the probability that *h* given *T* and given that *S* has *c*.¹⁰

The troublesome characteristics cited above appear to be ruled out by the probabilistic neutrality requirement. And as Swain points out, the characteristics that intuitively should matter are not ruled out:

Those characteristics that do seem relevant, however, also seem to be characteristics that are neutral when taken alone. In the original example involving Alfred, one such characteristic might be that of having normal eyesight. Although it is clear that this characteristic is relevant to determining whether Alfred reliably believes that it is snowing, it is, when taken alone, probabilistically neutral to whether it is snowing.

Whatever value we would assign to the probability that it is snowing, given only tautologous evidence *T*, we would assign precisely the same value to the probability that it is snowing, given *T* and given (only) the additional information that Alfred has normal eyesight.

Although the characteristic of having normal eyesight is probabilistically neutral when taken alone, Alfred's having this characteristic can have an effect on the conditional probabilities when taken in conjunction with the other conditioning evidence referred to in (4.1). Given that Alfred believes that it is snowing on the basis of reasons that include various perceptual states, the added information that Alfred has normal eyesight is positively relevant (probabilistically) to whether it is snowing. On the other hand, if Alfred's eyesight were poor and uncorrected, or if Alfred were a poor discriminator of snow from other forms

10

Ibid., p. 108.

of precipitation, these characteristics would be negatively relevant. All these characteristics however when taken alone, are probabilistically neutral to whether it is snowing.¹¹

The probabilistic neutrality restriction involves the determination of the effects of characteristics, of S, considered individually and in sets. This requires that the characteristics of a person be individuated in a fairly precise way. The problem is that the characteristics of a person can be individuated in a variety of ways. But what counts as one characteristic relative to one individuation scheme, may count as more than one characteristic relative to another. The probabilistic neutrality restriction seems to give the right result for the cases Swain discusses, but the success of the restriction hinges on the particular individuation scheme that Swain uses. Suppose we individuate S's characteristics as follows:

- c_1 : having a true belief acquired at t.
- c_2 : having acquired the belief that h at t (and no others).

Anytime S has a true belief that h, S will have c_1 and c_2 for some t. Both c_1 and c_2 meet the probabilistic neutrality requirement. Where h is, e.g., "It is snowing," the value we would assign to the probability that h, given

¹¹ Ibid., p. 108-109.

tautologous evidence T would not change if we add that S has c_1 , or that S has c_2 to that evidence. Of course the value of $\text{pr}(h)$ would increase if we were to add c_1 and c_2 , but as Swain points out, relevant characteristics will not be probabilistically neutral in sets. Since c_1 and c_2 are relevant characteristics as the theory defines this notion, and since anyone who has a true belief that h will have c_1 and c_2 , having a true belief that h will be sufficient for having a justified belief that h . Similarly, where S falsely believes that h , S will have analogous relevant characteristics that will insure that only true beliefs can be justified. Thus, Swain's theory falls victim to the same untoward result as does Goldman's theory, viz., all and only true beliefs are justified.

An additional problem remains for the reliable indication theory as providing a sufficient condition for justification--the lottery paradox. Swain recognizes this problem in connection with the initial formulation of his account and attempts to amend it in order to resolve the difficulty. Before discussing the proposed solution, we need to see just how the paradox arises for Swain's theory.

We begin by noting that Swain endorses the following principle of consistency for epistemically reliable belief:

(PEC) If, for each h_i in the set or propositions $[h_1, h_2, \dots, h_n]$, S 's believing that h_i on the

basis of R would be epistemically reliable, and if ' $h_1 \cdot h_2 \dots h_n \cdot q$ ' is inconsistent, and if S has some set of reasons, R^* , such that S's believing that ' $h_1 \cdot h_2 \dots h_n \cdot q$ ' is inconsistent on the basis of R^* would be epistemically reliable, then it is not the case that S's believing that q on the basis of R would be epistemically reliable.¹²

Let us assume that S reliably believes that the lottery is fair, that there are 1000 tickets in the lottery, and that exactly one ticket will win. On the basis of these reasons, S comes to believe that for each ticket t_i , t_i will lose. These will be reliable, since given S's reasons, for each t_i , it is more probable it will lose than it will not lose. Let q be the proposition that exactly one ticket will win. By assumption S reliably believes that q. Moreover on the basis of S's reasons R^* (presumably some propositions about the logical relations among the propositions of the following conjunction) S's believing that ' $t_1, t_2, \dots, t_{1000}$ will lose and q' is inconsistent would be epistemically reliable. But then (PEC) entails that it is not the case that S's believing that q is epistemically reliable. But this contradicts the assumption that S reliably believes that q.

To avoid this result Swain modifies his definition of epistemically reliable belief. He begins by tentatively

12

Ibid., p. 125.

introducing a definition of epistemic competition:

(4.13) h' is an epistemic competitor of h relative to person S , reasons R , and relevant characteristics C iff:

either (1) h and h' are such that the probability of each, given the usual facts about S , is greater than 0 and less than 1, and h' is negatively relevant to h given (i) S believes h on the basis of R and (ii) S has C

or (2) h is covered by (2b) of (4.5) and h' is equivalent to the denial of h .¹³

(Part (2) is designed to handle non-contingent propositions and will not concern us here.) This definition employs the crucial notion of negative relevance. Roughly h' is negatively relevant to h iff the $pr(h)$ is greater than $pr(h/h')$. According to this definition, all the propositions of the form ' t_i will lose' compete with each other. Swain then revises his definition of epistemic reliability to require that a reliably believed proposition outstrip all its competitors. This prevents any of S 's beliefs of the form ' t_i will lose' from being epistemically reliable.

As Swain notes (4.13) is unsatisfactory due to a technical problem discovered by Keith Lehrer. Lehrer shows that (4.13) is too weak in that it allows propositions to count as competitors of h which intuitively are irrelevant to the justification of h . Suppose r is a

13

Ibid., p. 131.

proposition that has a higher conditional probability than h but intuitively is irrelevant to h . Suppose q , a proposition with a lower conditional probability than h is negatively relevant to h . Then the proposition ' $r \vee q$ ' will be negatively relevant to h . Since r entails $r \vee q$, $\text{pr}(r \vee q)$ is greater than or equal to $\text{pr}(r)$. Thus the justification for q will be undermined by a proposition irrelevant to it. For in fact, it is the probability of r that undermines q .

Swain discusses an approach to solving this problem that he believes Lehrer dismisses prematurely. Lehrer argues that we cannot exclude as competitors to h propositions that are equivalent to disjunctions one of whose disjuncts is irrelevant to h . He demonstrates that this would exclude all propositions from being competitors¹⁴ to h except contradictions and contraries.

Swain suggests a way to weaken the restriction to avoid this consequence: "...the restriction that h ' [a competitor of h] not be equivalent to a disjunction of propositions one of whose disjuncts r has the following two characteristics: (a) r is irrelevant to h , given the usual facts about the subject; and (b) r is such that the conditional probability of r , given those facts about the

14

Ibid., p. 194.

subject, is greater than or equal to the probability of h , given those facts."¹⁵

This proposal would prevent a competitor from defeating h simply because the probability of an irrelevant disjunct equals or exceeds the probability of h . Propositions with irrelevant disjuncts can still be competitors of h on this restriction. But no harm results, Swain contends, since the restriction guarantees that the probability of the irrelevant disjunct must be less than the probability of h .

Although Swain's proposal sufficiently weakens the restriction to avoid the difficulties Lehrer discovered, we can see that in fact Swain's proposal is too weak. Let h' be a proposition negatively relevant to h , where $\text{pr}(h')$ is less than $\text{pr}(h)$. Suppose h' is disjoined with a proposition r that is irrelevant to h where $\text{pr}(r)$ is less than the $\text{pr}(h)$. Swain's restriction allows that the disjunction $h' \vee r$ is a competitor of h . But suppose the probability of r is sufficiently high so that $\text{pr}(h' \vee r)$ is greater than or equal to $\text{pr}(h)$. We would then have a case where an intuitively irrelevant proposition is responsible for undermining the justification for h . Thus a version of Lehrer's original problem emerges.

Notice that Swain cannot strengthen his restriction to exclude as competitors propositions equivalent to

¹⁵ Ibid., p. 130.

disjunctions one of whose disjuncts r is (1) irrelevant to h ; and (2) is such that the probability of r is responsible for increasing the probability of the entire disjunct to a value greater than or equal to h . Although this would escape Lehrer's original problem, it would fall victim to a version of Lehrer's second objection. As Lehrer notes, every genuine competitor h' that is not either a contradictory or contrary of h is equivalent to a disjunction of the form $(h' \cdot r) \vee (h' \cdot \neg r)$ where r is some proposition that offsets the negative relevance of h' . Now, suppose $\text{pr}(h' \cdot \neg r)$ is less than $\text{pr}(h)$, and $\text{pr}(h' \cdot r) < \text{pr}(h)$. Since (by the probability calculus) $\text{pr}(h' \cdot r) \vee (h' \cdot \neg r) = \text{pr}(h')$ (since they are logically equivalent), if $\text{pr}(h')$ is greater than or equal to $\text{pr}(h)$, it follows that $\text{pr}(h' \cdot r)$ is responsible for increasing $\text{pr}(h' \cdot r) \vee (h' \cdot \neg r)$ to a value greater than or equal to $\text{pr}(h)$. But $h' \cdot r$ is irrelevant to h (since r offsets the negative relevance of h'). As such, Swain's restriction precludes h' from being a genuine competitor of h . But by assumption h' is a genuine competitor of h . Thus Swain's proposal is too strong.

Perhaps Swain could invent another ad hoc restriction that would escape this latter problem. But as it now stands, it is fair to conclude that his analysis does not escape the lottery paradox.

CHAPTER 3

THE DOXASTIC CONNECTION

If my arguments are correct, then epistemic justification is not conceptually connected with objective truth-frequency. We need not conclude from this that justification and truth are in no way conceptually related. In fact, the foregoing considerations suggest an alternative account. The salient feature of those considerations seemed to be this: From the perspective of epistemic responsibility, a person S, can have an impeccable belief that certain conditions make the truth of a proposition, B likely. And under these circumstances S can be justified in believing B, even if those conditions do not in fact make the truth of B likely. As such, we might consider the possibility that the connection between justification and truth is to be found at the subjective or doxastic level. To put it schematically, one might propose that it is a conceptual truth that if C justifies B for S, then C entails that S believe that certain conditions obtain which make it probable that B is true.

This sort of account is congenial to coherence theories of justification.¹ Although coherence theories vary considerably, it is not unreasonable to cite as a characteristic feature, the requirement that the justification of any belief be a function of its relation to other beliefs.² So if the justification of a belief B for S entails that S have some other belief that certain conditions obtain which make the truth of B probable, then any justified belief will owe its justification to a relation it bears to another belief.³

As a representative of this general approach, we can examine a theory of justification advanced by Keith

1

Explanatory coherence theories would construe the connection between C and B as "C explains B." See Gilbert Harman, Thought (Princeton, 1973). I believe that what I say about probabilistic coherence theories will apply to explanatory coherence theories as well.

2

Because certain Foundation theories are cast in terms of prima facie justification, this characterization is not quite right. For prima facie justified beliefs will have to bear the relation of not being defeated to other beliefs if they are to be completely justified.

3

Certain Coherence theorists also require that the justification conditions for a belief include an objective connection to truth. See Keith Lehrer, "Knowledge, Truth and Ontology" (forthcoming) (p. 17) and Lawrence Bonjour, "The Coherence Theory of Empirical Knowledge" Philosophical Studies (1976) pp. 293-294. Insofar as they require this they are subject to the same objection that can be raised to Reliabilism.

⁴ Lehrer. For Lehrer, justification resides in the coherence of propositions accepted for the purposes of attaining truth and avoiding error. He analyzes the coherence relation in terms of probability.⁵ Ignoring certain sophistications, a proposition B is justified for S iff S accepts that given the system of propositions which he accepts in the interest of obtaining truth and avoiding error, B has a higher probability of being true than any statement with which it competes.⁶ Thus S is justified in believing B only if S accepts that certain conditions R (viz., those conditions described by the beliefs in S's acceptance system) make it more probable that B is true than any proposition that competes with B. This entails that S accepts that R makes it more probable than not that B is true since one of the competitors of B is \sim B. Here then is a representative of the view that a connection between justification and truth lies at the subjective or doxastic level.

⁴ In his book Knowledge (Oxford: 1974) Lehrer seems to endorse only a subjective connection to truth. In his later work, "Keith Lehrer--A Self Profile" in Radu J. Bogdan, Keith Lehrer (Dorchester: 1981) (pp. 79-85), and "Knowledge, Truth and Ontology" ibid (pp. 16-18) Lehrer endorses an objective connection as well.

⁵ Lehrer refers to probability as an objective feature of the world, e.g., a propensity. "Knowledge, Truth and Ontology" (p. 11).

⁶ Ignoring a technical emendation, P competes

While this view would seem to give us some insight into the nature of epistemic justification, a major difficulty lies in its reliance on an intellectualist⁷ model of justification. Recall that on this view, in order for S to be justified in believing B, S must believe that certain conditions obtain which make the truth of B probable. One might object that these supporting beliefs are not always necessary for justification to arise. Cases involving perceptual knowledge are often invoked in behalf of this weaker claim. Suppose that S perceives something which is red and so comes to believe "There is something red." Some philosophers have maintained that it is sufficient for S to be (at least prima facie) justified in this perceptual belief that S merely be⁸ appeared to redly or that S merely believe that he⁹ perceives something to be red. The proponent of the intellectualist model would argue that the mere fact that S is appeared to redly cannot justify a belief for S that

with Q if $\text{prob}(P/Q) < \text{prob}(P)$.

7

The term "Intellectualist Model of Justification" is used by Ernest Sosa in "The Raft and the Pyramid," Midwest Studies in Philosophy, ed., French, Vehling, and Wettstein (University of Minnesota: 1980) (pp. 3-26).

8

John Pollock, Knowledge and Justification (Princeton: 1974) (pp. 58-64).

9

Roderick Chisholm, Theory of Knowledge, 2nd edition, (Englewood Cliffs: 1876) (p. 78).

something is red. S must believe he is appeared to redly and moreover believe that his being appeared to redly makes it likely that there is something red. Similarly, the mere fact that S believes he perceives something to be red cannot justify the belief in question for S, unless S believes that his believing he perceives something to be red makes it likely that he does perceive something to be red. (i.e., S must believe that his perceptual beliefs are reliable.)

Although the claims of the intellectualist have some intuitive appeal, -- one feels that something like this must be involved in justification -- the strongest argument against them seems to be that they run the risk of skepticism. For in many cases of what we generally take to be e.g., perceptual knowledge, the psychological reality of the supporting beliefs required by the intellectualist model is questionable. There is no obvious sense in which most adult persons, not to mention children and animals have beliefs about how they are appeared to as well as beliefs about their reliability as perceivers, when they have perceptual beliefs.

To avoid skepticism, the defender of the intellectualist has two alternatives. He can argue that the relevant supporting beliefs are present unconsciously or he can argue that the beliefs are present dispositionally.

The first alternative is ambiguous between two ways in which beliefs can be said to be unconscious. According to psychoanalytic theory, beliefs are unconscious insofar as a certain psychological mechanism ceases to operate, then the belief surfaces in consciousness. It would be peculiar to claim that the beliefs required by the intellectualist model are unconscious for this reason.

It may also be claimed that certain beliefs are unconscious by appealing to models of human cognition that posit unconscious inference. The premises of these¹⁰ inferences may be construed as unconscious beliefs.

Could the beliefs required by the intellectualist model be unconscious in this way? Plainly an affirmative answer to this question cannot be based solely on the requirements of the intellectualist model. Cognitive psychologists currently view human cognition on the model of an information-processing mechanism. If the intellectualist seeks to appropriate a computational model of human cognition, he must provide some independent reason for supposing that such a model will be consonant with his own. At present, the details of these models are both sketchy and controversial. Thus, whether the

10

Cf. Gilbert Harman, op. cit. Harman construes the supporting beliefs as explanatory rather than probabilistic. (Chapters 7 and 8).

intellectualist can forge a doxastic connection out of the states posited by these psychological models is not a matter that can be glibly assumed. Rather it is a matter for further empirical research.

The second alternative for the intellectualist is to argue that the requisite beliefs are psychologically real in some dispositional sense. The problem with this approach is the difficulty in determining just when a dispositional belief is present. The clearest cases of dispositional belief are those in which the belief was previously occurrent and is presently stored in memory. But, obviously this won't do in perceptual cases. This cannot be the sense in which beliefs like "S is appeared¹¹ to redly" are dispositional. It seems that the intellectualist must claim that the supporting beliefs are dispositional in the sense that if the subject were to consider the relevant propositions, he would assent to them. So if S is justified in believing that there is something red before him, S would assent to the proposition that he is appeared to redly (or the proposition that he believes he sees something red), and the proposition that this fact makes it likely that that there is something red before him, if he were to consider these propositions.

¹¹

This point is made by Pollock op. cit. (p. 58).

Is this a psychologically more plausible picture? Again, it is not obvious that subjects have these required beliefs even in this dispositional sense. It is not always an easy matter to characterize the way in which you are appeared to. This is a skill that artists have to work at acquiring. But even if the required beliefs were present in this dispositional sense, it is easy to see that this sense of belief is too weak to subserve any epistemological end. Suppose it were true of S at (some time) t that S would assent to P if S were to consider P. This fact would not suffice for P to function as one of S's reasons for holding some belief at t . For it may be that S's assent to P would be prompted by his consideration of P. By reflecting, people can discover reasons for beliefs they hold. Such reasons cannot be said to have justified those beliefs prior to their discovery. For example, S may believe that it is going to rain on the basis of the pronouncements of a ouija board. Now it may be that S is a skilled meteorologist who has neglected to consider the proposition that the present weather conditions portend the approach of a rainstorm. Even if it were true that S would assent to this proposition if he were to consider it, this does not constitute grounds for saying that S is justified in believing that it is going to rain.

At this point, the intellectualist may simply want to insist that in those cases where the relevant beliefs are lacking, the subject is not justified. If and to what extent this position yields skepticism depends on the empirical issue of the extent to which those beliefs are lacking in actual cases. As a last resort, the intellectualist model can be viewed as an ideal to which as a matter of empirical fact, we only approximate. To take this tack, the intellectualist would, to a certain extent have to impugn some of our common sense judgments about when our beliefs are justified. This position would involve a kind of "soft" skepticism. Strictly speaking, a lot of the beliefs we think are justified are not. But in some sense, the beliefs are justified insofar as they approximate¹² the ideal model.

Admittedly this is a vague position which places the burden squarely on the shoulders of the intellectualist. What is needed is an account of the sense in which our belief systems do approximate the ideal model and an explanation of how this can in some sense confer justification on those beliefs. I think this would prove to be a difficult task.

12

This position is taken by Lawrence Bonjour, "Externalist Theories of Empirical Knowledge" in *Midwest Studies op. cit.*

CHAPTER 4

EPISTEMIC JUSTIFICATION WITHOUT A TRUTH CONNECTION

To reject the intellectualist model is to reject the thesis that justification is connected to truth at the doxastic level. Since we have previously seen that there is reason to deny that any objective connection exists, we may wonder what remains of the thesis that justification and truth are connected. It can still be maintained that S is justified in believing that B only if S is justified in believing that B is true. But as we noted earlier, this trivial connection does not seem to capture the basic intuition that justification must be connected to truth. Can one build a theory of justification that employs only this trivial connection? In fact, this is how one might plausibly construe certain Foundations theories. These theories typically start by sanctioning common sense judgments concerning when our belief are justified. They then proceed to posit a series of epistemic principles which validate these judgments by specifying just how those beliefs are justified. It is important to see that epistemic principles do not specify any connection between justification and truth beyond the trivial one.

Consider an epistemic principle proposed by¹
 Roderick Chisholm:

- (C) For any subject S, if S believes without ground for doubt, that he is perceiving something to be F, then it is evident for S that he perceives something to be F.

John Pollock has proposed:²

If S is appeared to redly, then S is prima facie justified in believing that there is something red before him.

1

Theory of Knowledge (p. 78). In "A Version of Foundationalism" Midwest Studies op. cit., Chisholm seems to acknowledge some form of nontrivial truth connection. After pointing out the trivial connection he writes:

There is still another point about the relation between epistemic justification and truth...if I want to believe what is true and not to believe what is false, then the most reasonable thing for me to do is to believe what is justified and not to believe what is not justified (p. 545).

Chisholm must intend the expression "the most reasonable thing for me to do" to be taken in a pragmatic rather than an epistemic sense. Otherwise the statement would be trivial. It is still unclear what Chisholm means here. At any rate, no truth connection is expressed in Chisholm's principles beyond the trivial one.

2

This principle is adapted from remarks Pollock makes in chapter three Knowledge and Justification.

These principles simply state that under certain conditions (stated in the antecedent) a certain belief (stated in the consequent) is justified. But the conditions described in the antecedent do not involve either an objective or a doxastic connection to the truth of the belief described in the consequent. Pollock is quite explicit about this. For him, epistemic principles are true in virtue of a meaning connection. But this connection is not between a statement and its truth conditions, rather the connection³ is between a statement and its justification conditions. Of course in each case the epistemic principles describe conditions under which S is justified in believing that P is true, but this is just the trivial connection.

Is a theory that employs only this trivial connection adequate? I believe we can see that it is not. This can be most clearly brought out by considering the⁴ details of Pollock's theory. For Pollock, epistemic principles describe conditions under which beliefs are prima facie justified. To say that a justification is prima facie⁵ is to say that it can be defeated under certain conditions. Pollock characterizes two types of defeaters

³ Op. cit. Chapter 1 (especially p.11-12). Chisholm is not explicit about this point.

⁴ What follows is from Chapter 2 op. cit.

⁵ Chisholm does not propose principles of prima facie justification. He handles defeaters by adding

(construed propositionally) for P being a prima facie reason for S to believe that Q. Type I defeaters are reasons for S to believe that Q is false. Type II defeaters are reasons for believing that the truth of P is not an indication of the truth of Q, independently of being a reason for believing that Q is false. So take Pollock's principle:

If S is appeared to redly, then S is prima facie justified in believing that there is something red (before him) (i.e., 'being appeared to redly' is a prima facie reason to believe "there is something red.")

The proposition "S is in a room with no red objects" is a type I defeater of this prima facie justification, since it is a reason for S to believe it is false that "There is something red." The proposition "S is in a room with a red light" is a type II defeater. Although it is not a reason to believe "There is something red" is false, it is a reason to believe that the truth of "S is appeared to redly" is not an indication of the truth of "There is something red (before S)."

The issue I want to raise is whether a theory like Pollock's can account for defeaters of prima facie justification. Type I defeaters are explainable by the trivial justification - truth connection. If S has a

clauses directly to the epistemic principle. See Theory of Knowledge, 2nd edition, (pp. 75-76).

reason to believe that Q is false, then it's clear why S is not justified in believing that Q is true. But matters are not so clear for type II defeaters. There is no doubt that Pollock has correctly identified a source of defeat for Prima facie justification in type II defeaters. If S has reason to believe there is a red light shining on X , then the fact that X appears redly to S does not justify S in believing that X is red. The difficulty is that there seems to be no way for a theory like Pollock's to account for the fact that type II defeaters defeat. Where P is a prima facie reason for S to believe that Q , essentially type II defeaters attack the connection between P and the truth of Q . But we have just seen that a theory like Pollock's does not require any such connection between P (the justification conditions for Q) and the truth of Q .⁶ If a theory requires that the

6

In Knowledge and Justification, Pollock does remark " P is a prima facie reason for S to believe-that- Q iff S is prima facie justified (in the sense of section 2.2) in believing-that ($P \Rightarrow Q$)" (p. 42) Pollock uses ($P \Rightarrow Q$) to symbolize the subjunctive conditional "it would not be true that P unless it were true that Q ". (p. 42) Section 2.2 tells us:

(2.2) " P is prima facie justified for S "
means "It is necessarily true that if S believes
(or were to believe) that P , and S has no
reason for thinking it is false that P , then
 S is (or would be) justified in believing that
 p " (p. 30).

I don't understand the basis for Pollock's remark, given his account of prima facie reasons in Chapters 1 and 2. Let P be " S is appeared to redly" and let Q be "There is something red before S ." Suppose S believes ($p = \sim Q$)

truth of P logically entail the truth of Q (like Descartes); or if a theory requires that P make the truth of Q probable (like Goldman's⁶); or if a theory requires that the subject believe that P makes the truth of Q probable (like Lehrer's); then such a theory can account for why the subject having a reason to believe that P is not connected to the truth of Q, defeats his justification for believing that Q is true on the basis of P. But in the absence of any such requirement the existence of type II defeaters is utterly mysterious.

It looks as if the existence of type II defeaters gives us good reason to think that any complete theory of epistemic justification must require a non-trivial truth connection. Now Pollock and Chisholm may want to claim that the existence of type-II defeaters is a primitive fact about epistemic justification.⁷ This raises the issue of what explanatory requirements we should place on a theory of epistemic justification. For the failure to explain type-II defeaters is an instance of a more general explanatory deficiency that certain philosophers have attributed to Foundations theories like Chisholm's and Pollock's. The objection is that those

because the Ouija Board tells him it is true. This does not give S a reason to believe it is false that $(P = \sim Q)$. But surely S is not justified in believing $(P = \sim Q)$.

7

Pollock has taken this position in conversation.

theories are not framed at a sufficiently general or abstract level to provide any insight into the nature of epistemic justification. According to Keith Lehrer:

The most important function [of a theory of justification], in my opinion, is to explain why certain beliefs are justified and others are not. Hence a theory of justification must be judged in terms of how well it explains this. A system of principles may be presented as a theory of justification when the principles are presented as ultimate . . . If a belief is justified according to the system, the only explanation for why it is justified is -- that is what the principles tell us. But the principles are unexplained.⁸

Chisholm and Pollock do not really provide analyses of justification in the sense that Goldman and Lehrer do. Their theories just amount to sets of epistemic principles, one concerning perception, one concerning memory etc. These principles tell us that beliefs of certain types are justified under certain conditions. But the principles are nowhere united by a general theory that explains why those beliefs are justified under those conditions. In a sense, Foundation theories do not tell us what justification is. This objection to Foundationalism is reminiscent of the complaints of Socrates who in the Platonic dialogues, repeatedly chides his dialectical opponents for merely citing instances of a particular concept rather than providing an account of the concept itself.

In all fairness to Chisholm and Pollock, it does not seem to be their intent to provide analyses of justification in the sense that Goldman and Lehrer do. Their project is to account for how knowledge is possible. Their epistemic principles are intended to serve this purpose. But certainly it would be desirable to have a general theory that would motivate epistemic principles -- a theory that would explain why, e.g., certain perceptual conditions yield justified beliefs.

My present purpose is to show that the truth-connection provides a basis for such an explanatory theory. This is illustrated by the theories of Goldman and Lehrer, both of which make essential use of a non-trivial truth connection. Those theories can explain the truth of epistemic principles by showing how the justifying conditions described in the antecedent are connected to the truth (constituted objectively or doxastically) of the beliefs described in the consequents. It will be helpful to consider an example. Consider Pollock's principle:

If S is appeared to ϕ -ly, then S is prima facie
justified in believing there is something ϕ
before him.

Presuming that Goldman would concede the truth of a principle very close to this, how would he explain its truth? Goldman holds that justified beliefs are produced

op. cit. (p. 2).

by reliable cognitive processes. Since perception is one example Goldman gives of a reliable process, he would explain the truth of the principle in those terms. If S' belief that there is something ϕ before him results from S' being appeared to ϕ -ly, then S' belief is produced by the reliable cognitive process - perception. As such S' belief is justified.

Lehrer as well would concede the approximate truth of Pollock's principle. How would he explain its truth? Lehrer holds that justification arises out of probabilistic relations within an acceptance system. Generally when one is appeared to ϕ -ly, one's acceptance system is such that the proposition "there is something ϕ " is assigned a higher probability than competing proposition. So on Lehrer's theory, it will generally be true that if S is appeared to ϕ -ly, S is justified in believing that there is something ϕ before him.

The argument would proceed in the same way for epistemic principles involving memory, introspection, et al. Thus Goldman and Lehrer claim a theoretical advantage over Foundationalists since they can in this way explain the principles that the Foundationalist must take as basic. And they achieve this advantage by constructing their theories on the basis of a non-trivial truth connection.

We seem to have come full circle. We began by noting that there is good reason to suppose that there is an internal connection between justification and truth. We then examined several proposals as to how to construe that connection and found them to be problematic. Finally we considered a theory that does not employ any such connection and found it wanting for precisely that reason.

What can we conclude. Intuitively it's hard to see how some truth connection could not exist. Being epistemically justified in believing a proposition is quite a different matter from being morally or pragmatically justified. And there is a strong temptation to say that the difference lies in the fact that unlike moral or pragmatic justification, epistemic justification is connected in some important way to the truth of the justified belief. Nonetheless it has turned out to be a difficult matter to say precisely what that connection comes to. While theories that ignore the connection incur an explanatory deficit, theories that are based on some initially plausible construal of the connection are problematic. This leaves open several possibilities: Perhaps a theory based on some other formulation of the truth-connection will avoid the shortcomings of those I have considered. On the other hand, perhaps an explanatory theory can be achieved without the use of a truth-connection, although it is hard to see how such a

theory could explain type-II defeaters. Finally we might suppose that there is no greater level of generality to be attained by a theory of epistemic justification than that provided by Foundations theories. There may be no more general truths about the nature of epistemic justification beyond those stated in the Foundationalist's epistemic principles.

CHAPTER 5
KNOWLEDGE AND INFORMATION

We have been exploring the connection between epistemic justification and truth. Fred Dretske has attempted to build a theory of knowledge that posits a very strong connection.¹ Since Dretske's theory is naturalistic, his analysis proceeds in terms of naturalistic analogues of epistemic concepts. The crucial notion in Dretske's theory is information which is intended to play the role of evidence. Close examination will reveal analogous problems to those we have already discovered.

Information is at root, an epistemic notion. Dretske tells us, "Roughly speaking, information is that commodity capable of yielding knowledge, and what information a signal carries is what we can learn from it . . . We say this [that a pamphlet contains information about how to probate a will] because we believe that someone (suitably equipped) could learn something about probating a will by consulting the pamphlet."²

¹ Like other naturalistic theories, Dretske attempts to avoid talking about epistemic justification.

² Ibid, pp. 44-45.

Information is thus intimately connected to truth. Dretske stresses that for a signal to carry the information that S is F, it must be true that S is F: "If everything I say to you is false, then I have given you no information...when [my utterance, "I have a toothache" is] false, it fails to carry the information that I have a toothache because it is incapable of yielding the knowledge that I have a toothache."³

These observations are descriptive of what Dretske refers to as 'the nuclear concept of information'. Dretske's goal is to develop an information-theoretic account of knowledge. Since information is an epistemic concept, Dretske's strategy is to provide a naturalistic theory of information and to employ the theory in a naturalistic account of knowledge. We must begin, then by examining Dretske's theory of information.

Dretske cautions us that the theory of information is not supposed to provide a definition of the word "information". The definition is something akin to the nuclear concept. Dretske is proposing a scientific theory, a theory that purports to tell us what information is ("...a more or less complete, precise, and systematic description of those entities and processes underlying

3

Ibid, pp. 44.

the phenomena of interest"). He calls his theory a semantic theory of information because it concerns the informational content of a given signal. It is presented in this concise formulation:

Informational Content: A signal r carries the information that s is f = the conditional probability of s 's being f , given r (and k), is 1 (but, given k alone, less than 1).⁴

In a footnote, Dretske says what he "means to be saying" by "the conditional probability of s being f (given k) is 1" is that "there is a nomic regularity between these event types, a regularity which nomically precludes r 's occurrence when s is not f ."⁵ The k variable stands for the knowledge the receiver of information may already have about the possibilities at the source.⁶ It will not be relevant to any of the issues discussed in the paper and so for simplicity's sake I will drop reference to it.

Dretske's theory of informational content amounts to the claim that a signal carries the

⁴ Ibid, p. 65.

⁵ Ibid, p. 245.

⁶ Dretske gives this example. There are four shells and a peanut is under only one of them. Suppose A knows that the peanut is not under one or two but B does not know this. If both A and B learn that the peanut is not under shell three, A receives the information that the peanut is under shell four but B does not.

information that s is f iff the occurrence of r nomically entails s is f . He points out that his theory captures the strict connection between information and truth since the theory insures that a signal cannot carry the information that s is f unless it is true that s is f . But it is the very requirement that gives rise to an obvious objection to Dretske's theory, viz., it appears to place much too strong a condition on a signal carrying information. To use one of Dretske's examples, the acoustic waves emanating from a radio speaker carry information about what is occurring at the radio station. But, presumably, their occurrence does not nomically entail the occurrence of any event at the radio station. One's percepts carry information about one's surroundings. But, again, one's percepts do not nomically entail anything about one's surroundings. At first blush, it looks as if Dretske's theory precludes the transmission of information in many cases where, intuitively, it seems correct to say that information is transmitted. Moreover, if knowledge is to be analyzed in information-theoretic terms, it looks as if Dretske's theory of knowledge will yield skepticism.⁷

7

The K parameter in the semantic theory will not help against skepticism. For K refers to knowledge the receiver already has. Thus even if it is claimed that where one knows s is f , one knows e.g. one is not hallucinating, this knowledge cannot be a factor in the probability determination. For this would not be knowledge one already has.

Since Dretske devotes an entire chapter to answering the charge of skepticism, I will defer discussion of the issue until I take up Dretske's theory of knowledge. Presently, we need to examine the considerations that lead Dretske to formulate his theory of information in such strong terms.

Dretske derives his semantic theory of information from communication theory which treats information in a purely quantitative way. Communication theory is silent on the issue of informational content. According to the theory, the amount of information contained in an event or state of affairs is a function of the reduction of possibilities. Imagine a roulette wheel with eight numbers. Suppose the marker is pointing to a particular number. Communication theory tells us that this state of affairs contain three bits of information, because it takes three "binary decisions" to reduce the eight possibilities to one. In order to see this, we can imagine that instead of spinning the roulette wheel, the number is chosen by three tosses of a coin. The first toss determines whether one through four, or five through eight are eliminated; the second toss (assuming the latter) whether five and six, or seven and eight are eliminated; and the third toss (assuming the latter) whether one or two is eliminated.

Of course, communication theory is concerned with the notion of one state of affairs (the signal) carrying information about another state of affairs (the source). The amount of information a signal carries about a source is defined by the probabilistic dependencies between the signal and the source--by how the occurrence of the signal reduces the possibilities at the source. A signal that reduces eight possibilities at the source to one, carries three bits of information about the source. A signal that reduces eight possibilities at the source to two, carries two bits of information etc. In order to derive the semantic theory of information from these purely quantitative considerations, Dretske considers how much information a signal must carry about a source in order to carry the information, expressed in terms of content, at the source. He concludes that (where s is some object at the source) if a signal carries the information that s is f , then the amount of information carried by the signal must be equal to the amount of information generated by s 's being f . So if s 's being f constitutes a reduction of n possibilities to one, the occurrence of the signal must also reduce those n possibilities to one.

Dretske claims this result is unavoidable because he takes, what he calls the xerox principle, to be axiomatic:

Xerox principle: If A carries the information that B, and B carries the information that C, then A carries the information that C.⁸

For Dretske, this is "a regulative principle, something inherent in and essential to the ordinary idea of information, something that any theory of information should preserve."⁹ Dretske tells us that the xerox principle is indispensable for the "flow of information." Without it, there could not be chains of communication.

Assuming the xerox principle, Dretske derives his conclusion by running a reductio and absurdum on the opposing thesis: Suppose a signal could carry the information that s is f without carrying as much information as is generated by s being f. Then A could carry the information that B and B could carry the information that C even though information is lost between A and B (The possibilities are not reduced to one). A problem arises because the information loss between the links in a communication chain is cumulative. And no matter how small the information loss between the individual links the amount of information lost between the source and distant links in the chain can be made considerable. But assuming the xerox principle, it follows that a signal (viz., a distant link) can carry information about a

8

Ibid, p. 57.

9

Ibid, p. 57.

source even though the amount of information it contains about the source is negligible (even though the signal only reduces the possibilities at the source by a negligible amount). Since this is absurd, it follows that a signal can carry the information that s is f only if it carries all the information contained in s being f .

With this result, it is a trivial step to the semantic theory of information. If a signal carries as much information about a state of affairs at the source as is contained in that state of affairs, it must eliminate all possibilities at the source except the actual state of affairs. Dretske claims that the only way to insure this is to require that the probability of the state of affairs conditional on the occurrence of the signal, equal one.

At this point, some comments should be made about the role of probability in Dretske's theory. Dretske seems to be relying on some version of the logical interpretation of probability. On this interpretation, the probability of a proposition is a measure of the proportion of possible worlds in which it is true. However, Dretske seems to be concerned only with nomic possibilities. As such, the probability of a proposition can be taken as a measure of the proportion of nomically possible worlds in which it is true. Dretske wants to make this

probability conditional on the occurrence of a signal. We can then take him to mean that the signal's occurrence must eliminate every nomically possible world other than those world in which 's is f' is true. Then the proportion of possible world remaining in which 's if f' is true will be one. Thus the conditional probability of s being f will be one. As we noted earlier, Dretske takes this to mean that the occurrence of the signal nomically entails s being f.

A peculiar feature of Dretske's theory is that it conflicts with some of his own examples. Dretske illustrates the principles of communication theory by discussing an example where a group of eight employees are told to select one from among the group of eight employees are told to select one from among the group to perform a task. The employees make their selection and write the name (e.g., Herman) on a slip of paper and deliver it to their employer. Dretske claims that the message or signal received by the employer (the slip of paper with 'Herman' written on it) carries three bits of information concerning the state of affairs at the source. (Herman was selected). Since the signal reduces the possibilities at the source to one, presumably Dretske would want to say that the signal carries the information that Herman was selected. But certainly it is not in virtue of any nomic

connection that the possibilities at the source are reduced to one. Given a note written by the employees with the word "Herman" on it, it does not follow nomically that Herman was selected. The connection between signal and source is clearly conventional rather than nomic. So it's puzzling that Dretske formulates his theory in terms of nomic connections.

Moreover its not at all clear what motivates Dretske to limit the informational connections to nomic connections. The intuitive idea is that a signal carries information about a source just in case the signal enables one to learn something about the source. And we are to understand the phrase "enable one to learn" in a very weak sense. Dretske informs us that "...the information contained in a signal [is not] dependent on the receiver's actually learning something from that signal. The recipient may not be able to decode or interpret the message."¹⁰ Dretske illustrates this point by imagining that the employees in the aforementioned example devise a secret code in which to communicate the results of the selection process. They decide to write 'Herman' if Shirley is selected, 'Shirley' if Doris is selected etc. In effect, the employees alter the conventions of note

10

Ibid, p. 57.

writing. If the employer is not apprised of the code, he will not learn who was selected from the message. Nonetheless, Dretske allows that the message carries the information that Herman was selected. He states, "One¹¹ simply has to know the code to extract the information". (Indeed Dretske could not make the presence of information dependent on someones actually learning something, since Dretske wants to use the notion of information to analyze the epistemic concept, "knowledge." If "information" were analyzed in terms of someone actually learning, this would render the epistemic analysis circular.)

Dretske's concept of information appears to allow for signals to carry information about a state of affairs provided simply that there exists some system or code that maps states of affairs onto signals which excludes many-to-one mappings. Whenever such a mapping exists, one can learn of the states of affairs from the signals, simply by learning the mapping. Such a mapping would seem to be a sufficient condition for the existence of information.

This result has serious implications for Dretske's philosophy of mind. For example Dretske seeks to account, ultimately, for the intentionality of cognitive attitudes in terms of the intentionality of the underlying informational structures. Dretske takes the intentionality of

¹¹ Ibid, p. 57.

informational structures to be a consequence of the fact that purely extensional correlations are not manifestations of nomic regularities. Thus the semantic theory of information will not sanction purely extensional relations as informational. But as we have seen, in light of the intuitive concept of information, the semantic theory is unduly restrictive. Relaxing the requirements in accord with Dretske's discussion of the intuitive concept yields the result that cognitive structures do not exhibit intentionality. Extensional relations seem to fit quite readily the requirements for information-carrying that we could glean from Dretske's discussion. If a signal is related only extensionally to a state of affairs, one can learn of the state of affairs from the signal simply by learning the extensional code.

Even with the weakened semantic theory of information, the problem of skepticism remains. Eventually we will see Dretske has to qualify the strict probability requirement. Unfortunately the qualification will undermine Dretske's theory of information and vitiate the theoretical advantages Dretske claims for his information-theoretic epistemology. But the discussion of skepticism must again be postponed until we have examined Dretske's epistemology.

Dretske's theory of knowledge is concisely stated as:

K knows that s is f = K's belief that s is f
is caused (or causally sustained) by the information that s is f.¹²

Dretske restricts the analysis to cases of de re belief. The analysis is aimed at cases of "...knowing of something that it is f where the something known to be f is fixed by perceptual (noncognitive) facts.¹³

One problem with this analysis that is immediately apparent is the requirement that information, an abstract entity be causally efficacious. Dretske handles this difficulty by noting that a signal carries information in virtue of having certain physical properties. The information is said to be causally efficacious in just those cases where the physical properties are causally efficacious.

Since Dretske's analysis is essentially causal, it is beset by the usual technical difficulties of specifying the nature of the causal and causally sustaining relationships. Dretske sidesteps this problem seeking to avoid a "tedious digression on the nature of causal relations, causal sufficiency, overdetermination, and counterfactuals."¹⁴

¹² Ibid, p. 86.

¹³ Ibid, p. 86.

¹⁴ Ibid, p. 90.

Nonetheless, the burdens of a causal theory of knowledge exceed those presented by the technical specifications. The essential task lies in cashing out the normative aspects of epistemic concepts like justification and knowledge in terms of naturalistic concepts like causation and in this case, information. The causal relations that are suitable must be carefully restricted lest the analysis be subject to easy counter-example. Suppose a signal carries the information that *s* is *f* in virtue of the signal being composed of sound waves of a certain frequency. These sound waves trigger a landslide that results in a certain subject, *k*, being struck by a rock. The concussion in turn causes *k* to believe that *s* is *f*. Clearly we would not want to count *k*'s belief that *s* is *f* as an instance of knowledge. The causal etiology of the belief is unsuitable. We can see why by returning to the theory of information.

Recall that Dretske notes that a signal can carry information even though the recipient of the signal is unable to extract the information--even though the recipient can extract the information only if he is aware of the connection between the signal and the state of affairs. This was made explicit by Dretske in the case of the employees who deliver the message to their boss in a specially devised code. What it takes for the employer

to extract the message, i.e., what it takes for the employer to come to know the message, is knowledge of the code.

Although Dretske does not discuss this point in the context of nomic connection, it does not take much to see that the issue is the same. Light rays from distant galaxies carry information about states of affairs in those regions of the universe. Those ignorant of astronomy can receive those signals without truly coming to know anything about those state of affairs. They are ignorant of the 'astronomical code' (the nomic connections) necessary for extracting the information. The astronomer, on the other hand, is sufficiently versed in the relevant nomic connections to enable him to look through a telescope and learn things about distant galaxies.

I think it is fair to conclude that knowledge arises from the receipt of information only if the recipient appreciates the nomic connections that define the informational relationship. This fact must be reflected in Dretske's theory if it is going to have any chance of succeeding. It is an epistemological truism that evidence provides one with knowledge only if one appreciates the evidence. If we view information as roughly a naturalistic analogue of evidence, we could say that the

process by which information causes belief must be a
¹⁵
 naturalistic analog of evidence appreciation.

In subsequent chapters Dretske proposes a theory of the structure of belief and an account of belief formation. In order to evaluate his epistemology, we need to examine this account. For Dretske, a belief is a structure having a certain semantic content, e.g. *t* is *f*. Ignoring a technical emendation Dretske claims:

Structure *s* has the fact that *t* is *f* as its
semantic content = *s* carries the information that
t is *f* in digital form.¹⁶

Roughly, a structure *s* carries the information that *t* is *f* in digital form, if *s* carries no more specific information about *t* than that *t* is *f*. This is to be contrasted with a structure *s* carrying information in analog form. If *s* carries the information that *t* is *f* in analog form, then the information that *t* is *f* is

15

Dretske comes close to acknowledging this himself:

The idea of information causing (or causally sustaining) belief is intended to capture what is worth capturing in the doctrine that for a person's belief to qualify as knowledge, these must not only be evidence to support it, the belief must be based on that evidence. (p. 91)

But in most cases, a belief cannot be based on evidence unless the subject appreciates the evidence.

16

Ibid, p. 177.

nested in more specific information about *t*. Dretske uses this example: A speedometer carries the information that a vehicle is traveling at 43 m.p.h. in digital form. At the same time, the speedometer carries the information that the vehicle is traveling at a velocity between 25 and 50 m.p.h. in analog form. This latter information is nested in the more specific information that the vehicle is traveling at 43 m.p.h.

Belief formation is the process of converting information encoded analogically to information encoded digitally. For example, sensory information is initially stored in analog form. A sensory representation of a tree is rich in detailed information about the size, shape, texture etc. of the distant object. The information that *t* is a tree is nested in this more specific information. The belief that *t* is a tree emerges from a process that digitalizes this information. A structure is formed that encodes the information that *t* is a tree, previously encoded analogically, in digital form.

Essentially, a structure *s* has *t* is a tree as its semantic content because it is selectively sensitive to the information that *t* is a tree. *S* is responsive to the information that *t* is a tree stored in a variety of analogical forms, e.g. *t* is a brown tree, *t* is a tall tree. *S* is not sensitive to those features of the

signal that carry the information that *t* is brown or *t* is tall. That is what enables it to develop from structures that carry the information *t* is a brown tree as well as structures that carry the information that *t* is a tall tree.

Beliefs cannot be strictly identified with these semantic structures. A structure that has "*t* is *f*" as its semantic content, carries the information that "*t* is *f*". This entails that "*t* is *f*" is true since information entails truth. But beliefs can be false. As such, Dretske needs a way of accounting for how structures can develop with a false semantic content. Dretske accomplishes this by distinguishing between structure types endowed with a specific semantic content and concrete tokens which inherit their semantic content from their type. A certain type of semantic structure acquires its content from its informational origins, from the information to which it is selectively sensitive. Once it is formed, its semantic content is bequeathed to its subsequent tokens regardless of the informational origins of those tokens, i.e., regardless of whether those tokens actually carry the information. Thus a semantic structure token can have "*t* is *f*" as its content even though "*t* is *f*" is false. This would be an instance of a false belief.

Given these details of the causal process that lead from the receipt of information to the formation of belief, we can now evaluate whether this process can bear the necessary epistemic weight--whether Dretske has provided a naturalistic analogue of evidence-appreciation. I think it is clear he has not. Dretske tells us nothing about the process that converts analogically encoded information to digitally encoded information that indicates that it is in any way epistemically sensitive. The digitalization process is just the process by which any belief is formed. But not all beliefs count as instances of knowledge. We have seen that the receipt of information results in knowledge only if the recipient appreciates the connection between signal and source. If the receipt of information produces a belief about the source, it does not follow that the belief counts as knowledge even if the belief is true. Dretske claims that if the process is initiated by the relevant information, the subject comes to know. But Dretske has not apprised us of any feature of the process when it is actually initiated by the information that could reasonably be construed as appreciating the information/evidence.

It is true that the belief that t is f is a token of a semantic structure that inherits a specific content " t is f " from its selective sensitivity to the

information that t is f . And it might be claimed that the tokening of a type of semantic structure selectively sensitive to the information that t is f , is a suitable naturalistic realization of "evidence appreciation." Whether this is a viable position will depend on exactly what is entailed by a structure being selectively sensitive to certain information. Dretske tells us, "A structure's selective sensitivity to the information that s is f is just another way of describing the fact that this structure develops a digital representation of the f -ness of things in the kind of circumstances characteristic of the learning situation."¹⁷ For Dretske, to develop a structure selectively sensitive to the information that t is f is to acquire the concept of f -ness. But if one acquires the concept of f -ness it does not follow that one can appreciate the presence of f -ness anytime one receives the information that something is f . To receive the information that t is f , one need only receive a signal whose occurrence nomically entails the existence of something which is f . And certainly it does not follow from the fact that one has the concept of f -ness, that one is able to recognize all the signals whose occurrence nomically entails that something is f .

17

Ibid, p. 219.

The forms in which a specific informational content can be transmitted vastly exceed the knowledge one acquires when one acquires a concept. Thus the tokening of a semantic structure selectively sensitive to the information that t is f , even if it is occasioned by the receipt of the information that t is f , cannot by itself be taken as a suitable realization of the process whereby one appreciates the nomic dependencies that yield information.

This deficiency in Dretske's theory is evidenced in the theories vulnerability to counter example. Suppose a subject k , after a period of receiving the information that something is an acid, develop the capacity for digitalizing this information. A semantic structure evolves with the content " S is an acid." K has little knowledge of chemistry--specifically, he knows nothing about litmus tests. He dips a piece of litmus paper in a solution and the reading on the paper carries the information that the solution is an acid. The percept produced in K carries the information, encoded in analog form, that the solution is an acid. This information activates a token of the semantic structure with the content " S is an acid." K comes to believe that S is an acid. According to Dretske's theory K thereby comes to know that the solution is an acid. This is

strongly counterintuitive. Since K knows nothing about litmus tests, his belief, from an epistemic point of view, is quite arbitrary.

If anyone's intuitions are unswayed by this example, the situation can be made even clearer. Suppose that a chemist (falsely) tells K that the litmus paper indicates that the solution is a base. Nonetheless, K is caused by receiving the information to believe S is an acid. Not only does K not appreciate the (litmus paper) evidence, he possesses contradictory evidence which he does appreciate. Clearly K does not know that S is an acid. Yet Dretske's theory predicts that he does know.

I think these examples show that Dretske's information-theoretic criterion is not sufficient for knowledge. The question remains whether it is necessary for knowledge. Dretske argues that his theory, in virtue of the information requirement enjoys certain important advantages over conventional theories. As I indicated previously, the theory seems to be too strong--the information requirement leads to skepticism. Dretske explicitly addresses the skeptical implications of his theory and sketches a solution. I will argue that his solution in effect recants most of what served to distinguish Dretske's theory from its theoretical foils. This, in turn, undermines the advantages Dretske claims for his theory.

We can begin by citing those alleged advantages:

(1) The Gettier Problem

The Gettier problem has proved to be most intractable. It has resisted so many attempted solutions that most epistemologists have simply stopped talking about it. Nonetheless, the inadequacies of the traditional justified, true belief analysis are sufficiently well documented to make a solution to the Gettier problem a desideratum of any theory of knowledge. Dretske claims that his theory, in virtue of the information requirement, affords a solution.

Gettier cases typically exploit the fact that justification can be transmitted by reasoning through a false premise. Dretske discusses a case where K is justified in believing a marker is on square two of a checker board. Recognizing it as a deductive consequence of what he believes, K infers that the marker is on square two or three. Since the marker is on square three, K has a justified true belief that falls short of knowledge.

Dretske argues that his analysis is not subject to this counter-example. Since it is false that the marker is on square two, K could not have received the information that it is on square two. So even if K justifiably infers the true proposition that the marker is on square two or three, it does not follow K comes to

know this proposition as it does not follow that K has received the information that the marker is on square two or three.

(2) Conjunctivity

The conjunction principle ($K_p \cdot K_q \supset K_{pq}$) has undergone some critical scrutiny in connection with discussions of the lottery paradox. However, the vast majority of philosophers agree that the cost of abandoning the principle is too great. By assuming the conjunction principle, Dretske believes he can show that the receipt of information is a necessary condition of knowledge. The reasoning closely parallels that employed in the argument for the semantic theory of information that pivoted on assuming the xerox principle. Again Dretske runs a reductio, this time on the assumption that the knowledge that *s* is *f* does not require the receipt of the information that *s* is *f*, i.e., knowledge does not require that the signal that causes the belief that *s* is *f*, make the probability that *s* is *f* equal to one.¹⁸

18

It must be remembered that for Dretske (in this context) "information" is a technical term. When Dretske talks about the loss of information, he refers only to the reduction of the probability of the state of affairs at the source, given the signal. Although Dretske himself warns the reader about this, I think it is still easy to credit his arguments with more cogency than they deserve by taking "information" as referring to the ordinary concept. It is easy to fall into this when assessing Dretske's claim that it's absurd to hold that K can know *s* is *f* when K has received almost no

Suppose s can come to know that s is f without having received the information that s is f . Then s can come to know a number of propositions without having received the corresponding information. Assuming the conjunction principle s would come to know the conjunction of those known propositions. But Dretske points out that the conjunctive proposition can represent an enormous loss of information. For the information losses of the individual propositions accumulate in the conjunction. (The probability of the conjunction is a function of the probabilities of the conjuncts). But it is absurd, Dretske contends, to hold that one can know a proposition is true, when one has received almost no information about the state of affairs described by the proposition. (When the probability of the proposition, given the information one has received, is quite low). Thus one cannot come to know s is f without receiving the information that s is f .

(3) Communication

Much of what we learn comes to us through communication. Any acceptable theory of knowledge should allow for the transmission of knowledge through communication. Assuming this, Dretske again runs a reductio on the position that knowledge that s is f can arise

information about s .

without the receipt of the information that s is f . And again the argument parallels the argument for the semantic theory.

Suppose one can know s is f without receiving the information that s is f . Then imagine a chain of communication with a loss of information between the adjacent links. Again the loss of information will accumulate in the chain until eventually the distant links in the chain will be receiving virtually no information about the source. Assuming communication can transmit knowledge, we are left with the absurd result that the distant links in the chain can still come to have knowledge about the source.

The advantages Dretske claims for his theory will come to naught if the theory entails skepticism. And as we saw earlier, Dretske's theory seems to yield just this result. The semantic theory of information requires that the probability of s is f given the occurrence of the signal r , be 1, if r is to carry the information that s is f . The difficulty seems to be that it's problematic whether there are any actual cases that meet this strict requirement. Thus, it would seem that signals never carry information (semantically). And since knowledge is informationally caused belief, it would seem that there are no instances of knowledge.

Dretske recognizes this skeptical argument and attempts to defuse it. Following Peter Unger, Dretske holds that knowledge is an absolute concept. Knowledge arises only when there is no possibility of error. Dretske holds that knowledge inherits this property from the information on which it depends. For the semantic theory holds that a signal carries the information that s is f only if all (nomic) possibilities other than s is f have been eliminated. The skeptical thesis amounts to the claim that in actual situations, there will always be uneliminated alternative possibilities.

Dretske's anti-skeptical strategy consists in parting company with Unger, by proposing that the absolute character of information and knowledge is qualified by a relative or pragmatic element. To know, or to have received information it must be the case that all relevant alternative possibilities have been eliminated, and not every possibility is relevant. Dretske has a technical term for those states of affairs whose alternative possibilities are not relevant. He calls them channel conditions. "The channel is that set of existing conditions that have no relevant alternative states..."¹⁹

19

Ibid, p. 123.

It looks as if Dretske has the following in
 20 mind: The skeptical thesis exploits the fact that the conditional probability of a state of affairs will always be less than one, owing to the persistence of alternative possibilities. But these possibilities are not relevant. And the probability of a state of affairs is conditional not only on the occurrence of the signal, it is conditional on the assumption that the nonrelevant possibilities do not obtain. This allows for states of affairs to have probability one, thereby satisfying the requirements of the semantic theory. Signals can carry information and when that information causes belief in a subject, that subject comes to know.

Dretske's solution naturally raises the question of what is to count as a relevant alternative. One wants to know what the criteria of relevance are. Only then can it be determined whether the alternative possibilities that generate the skeptical result can plausibly be construed as not relevant.

As Dretske notes, the notion of a relevant alternative has played a prominent role in recent

20

I found the chapter on skepticism, "The Communication Channel" quite vague and confusing. What follows is my attempt to distill a straightforward position from a long series of rambling remarks.

discussion of the Gettier problem. Alvin Goldman considers whether someone who perceives a barn while driving through the Wisconsin countryside, has his knowledge that there is a barn before him, threatened by his inability to rule out the possibility that what he is perceiving is an elaborately constructed barn facsimile.²¹ The issue is whether a person's inability to rule out such alternatives is relevant to the question of whether he knows. Most people would agree that generally it is not. However if barn facsimiles actually happen to exist in the Wisconsin countryside, then such a possibility becomes relevant. Whether or not a possibility is relevant turns out to be a matter of degree. For example suppose the facsimiles are extremely rare. Suppose the facsimiles exist only in Iowa?...in Sweden?

One of the upshots of the vast Gettier literature is that whether a person knows, i.e., whether a possibility is relevant, depends on pragmatic features such as the beliefs and knowledge of the social group and the availability of unpossessed evidence. Dretske claims that information has pragmatic elements as well.

Whether or not a signal carries a piece of information depends on what the channel is between source and receiver, and the question of whether an

21

Alvin I. Goldman, "Discrimination and Perceptual Knowledge," Journal of Philosophy, 1976, pp. 771-791.

existing condition is stable or permanent enough to qualify as part of the channel...is a question of degree, a question about which people (given their differing interests and purposes) can reasonably disagree, a question that may not have an objectively correct answer. When a possibility becomes a relevant possibility is an issue that is, in part at least, responsive to the interests, purposes, and, yes, values of those with a stake in the communication process. The flow of information, just like the cognitive exploits it makes possible, is a process that exhibits some sensitivity to the variable purposes of those who send and receive this information.²²

Dretske points out that this relativity does not conflict with the absoluteness of knowledge and information. Knowledge and the information upon which it depends still require the elimination of all relevant alternatives. Flexibility enters in only in consideration of what will count as a relevant alternative.

Be that as it may, Dretske's remarks on what determines whether or not an alternative is relevant fall drastically short of being a systematic account. This makes it very difficult to assess Dretske's defense against the charge of skepticism. We are in no position to determine whether the alternative possibilities are going to be relevant or not. Moreover it looks dubious that the notion of a relevant alternative can do the work Dretske intends it to do. For we have no reason to believe that in every alleged case of information transmission, or knowledge, the alternative

²²

Op. cit. pp. 132-133.

possibilities that diminish the probabilities will fail to be relevant. If we start with the intuition that a subject knows in a certain case, we can then go on to brand the uneliminated alternatives as irrelevant. But Dretske cannot use this criterion on pain of circularity. Given the vagueness of Dretske's remarks, I think it's fair to characterize his case against skepticism as spacious.

But there is a deeper problem. Dretske's anti-skeptical arguments succeed only at the expense of his semantic theory of information and his information-theoretic epistemology. For Dretske, the connection to truth is central to what he calls the nuclear concept of information. "In this [the nuclear] sense of the term, false information and misinformation are not kinds of information...to speak of certain information as being reliable is to speak redundantly."²³ Dretske tells us that the terms in which dictionaries most frequently define "information" all "...have a common nucleus. They all point in the same direction--the direction of truth."²⁴ As Dretske notes, information could not bear the theoretical burden he sets out for it, if it did not

23

Op. cit. p. 45.

24

Op. cit. p. 45.

entail truth. "Information is what is capable of yielding knowledge, and since knowledge requires truth,²⁵ information requires it also."

It is this nuclear concept of information that the semantic theory of information must capture. As Dretske states, this is the central part of the motivation for the probability one requirement: "Condition B [the truth requirement] is satisfied because if the conditional probability of s being f is one, then s is f ."²⁶ The problem is that Dretske's case against skepticism undermines this claim. Recall that Dretske's solution consists in excluding certain possibilities as irrelevant. But this clearly does not entail that these possibilities do not obtain. Relevance is a pragmatic notion, and unless we are willing to countenance some version of a pragmatic conception of truth, irrelevant possibilities can still obtain. This is not to deny that the conditional probability of a state of affairs can be one. It is just that on Dretske's interpretation, a state of affairs having probability one does not entail that the state of affairs obtains.

²⁵ Op. cit. p. 45.

²⁶ Op. cit. p. 65.

Consider ordinary perception. Suppose it appears to K that s is f. It is certainly nomically possible that K have such a percept and it not be the case that s is f. K could be deceived by a hallucination or a perceptual equivalent. Now if Dretske wants to avoid skepticism by making the probability that s is f, conditional on s not being deceived in these ways, the probability that s is f may equal one. But s may be hallucinating nonetheless.

In view of this, the semantic theory fails to give an adequate account of information. It does not preserve the essential link between information and truth. We can now investigate whether the theory can serve as a foundation for a theory of knowledge.

An initial difficulty is that knowledge entails truth. One cannot know that s is f if it is false that s is f. Since the revised antiskeptical account of information allows that one can receive the information that s is f when s is not f, one could acquire an informationally caused belief that s is f, when s is not f. In such a case one would not know s is f.

It looks as if Dretske's theory of knowledge will have to be supplemented with an independent requirement of truth:

K knows that s is f = (1) k's belief that s is f is caused (or causally sustained) by the information that s is f; (2) s is f.

Unfortunately, the purported advantages of the theory are lost. Consider first the Gettier problem. Dretske held that his analysis was immune to Gettier counter-examples because one cannot receive the information that s is f when s is not f . While reasoning through a false premise may transmit justification, it does not transmit information. But this is no longer true given Dretske's solution to skepticism. The transmission of information does require probability one, but probability one does not entail truth.

Returning to Dretske's example, suppose the marker is on square three. S is having a rare hallucination that makes it appear as if the marker is on square two. Because the possibility of hallucination is irrelevant (at least it will have to be in most cases if Dretske's anti-skeptical argument will work in the case of ordinary perception), the conditional probability of the marker being on square two, given s' percept, is one. It follows that s' percept carries the information that the marker is on square two. But then s' percept carries the information that the marker is on square two or three.²⁷ If s' percept causes him to believe the marker is on square two or three, s will have an

27

If s' percept makes the probability that the marker is on square two, equal to one, then trivially it

informationally caused true belief that is not a case of knowledge. Thus we have an information-theoretic analog of a Gettier counter-example.

The second advantage Dretske claims for his theory is that it preserves the validity of the conjunction principle ($K_p \cdot K_q \supset K_{pq}$). Dretske alleges the conjunction principle runs into trouble on any theory that permits knowledge of a proposition with less than probability one. As such propositions are conjoined, the conditional probability of the conjunction diminishes indefinitely. But one cannot know a proposition with a very low conditional probability.

After Dretske's remarks on skepticism, it becomes puzzling exactly what the advantage of Dretske's theory comes to. Dretske's theory can preserve the validity of the conjunction principle because it requires probability one for known propositions. In order to avoid skepticism, Dretske makes the probability conditional on the assumption that a host of possibilities (the irrelevant ones) do not obtain. But given this qualification, it is no longer clear how the strict probability requirement distinguishes Dretske's theory from more conventional

makes the probability that the marker is on square two or three, equal to one. Dretske refers to this as the information that the marker is on square two or three being analytically nested in the information that the marker is on square two. See pp. 70-71.

ones. When most epistemologists insist that a probability one requirement is too strict (entails skepticism) they have in mind precisely those possibilities Dretske himself must exclude as irrelevant, if his theory is to avoid skepticism. If these possibilities are excluded, most epistemological theories would be consistent with the requirement that know proposition have probability one. Of course it is hard to be precise about this since Dretske is not precise about the criteria of relevance for alternative possibilities. But I think its fair to say that Dretske's remarks on skepticism obscure the difference between his theory and those he criticizes. As such, any claim to advantage is suspect.

Similar considerations apply to Dretske's comments about chains of communication. While its true knowledge must be transmitted through (some) communication chains, the possibilities of error in such chains are precisely those that Dretske must exclude as irrelevant if he is to avoid skepticism. So the strict probability requirement is no longer controversial.

Finally, we are able to see that Dretske's indictment of theories that lack the strict probability requirement is ill-conceived. Those theories are deficient if it is assumed that the probability of a

belief, that counts as an instance of knowledge, most equal one, given everything that's nomically possible. But as Dretske himself points out, not every possibility is relevant to the question of whether someone knows. To claim otherwise results in skepticism.

REFERENCES

- Chisholm, Roderick. Theory of Knowledge, 2nd Edition,
(Englwood Cliffs: 1976).
- Dretske, Fred. Knowledge and the Flow of Information,
(Oxford, 1981).
- Goldman, Alvin. "What is Justified Belief?" in
George S. Pappas (ed.), Justification and
Knowledge (Dordrecht, 1979).
- Harman, Gilbert. Thought (Princeton, 1973)
- Lehrer, Keith. Knowledge (Oxford, 1974).
- Pollock, John. Knowledge and Justification
(Princeton, 1974).