JOHN L. POLLOCK

EPISTEMIC NORMS*

1. INTRODUCTION

Historically, the main concern in epistemology has been to explain how we are justified in holding the various kinds of beliefs we have about the world. When we ask whether a belief is justified, what we want to know is whether it is all right to believe it. Justification is a matter of "epistemic permissibility". It is not a novel observation that epistemic justification is a normative notion, but by emphasizing the normative character of epistemic justification and downplaying its role in knowledge I hope to avoid some confusions that (I will argue) have plagued recent discussions of epistemic justification. I will think of epistemic justification as being concerned with questions of the form, "When is it permissible (from an epistemological point of view) to believe that *P*?" This is the concept of epistemic justification that I am concerned with exploring.

Norms are general descriptions of the circumstances under which various kinds of normative judgments are correct. Epistemic norms are norms describing when it is epistemically permissible to hold various beliefs. A belief is justified iff it is licensed by correct epistemic norms. Assuming that what justifies a belief is the reasoning underlying it ("reasoning" construed broadly), epistemic norms are the norms governing "right reasoning". Epistemologists have commonly supposed that epistemic norms are much like moral norms and that they are used in evaluating reasoning in the same way moral norms are used in evaluating actions. One of the main contentions of this paper will be that this parallel is not at all exact and that epistemologists have been misled in important ways by supposing the analogy to be better than it is. A proper understanding of epistemic norms will provide us with a radically new perspective on epistemology, and from the point of view of this perspective new light can be thrown on a number of central epistemological problems.

Perhaps the most fundamental disagreement about epistemic norms is that involved in the internalism/externalism debate. Contemporary

Synthese **71** (1987) 61–95.

epistemologists have divided into two camps. In the one camp we have the traditionally oriented *internalists*, according to whom the justifiedness of a belief is a function exclusively of the internal states of the believer.[1] This means that if one is in the same internal states in two possible circumstances, then no matter how those circumstances differ with respect to things other than one's internal states, there will be no difference in what beliefs are justified under those circumstances. In the other camp we have the *externalists* who maintain that more than the internal states of the believer can be relevant to the justifiedness of his beliefs. The internalism/externalism distinction is notoriously unclear in one respect. It is formulated in terms of an undefined notion of an internal state. It is fairly clear what kinds of states people have in mind when they talk about internalism and externalism, but it is hard to give a general characterization of these states. I will return to this matter in section 5.

The internalist tries to map out the structure of our epistemic norms entirely in terms of relations between internal states. For example, an internalist may tell us that something's looking red to me gives me a prima facie reason to think that it is red. The internalist alleges that by compiling a list of such epistemic rules he has described our epistemic norms and thus provided an adequate account of epistemic justification. Externalists typically object that such lists of rules leave the concept of justification unexplained and mysterious. The externalist insists instead that the purpose of reasoning is to achieve certain epistemic goals (most notably the acquisition of true beliefs) and hence correct epistemic norms should be those enabling us to achieve these goals. There are two ways external considerations could be brought to bear on epistemic norms, and they have not been clearly distinguished in most recent discussions of externalism. On the one hand, our epistemic norms could be *formulated in terms of* external considerations. A typical example of such a norm might be, "It is permissible to hold a belief if it is generated by a reliable belief-forming process". I will call this variety of externalism *belief externalism*. In contrast to this, *norm externalism* acknowledges that the rules comprising our epistemic norms must be internalist, but employs external considerations in the selection of the norms themselves. The distinction between belief and norm externalism is analogous to the distinction between act and rule utilitarianism. Externalism (simpliciter) is the disjunction of belief externalism and norm externalism.

A number of philosophers who are usually considered externalists appear to vacillate between belief externalism and norm externalism. The difference between these two varieties of externalism will prove important.[2]

Reliabilism is that version of belief externalism that seeks to formulate epistemic norms in terms of considerations of reliability.[3] Reliabilism is the most common variety of belief externalism, but it is important to realize that it is not the only possible variety. Any theory proposing non-internalist norms is a version of belief externalism. Non-internalist norms need not proceed in terms of reliability. Some examples of belief externalists who are not reliabilists will be discussed in section 3.

To my mind the most telling objection to existing internalist theories is that they are simultaneously incomplete and ad hoc. They are incomplete in that they leave the concept of epistemic justification unanalyzed, and they are ad hoc in that they propose arrays of epistemic rules without giving any systematic account of why those should be the right rules.[4] The methodology of internalism has been to describe our reasoning, rather than to justify it or explain it. These two points are connected. As long as we take the concept of epistemic justification to be primitive and unanalyzed, there is no way to *prove* that a particular epistemic rule is a correct rule. All we can do is collect rules that seem intuitively right, but we are left without any way of justifying or supporting our intuitions. Herein lies the main attraction of externalism. Externalist theories begin by proposing analyses of epistemic justification from which epistemic rules can be derived. Epistemic justification is no longer taken as primitive, and there is no longer any need to simply posit epistemic rules. Of course, the success of this approach turns upon whether externalist analyses of epistemic justification can be successful. There is a wide variety of externalist theories, and each is subject to its own difficulties. One cannot, however, do a successful job of refuting generic externalism by refuting individual externalist theories one at a time. Instead, my strategy will be to raise a general difficulty that, I believe, will demonstrate the impossibility of any externalist theory. Confronting this difficulty will ultimately enable us to understand the source of epistemic norms and the nature of epistemic justification, thus resolving what I take to be the primary and most glaring problem for internalism and leading us to a novel "naturalistic" internalism.

2. HOW DO EPISTEMIC NORMS REGULATE?

In order to get a grasp on the nature of epistemic norms, let us begin by asking their purpose. It is important to distinguish between two uses of norms (epistemic or otherwise). On the one hand, there are third person uses of norms wherein we use the norms to evaluate the behavior of others. Various norms may be appropriate for third person evaluations, depending upon the purpose we have in making the evaluations. For example, we may want to determine whether a person is a good scientist because we are trying to decide whether to hire him. To be contrasted with third person uses of norms are first person uses. First person uses of norms are, roughly speaking, action-guiding.[5] For example, I might appeal to *Fowler's Modern English Usage* to decide whether to use "that" or "which" in a sentence. Epistemological questions are inherently first person. The traditional epistemologist asks. "How is it possible for me to be justified in my beliefs about the external world, about other minds, about the past, etc.?" These are questions about what to believe. Epistemic norms are the norms in terms of which these questions are to be answered, so these norms are used in a first person reason-guiding capacity.

If reasoning is governed by epistemic norms, just how is it governed? There is a model of this regulative process that is often implicit in epistemological thinking, but when we make the model explicit it is *obviously* wrong. This model assimilates the functioning of epistemic norms to the functioning of explicitly articulated norms. For example, naval officers are supposed to "do it by the book", which means that whenever they are in doubt about what to do in a particular situation they are supposed to consult explicit regulations governing all aspects of their behavior and act accordingly. Explicitly articulated norms are also found in driving manuals, etiquette books, etc. Without giving the matter much thought, there is a tendency to suppose that all norms work this way, and in particular to suppose that this is the way epistemic norms work. I will call this "the intellectualist model".[6] It takes little reflection to realize that epistemic norms cannot function in accordance with the intellectualist model. If we had to make an explicit appeal to epistemic norms in order to acquire justified beliefs we would find ourselves in an infinite regress, because to apply explicitly formulated norms we must first acquire justified beliefs about how they

apply to this particular case. For example, if we are to reason by making explicit appeal to a norm telling us that it is permissible to move from the belief that something looks red to us to the belief that it is red, we would first have to become justified in believing that that norm is included among our epistemic norms, and we would have to become justified in believing that we believe that the object looks red to us. In order to become justified in holding those beliefs, we would have to apply other epistemic norms, and so on ad infinitum. Thus it is clear that epistemic norms cannot guide our reasoning in this way.[7]

If the intellectualist model is wrong, then how do epistemic norms govern reasoning? At this point we might raise the possibility that they do not. Perhaps epistemic norms are only of use in third person evaluations. But it cannot really be true that epistemic norms play *no role at all* in first person deliberations. We can certainly subject our reasoning to self-criticism. Every philosopher has detected invalid arguments in his own reasoning. This might suggest that epistemic norms are only relevant in a negative way. Our reasoning is innocent until proven guilty. We can use reasoning to criticize reasoning, and hence we can use reasoning in applying epistemic norms to other reasoning, but we cannot be required to reason about norms *before* we can do any reasoning. This would avoid the infinite regress.

But as theoretically attractive as the "innocent until proven guilty" picture might be, it cannot be right. There are a number of natural processes that lead to belief formation. Among these are such "approved" processes as vision, inductive reasoning, deductive reasoning, memory, etc., and also some "unapproved" but equally natural processes like wishful thinking. The latter is just as natural as the former. For example, a friend of mine recently drove to Albuquerque. The morning she left the weather turned unseasonably cold. As she was leaving we joked about her driving into a snow storm. As it turned out, that is exactly what happened. After she had left and I learned how bad the weather was, it occurred to me to wonder whether she had taken a coat (in fact, she had not). I found myself thinking, "Oh, she must have", and dismissing the matter from my mind. Then I realized that was just wishful thinking. I had no reason to believe she had taken a coat. The point here is that wishful thinking is a natural belief-forming process, but we do not accord it the same status as

some other belief-forming processes like vision. Although we have a natural tendency to form beliefs by wishful thinking, we also seem to "naturally" know better. This is not just a matter of after-the-fact criticism. We know better than to indulge in wishful thinking at the very time we do it. It seems that *while* we are reasoning we are being guided by epistemic norms that preclude wishful thinking but permit belief formation based upon perception, induction, etc. This is of more than casual significance, because it might be impossible to rule out wishful thinking by after-the-fact reasoning. This is because the after-the-fact reasoning might include wishful thinking again, and the new wishful thinking could legitimize the earlier wishful thinking. If epistemic norms play no regulative role in our reasoning while it is going on, there is no reason to think they will be able to play a successful corrective role in after-the-fact evaluations of reasoning. In order for the corrective reasoning to be successful it must itself be normatively correct. Epistemic norms must, and apparently do, play a role in guiding our epistemic behavior at the very time it is occurring. But how can they?

Epistemic norms cannot play a merely negative, corrective, role in guiding reasoning, nor can they function in a way that requires us to make judgments before we can make judgments. What is left? I think that our perplexity reflects an inadequate understanding of the way action-guiding norms usually function. The case of making an explicit appeal to norms in order to decide what to do is the exception rather than the rule. You may make reference to a driving manual when you are first learning to drive a car, but once you learn how to drive a car you do not look things up in the manual anymore. You do not usually give any explicit thought to what to do – you just do it. This does not mean, however, that your behavior is no longer being guided by those norms you learned when you first learned to drive. Similarly, when you first learned to ride a bicycle you were told to turn the handlebars to the right when the bicycle leaned right. You learned to ride in accordance with that norm, and that norm still governs your bike riding behavior but you no longer have to think about it. The point here is that *norms can govern your behavior without your having to think about them*. The intellectualist model of the way norms guide behavior is almost always wrong. This is an obvious point, but it has

been insufficiently appreciated. It is of major importance in understanding epistemic norms. Reasoning is more like riding a bicycle than being in the navy.

What makes it possible for your bike riding behavior to be governed by norms without your thinking about the norms is that you *know how* to ride a bicycle. Knowing how to ride a bicycle consists of knowing what to do under various circumstances, e.g., knowing to turn right when the bike leans right. Knowing what to do and being aware of it constitutes knowing what you *should* do. Moral philosophers have talked about different senses of "should", distinguishing particularly between moral uses of "should" and goal directed uses of "should". An example of the latter is "If you want the knife to be sharp then you should sharpen it on the whetstone". But the use of "should" in "In riding a bicycle, when the bicycle leans to the right you should turn the handlebars to the right" is of neither of these varieties. It is perhaps more like the goal directed kind of "should", but we are not saying that that is what you should do to achieve the goal of riding a bicycle. Rather, that is part of what is involved *in* riding a bicycle – that is *how* to ride a bicycle.

What we know in knowing how to ride a bicycle can be regarded as normative – we know what we should do under various circumstances. Knowing what we should do under various circumstances does not involve our being able to give a general description of what we should do under various circumstances. This is just to make the familiar observation that knowing how to ride a bicycle does not automatically enable one to write a treatise on bicycle riding. This is true for two different reasons. First, knowing how to ride a bicycle requires us to know what to do in each situation *as it arises*, but it does not require us to be able to say what we should do before the fact. Second, even when a situation has actually arisen, our knowing what to do in that situation need not be propositional knowledge. In the case of knowing that we should turn the handlebars to the right when the bicycle leans right, it is plausible to suppose that most bicycle riders do have propositional knowledge of this; but consider knowing how to hit a tennis ball with a tennis racket. I know how to do it – as the situation unfolds, at each instant I know what to do – but even at that instant I cannot give a description of what I should do. Knowing what to do is

the same thing as knowing to do it, and that need not involve propositional knowledge.

We are now in a position to give a rough explanation of how action-guiding norms can govern behavior in a non-intellectualist manner. When we learn how to do something X, we "acquire" a plan of how to do it, and that plan becomes internalized. When we subsequently undertake to do X, our behavior is automatically channeled into that plan. This is just a fact of psychology. We form habits or conditioned reflexes. Norms for doing X constitute a description of this plan for doing X. The sense in which the norms guide our behavior in doing X is that the norms describe the way in which, once we have learned how to do X, our behavior is automatically channeled in undertaking to do X.

Now let us apply this to epistemic norms. We know how to reason. That means that under various circumstances we know what to do in reasoning. This entails that there are things we should do, and hence that there are epistemic norms that guide our reasoning. The way epistemic norms can guide our reasoning without our having to think about them is no longer mysterious. They describe an internalized pattern of behavior that we automatically follow in reasoning, in the same way we automatically follow a pattern in bicycle riding. This is what epistemic norms are. They are the internalized norms that are used automatically when we reason. Once we realize that they are just one more manifestation of the general phenomenon of automatic behavior governed by internalized norms, epistemic norms should no longer seem puzzling. We would like to have a better understanding of the psychological process wherein behavior is generated in conformance with internalized norms, and I will say more about this below. But in the meantime, much of the mystery surrounding epistemic norms evaporates once we recognize that the governing process is a general one and its application to epistemic norms and reasoning is not much different from its application to any other kind of action-guiding norms. Of course, unlike most norms our epistemic norms may be innate, in which case there is no process of internalization that is required to make them available for use in guiding our reasoning.[8]

I have described how our epistemic norms work. This is to describe our *actual* epistemic norms. Internalists typically assume that what-

ever our actual epistemic norms are, they are the correct epistemic norms. I have taken it to be part of the definition of internalism that our epistemic norms are at least not subject to criticism on externalist grounds. Of course, this is precisely where internalists disagree with norm externalists. Let us turn then to a reconsideration of externalism in the light of our new understanding of epistemic norms.

3. THE REFUTATION OF EXTERNALISM

3.1. *Belief externalism*

Now that we understand how epistemic norms work in guiding our reasoning, it is easy to see that they must be internalist norms. This is because when we learn how to do something we acquire a set of norms for doing it and these norms are internalized in a way enabling our central nervous system to follow them in an automatic way without our having to think about them. This has implications for the content of our norms. For example, I have been describing one of our bike riding norms as telling us that if the bicycle leans to the right then we should turn the handlebars to the right, but that is not really what we learn when we learn to ride a bicycle. The automatic processing systems in our brain do not have access to whether the bicycle is leaning to the right. What they do have access to are things like (1) our *thinking* that the bicycle is leaning to the right, and (2) certain balance sensations emanating from our inner ear. What we learn (roughly) is to turn the handlebars to the right if we either experience those balance sensations or think on some other basis that the bicycle is leaning to the right. In general, the circumstance-types to which our norms appeal in telling us to do something in circumstances of those types must be directly accessible to our automatic processing systems. The sense in which they must be directly accessible is that our automatic processing system must be able to access them without our first having to make a *judgment* about whether we are in circumstances of that type. We must have non-epistemic access.[9]

This general observation about action-guiding norms has immediate implications for the nature of our epistemic norms. It implies that reason-guiding epistemic norms cannot appeal to external considerations of reliability. This is because such norms could not be internalized. Like *leaning to the right*, considerations of reliability are

not directly accessible to our automatic processing systems. There is in principle no way that we can learn to make inferences of various kinds only if they are *in fact* reliable. Of course, we could learn to make certain inferences only if we *think* they are reliable, but that would be an internalist norm appealing to *thoughts* about reliability rather than an externalist norm appealing to reliability itself.[10] Similar observations apply to any externalist norms. Consequently, it is in principle impossible for us to actually employ externalist norms. I take this to be a conclusive refutation of belief externalism.

I introduced the internalism/externalism distinction by saying that internalist theories make justifiedness a function exclusively of what internal states the believer is in, where internal states are those that are "directly accessible" to the believer. The notion of direct accessibility was purposely left vague, but it can now be clarified. I propose to define internal states to be those states that are directly accessible to the mechanisms in our central nervous system that direct our reasoning. The sense in which they are *directly* accessible is that access to them does not require us to first have beliefs about them. This definition makes the internalist/externalist distinction precise in a way that agrees at least approximately with the way it has generally been used, although it is impossible to make it agree with everything everyone has said about it because different philosophers have drawn the distinction in different ways.

I have characterized internalist theories in terms of direct accessibility, but I have not said anything in a general way about which states are directly accessible. It seems clear that directly accessible states must be in some sense "psychological", but I doubt that we can say much more than that from the comfort of our armchair. That is an empirical question to be answered by psychologists. Despite the fact that we do not have a general characterization of direct accessibility, it is perfectly clear in many specific cases that particular states to which philosophers have appealed are not directly accessible. In light of this, the preceding refutation of belief externalism can be applied to a remarkably broad spectrum of theories, and it seems to me to constitute an absolutely conclusive refutation of those theories. I have indicated how it applies to theories formulating epistemic norms in terms of reliability. It applies in the same way to a much wider class of theories that proceed generally in terms of probability. For example, a few philosophers endorse the *Simple Rule*:

> A belief is justified iff what is believed is sufficiently probable.[11]

If the simple rule is to provide us with a reason-guiding norm, then a belief's being sufficiently probable must be directly accessible. No objective probability can have that property. Thus it is impossible to use the simple rule, interpreted in terms of objective probabilities, as a reason-guiding norm. This objection could be circumvented by replacing the simple rule by its "doxastic counterpart":

> A belief is justified iff the epistemic agent believes it to be highly probable.

But this rule formulates an internalist norm (albeit, an implausible one).[12]

It might be supposed that we could breath life back into the simple rule by interpreting it in terms of subjective probability. Here we must be careful to distinguish between subjective probability as actual degree of belief and subjective probability as rational degree of belief. Interpreted in terms of actual degrees of belief, the simple rule would amount to the claim that a belief is justified iff it is firmly held, which is an internalist norm, but a preposterous one. Interpreted in terms of rational degrees of belief it becomes an externalist norm. Rational degree of belief is the unique degree of belief one rationally ought to have in a proposition, given one's overall doxastic state. I have serious doubts about the intelligibility of this notion. Is there any reason to believe that there is a unique rational degree of belief a person ought to have in a proposition? But even if we waive this difficulty, ascertaining what this unique rational degree of belief should be is immensely difficult. The rational degree of belief one ought to have in a proposition is certainly not a directly accessible property of it, and hence this version of the simple rule also succumbs to our general objection to belief externalism.

Many other epistemological theories succumb to this objection to belief externalism. For example, Keith Lehrer's coherence theory proceeds in terms of probability and hence is akin to reliabilist theories in various ways, but it makes use of probability in a complicated way that disqualifies it from being a reliabilist theory. Leaving out a few details, Lehrer's proposal is:

r competes with h for S iff prob(h/r) is less than prob(h) and the disjunction d which is logically equivalent to r and contains as disjuncts members m_1, m_2, and so forth of the epistemic partition of h for S in numerical order, is such that no disjunction d' of any of those members can be formed where prob(h/d') = prob(h).

S is completely justified in believing h if and only if prob(h) is greater than prob($\sim h$) and for any r, if r competes with h for S, then prob(h) is greater than prob(r).[13]

But a proposition's being more probable than any of its competitors is most assuredly not a directly accessible property of it, and hence Lehrer's theory becomes incapable of supplying us with a reason-guiding norm.[14]

This same kind of objection applies to a broad class of coherence theories. In Pollock (1979), I distinguished between *holistic coherence theories* and *linear coherence theories*.[15] A linear coherence theory takes a classical view of reasons according to which one belief is a reason for a second by virtue of some internal relation between them, and differs from a foundations theory only in the overall use it makes of reasons and reasoning. Such a theory is immune from the present objections. A holistic coherence theory, on the other hand, adopts a holistic view of reasons according to which what licenses a belief is its being suitably related to the set of *all* the beliefs one holds. Lehrer's coherence theory is of the holistic variety. A holistic coherence theory requires a relationship between a justified belief and the set of all the beliefs one holds, but that will not normally be a directly accessible property of the justified belief, and hence the norm proposed by the holistic theory will be an externalist norm. Thus it cannot be reason-guiding.

The present account of epistemic norms is efficient in dispatching a wide variety of epistemological theories, but it also has some positive consequences. I take foundationalist theories to require that all justification derives ultimately from "epistemologically basic beliefs". These are typically taken to be beliefs about how we are appeared to, what we seem to remember, and so forth. In earlier publications I have rejected foundationalist theories on the grounds that we rarely have such beliefs, and in their place I have endorsed *direct realism*, according to which justification typically derives from nondoxastic states like *being appeared to really* (without your having to believe that you are appeared to really).[16] The need for the move from foundationalism to direct realism seems to me to be compelling, but the move itself can seem puzzling. How can it be possible for nondoxastic states to justify

beliefs when we are not aware that we are in them?[17] This really only seems puzzling because we are implicitly assuming the intellectualist model of the way epistemic norms regulate belief. Given the way epistemic norms actually operate, all that is required is that the input states be directly accessible. Belief states are directly accessible, but so are a variety of nondoxastic states like perceptual states and memory states. Thus there is no reason why epistemic norms cannot appeal to those states, and the move to direct realism ceases to be puzzling.

Is there any way to salvage belief externalism in the face of the objection that it cannot give reasonable accounts of first person reason-guiding epistemic norms? The possibility remains that belief externalism might provide norms for third person evaluations. I think it is noteworthy in this connection that externalists tend to take a third person point of view in discussing epistemology. If externalist norms played a role in third person evaluations, we would then have both externalist and internalist norms that could be applied to individual beliefs and they might conflict. What would this show? It would not show anything – they would just be different norms evaluating the same object from different points of view. I can imagine a persistent externalist insisting, "Well, if the two sets of norms conflict, which way should we reason – which set of norms should we follow?" But that question does not make any sense. Asking what we should do is asking for a normative judgment, and before we can answer the question we must inquire to what norms the "should" is appealing. To make this clearer consider an analogous case. We can evaluate beliefs from both an epistemic point of view and a prudential point of view. Suppose Helen has good reasons for believing that her father is Jack the Ripper, but suppose coming to believe that would be psychologically crushing. Then we might say that, epistemically, she should believe it, but prudentially she should not. If one then insists upon asking, "Well, should she believe it or not?", the proper response is, "In what sense of 'should', epistemic or prudential?" Similarly, if externalist and internalist norms conflict and one asks, "Which way should we reason?", the proper response is to ask to which set of norms the "should" is appealing. The point is that different norms serve different purposes, and when they conflict that does not show that there is something wrong with one of the sets of norms – it just shows that the different norms are doing different jobs. The job of internalist norms is reason-guiding, and as such they are the norms traditionally sought in

epistemology. Externalist norms (if any sense can be made of them) may also have a point, but they cannot be used to solve traditional epistemological problems pertaining to epistemic justification.

3.2. *Norm externalism*

Recall that there are two kinds of externalism. Belief externalism advocates the adoption of externalist norms. I regard belief externalism as having been decisively refuted by the preceding considerations. Norm externalism, on the other hand, acknowledges that we must employ internalist norms in our reasoning, but proposes that alternative sets of internalist norms should be evaluated in terms of external considerations. For example, it may be alleged that one set of internalist norms is better than another if the first is more reliable in producing true beliefs. Both internalism and norm externalism endorse internalist norms, but they differ in that the internalist alleges that our epistemic norms are not subject to criticism on externalist grounds. It is hard to see how they could be subject to criticism on internalist grounds, so the internalist has typically assumed that our epistemic norms are immune from criticism – whatever our actual epistemic norms are, they are the correct epistemic norms. That, however, seems odd. On the surface, it seems it must be at least logically possible for two people to employ different epistemic norms. They could then hold the same belief under the same circumstances and on the basis of the same evidence and yet the first could be conforming to his norms and the second not conforming to his. If a person's epistemic norms are always beyond criticism, it would follow that the first person is justified in his beliefs and the second is not, despite the fact that their beliefs are based upon the same evidence. That would at least be peculiar. Because it seems that it must be possible for different people to employ different epistemic norms, this makes a strong *prima facie* case for norm externalism.

Action-guiding norms are not generally immune from criticism. Typically, action guiding norms tell us how to do one thing *by* doing something else.[18] For example, knowing how to ride a bicycle consists of knowing what more basic actions to perform – leg movements, arm movements, and the like – in order to ride the bicycle. An action that is performed by doing something else is a *nonbasic* action. Norms describing how to perform nonbasic actions can be subject to external

evaluation. There may be more than one way to perform the nonbasic action, and some ways may be better (more efficient, more reliable, etc.) than others. If I know how to do it in one way and you know how to do it in another way, you know how to do it better than I if the norms governing your behavior are better than the norms governing mine. For example, we may both know how to hit the target with a bow and arrow, but you may know how to do it more reliably than I.[19] It thus becomes an empirical question whether acting in accordance with a proposed norm will constitute your doing what you want to be doing and whether another norm might not be better.

Reasoning is not, strictly speaking, an action, but it is something we do, and we do it by doing other simpler things. We reason by adopting new beliefs and rejecting old beliefs under a variety of circumstances. Our norms for reasoning tell us when it is permissible or impermissible to do this. It seems that the norms we actually employ should be subject to external criticism just like any other norms. The norm externalist proposes that we should scrutinize them and possibly replace them by other norms. Because of the direct accessibility problem, we cannot replace them by norms making an explicit appeal to reliability, but we might discover that (1) under certain circumstances inferences licensed by our natural norms are unreliable, and (2) under certain circumstances inferences not licensed by our natural norms are highly reliable. The norm externalist proposes that we should then alter our epistemic norms, adopting new internalist norms allowing us to make the inferences described under (2) and prohibiting those described under (1).

We must distinguish between two construals of the norm externalist proposal. He might be telling us that when we *discover* old reasoning patterns to be unreliable or new reasoning patterns to be reliable then we should alter our norms and our reasoning accordingly. Alternatively, he might be telling us that if old patterns simply *are* unreliable and new patterns *are* reliable, independently of our knowing or believing that they are, then we should alter our reasoning. The first construal seems like an eminently reasonable proposal, and it is one that has been made explicitly by various externalists. For example, in discussing how reliabilist considerations bear on reasoning, Goldman (1980, p. 47) writes:

> At the start a creature forms beliefs from automatic, preprogrammed doxastic processes.... Once the creature distinguishes between more and less reliable belief-

forming processes, it has taken the first step toward doxastic appraisal.... The creature can also begin doxastic self-criticism, in which it proposes *regulative* principles to itself.

But this involves a fundamental misconception. Our epistemic norms are not subject to criticism in this way. Our *reasoning* is subject to such criticism, and the criticism can dictate changes in our reasoning, but this does not lead to changes in our epistemic norms. This is because, unlike other norms, our epistemic norms already accommodate criticism based on reliability. The point is twofold. First, discovering that certain kinds of inferences are unreliable under certain circumstances constitutes a defeater for those inferences and hence makes us unjustified in reasoning in that way, and this is entirely in accordance with our natural unmodified epistemic norms. For example, we discover that color vision is unreliable in dim lighting, and once we discover this we should cease to judge colors on that basis under those circumstances. But this does not require an alteration of our epistemic norms, because color vision only provides us with defeasible reasons for color judgments, and our discovery of unreliability constitutes a defeater for those reasons. This is entirely in accordance with the norms we already have. Second, discovering that some new inferences are reliable under certain circumstances provides us with justification for making those inferences under those circumstances, but this is licensed by the norms we already have. That is precisely what induction is all about. For example, I might discover that I am clairvoyant and certain kinds of "visions" provide reliable indications of what is about to happen. Once I make this discovery it becomes reasonable for me to base beliefs about the future on such visions. Again, this is entirely in accordance with the norms we already have and does not require us to alter those norms in any way. The general point is that the kinds of reliability considerations to which the norm externalist appeals can lead us to reason differently (refrain from some old inferences and adopt some new inferences), but this does not lead to any change in our epistemic norms. Epistemic norms are unique in that they involve a kind of feedback, having the result that the sort of external criticism that could lead to the modification of other action-guiding norms does not necessitate any modification of epistemic norms.

I have had several externalists respond to this objection by protesting that they do not see the point of distinguishing between considerations of reliability leading us to alter our reasoning, and those

considerations leading us to alter our norms. But if all the externalist means is that considerations of reliability can lead us to alter our reasoning, then he is not disagreeing with anyone. In particular, he is not disagreeing with paradigmatic internalists like Chisholm and me. Norm externalism becomes nothing but a pretentious statement of a platitude.

The alternative construal of norm externalism takes it to be telling us that if old patterns of reasoning are unreliable and new patterns are reliable, then regardless of whether we *know* these facts about reliability, we should not reason in accordance with the old patterns and we should reason in accordance with the new patterns. What could the point of this claim be? It cannot be taken as a recommendation about how to reason, because it is not a recommendation anyone could follow. We can only alter our reasoning in response to facts about reliability if we are appraised of those facts. However, normative judgments do not always have the force of recommendations. This is connected with the distinction that is often made in ethics between subjective and objective senses of "should". To say that a person subjectively should do X is to say, roughly, that given what he believes (perhaps falsely) to be the case he has an obligation to do X. To say that he objectively should do X is to say, roughly, that if he were apprised of all the relevant facts then he would have an obligation to do X. Judgments about what a person subjectively should do can serve as recommendations, but judgments about what a person objectively should do can only serve as external evaluations having some purpose other than guiding behavior.[20] The subjective/objective distinction can be regarded as a distinction between evaluating the person and evaluating his act. The subjective sense of "should" has to do with moral responsibility, while the objective sense has to do with what act might best have been performed.

We can draw a similar subjective/objective distinction in epistemology. The epistemic analogue of moral responsibility is epistemic justification. A person is being "epistemically responsible" just in case his beliefs are justified. In other words, epistemic justification corresponds to *subjective* moral obligation. What determines whether a belief is justified is what else the epistemic agent *believes* about the world (and what other directly accessible states he is in) – not what is in fact true about the world. This seems to show that whatever considerations of *de facto* reliability may bear upon, it is not epistemic

justification. They must instead bear upon the epistemic analogue of objective obligation. What is the analogue? There is one clear analogue: objective epistemic justification is a matter of what you should believe if you were apprised of all the relevant truths. But what you should believe if you were apprised of all the relevant truths is just *all the truths*. In other words, the epistemic analogue of objective justification is *truth*. There is nothing here to give solace to a norm externalist.

Goldman draws a somewhat different distinction between two senses of "justified" in epistemology.[21] He distinguishes between "theoretical" evaluations of reasoning and "regulative" evaluations (the latter being reason-guiding). He suggests that the theoretical sense of justification is the sense required for knowledge and that it is to be distinguished from the reason-guiding sense. He suggests further that his Historical Reliabilism concerns the theoretical sense.[22] The proposal is that it is knowledge that provides the point of a norm externalist's evaluation of epistemic norms in terms of considerations of reliability unknown to the epistemic agent. I do not believe that, but even if it were true it would not affect my overall point. The sense of epistemic justification with which I am concerned in this paper is the reason-guiding sense, and if it is acknowledged that norm externalism bears only upon another sense of justification then my main point has been conceded.

To summarize the discussion of externalism, one can be an externalist by being either a belief externalist or a norm externalist. These exhaust the ways in which externalist considerations might be brought to bear on our epistemic norms. The belief externalist tries to formulate epistemic norms directly in terms of externalist considerations, but it is impossible to construct reason-guiding norms in this way. The norm externalist proposes instead to recommend changes in reason-guiding norms on the basis of considerations of reliability. But this appeal to reliability is redundant because it is already incorporated in our unadulterated internalist norms. Thus, as far as I can see, externalism has nothing to contribute to the solution to traditional epistemological problems. Justified beliefs are those resulting from normatively correct reasoning. Consequently, any evaluation of the justifiedness of a belief must be reason-guiding and hence must be beyond the pale of externalism.

4. EPISTEMIC NORMS AND THE INDIVIDUATION OF CONCEPTS

The account of epistemic norms proposed in section 2, coupled with the apparent failure of norm externalism, leaves us with a puzzling problem. Internalists have typically assumed that whatever epistemic norms we actually employ are automatically correct. But that seems hard to reconcile with the seemingly obvious fact that it is at least logically possible for different people to employ different norms. Surely, if Smith and Jones believe *P* for the same reasons, they are either both justified or both unjustified. There is no room for their justification to be relative to idiosyncratic features of their psychology resulting in their employing different epistemic norms. This seems to imply that there is just one set of correct epistemic norms, and the norms a person actually employs may fail to be correct. This conclusion would seem to be obvious if it were not for the fact that there is no apparent basis for criticizing a person's norms. That is precisely what norm externalism tries unsuccessfully to do. The reliabilist considerations to which the norm externalist appeals are the only plausible candidates for considerations of use in criticizing and correcting epistemic norms, and we have seen that our epistemic norms cannot be corrected in this way. Of course, I might criticize Jones' norms simply because they disagree with mine, but he could equally criticize mine because they disagree with his. Are we committed to a thoroughgoing epistemological relativism then? That is at least unpalatable.

The solution to the problem of relativism can be found by turning to another problem. This is the problem of how concepts are individuated. The standard view takes concepts to be individuated by their truth conditions. The claim of this theory is that what makes a concept the concept that it is are the conditions that must be satisfied for something to exemplify that concept. These conditions comprise its truth conditions. The precise content of the truth condition theory of concepts deserves closer inspection than it usually receives. There is one sense in which the truth condition theory of concepts is true but also completely trivial and uninteresting. The truth condition of the concept *red* is the condition of *being red*, and the truth condition of the concept *blue* is the condition of *being blue*. The following is undeniable:

red = *blue* iff *being red* = *being blue*

but it is hardly illuminating. Rather than explaining the concept, the truth conditions presuppose the concept. We might just as well define the "identity condition" of a physical object to be the condition of *being that object* and then claim that physical objects are individuated by their identity conditions. That is about as unilluminating as you can get.

Typically, philosophical logicians slide back and forth between the vacuous claim that concepts are individuated by their truth conditions and the considerably more contentious claim that concepts can be informatively characterized by (and only by) giving truth condition analyses of them. A truth condition analysis of a concept is an informative statement of necessary and sufficient conditions for something to exemplify the concept. I think it is fair to say that many philosophical logicians do not clearly distinguish between the vacuous claim and the contentious claim, or at least take the vacuous claim to somehow directly support the contentious claim. But I see no reason to think there is any connection between the two claims.

There is another strand to this story. Traditionally, the only logical relations between concepts that were recognized by philosophers were entailment relations. Concepts, "logical items", were supposed to be individuated by their logical properties, and it seemed that the only logical properties concepts possessed were those definable in terms of their entailment relations to other concepts. This generates the picture of a "logical space" of concepts, the identity of a concept being determined by its position in the space, and the latter being determined by its entailment relations to other concepts. The claim that concepts must have definitions is just a more specific version of this general picture – one alleging that the position of a concept in logical space is determined not just by one-way entailments but by two-way logical equivalences. Some version of this picture has been prevalent throughout much of twentieth century philosophy, and it still plays a prominent role in philosophical logic. I will call this general picture of the individuation of concepts *the logical theory of concepts*. It has typically been either conflated with or identified with the truth condition theory.

The short answer to all of this is that most concepts do not have the kind of definitions required by the logical theory of concepts. Analytic

philosophy in the mid-twentieth century concerned itself almost exclusively with the search for such definitions, and if we can learn anything from that period it is that the search was largely in vain. It is a very rare concept that can be given an informative definition stating truth conditions. This is related to a more purely epistemological problem. What makes something a *good* reason for holding a belief is a function of the content of the belief. If the content of the belief is determined by entailment relations, then those entailment relations must also determine what are good reasons for holding that belief. The only kinds of reasons that can be derived from entailment relations are reasons that are themselves entailments – *conclusive reasons*. Thus, epistemologists were forced to the conclusion that all reasons must be entailments, and they were forced to try to solve all epistemological problems in terms of conclusive reasons. Phenomenalism is a paradigm case of this. Phenomenalist theories tried to give truth condition analyses of physical object beliefs in terms of sense data or ways of being appeared to. Phenomenalist theories are the only possible theories of perceptual knowledge if we suppose that reasons must always be derived from definitions. But phenomenalist theories were invariably unsuccessful. Nowadays everyone will grant that physical object beliefs are not entailed by the way we are appeared to. Perceptual knowledge cannot be explained in terms of entailments; it can only be explained in terms of defeasible reasons. Defeasible reasons cannot be derived from entailments, so it follows that there must be more to the identity of a concept than its entailment relations.

Some philosophers still resist most examples of prima facie reasons. For example, there is less than universal agreement that ⌜x looks-red to me⌝ is a prima facie reason for me to believe ⌜x is red⌝. But there is one example of a prima facie reason that absolutely everyone must acknowledge, and that is induction. These days, no sane philosopher would try to account for inductive reasons entirely in terms of logically conclusive reasons. By the very nature of induction, our inductive reasons do not entail our inductive conclusions. This, of course, was Hume's point. Some philosophers earlier in this century tried to circumvent it by supposing that induction involves an implicit premise stating a principle of "the uniformity of nature".[23] Quite apart from the fact that they could never formulate such a principle, it was finally acknowledged that there was no noninductive way we could ever justify belief in such a principle. The upshot of this is that we must

acknowledge there to be at least some defeasible reasons. And as long as we agree that good reasons for a belief are a function of the content of the belief, we must acknowledge that the content of a belief is not completely determined by its entailment relations.

The next thing to notice is that the logical theory of concepts makes conclusive reasons just as mysterious as prima facie reasons. This has generally been overlooked, but it is really rather obvious. Epistemologists have noted repeatedly that logical entailments do not always constitute reasons. Some entailments are conclusive reasons and others are not reasons at all. The latter is because P may entail Q without the connection between P and Q being at all obvious. For example, mathematicians have proven that the Axiom of Choice entails Zorn's Lemma. These are abstruse mathematical principles apparently dealing with quite different subject matters, and just looking at them one would not expect there to be any connection between them. If, without knowing about the entailment, one were so perverse as to believe Zorn's Lemma on the basis of the Axiom of Choice, one would not be justified in this belief. Once the entailment is known, you can become justified in believing Zorn's Lemma *partly* by appeal to the Axiom of Choice, but your full reason for believing Zorn's Lemma will be the conjunction of the Axiom of Choice and the fact that if the Axiom of Choice is true then Zorn's Lemma is true. You are believing Zorn's Lemma on the basis of this conjunction rather than just on the basis of the Axiom of Choice. You can never become justified in believing Zorn's Lemma on the basis of the Axiom of Choice alone, so the latter is not a reason for the former.

On the other hand, if I justifiably believe both P and $(P \supset Q)$, I *can* justifiably believe Q on the basis of these other two beliefs. In this case I do not have to believe Q on the basis of the more complicated belief:

$$P \text{ and } (P \supset Q) \text{ and if } [P \ \& \ (P \supset Q)] \text{ then } Q.$$

To suppose that each instance of reasoning in accordance with *modus ponens* must be reconstructed in this way would lead to an infinite regress.[24] Thus some entailments are conclusive reasons and others are not. But the logical theory of concepts gives us no way to make this distinction. It characterizes concepts in terms of their entailment relations to other concepts, but, *a fortiori*, all entailment relations are

entailment relations. There is nothing about the entailment relations themselves that could make some of them reasons and others not. Thus conclusive reasons become just as mysterious as prima facie reasons on the logical theory of concepts. This seems to indicate pretty conclusively that the logical theory of concepts is wrong. There has to be more to concepts than entailment relations.

To argue that the logical theory of concepts is wrong is not yet to say what is right. The theory I want to endorse in its place is *the epistemological theory of concepts*. This theory begins by noting that concepts are both logical and epistemological items. That is, concepts are the categories whose interrelationships are studied by logic, and they are also the categories in terms of which we think of the world. The interrelationships studied by logic can all be reduced to entailment relations. Thus logic need not take note of any other features of concepts. Logic can get along with a cruder picture of concepts than can epistemology. But a complete account of concepts must accommodate both logic and epistemology. There is good reason to think that the role of concepts in epistemology is fundamental. Not all entailment relations are conclusive reasons, but it seems likely that all entailment relations derive from "simple" entailment relations, where the latter just are those that are conclusive reasons. Thus a theory of concepts adequate for epistemology will very likely be adequate for logic as well. The question then becomes, "What kind of theory of concepts is adequate for epistemology?" In epistemology, the essential role of concepts is their role in reasoning. They are the categories in terms of which we think of the world, and we think of the world by reasoning about it. This suggests that concepts are individuated by their role in reasoning. What makes a concept the concept that it is is the way we can use it in reasoning, and that is described by saying how it enters into various kinds of reasons, both conclusive and prima facie. Let us take the *conceptual role* of a concept to consist of (1) what are reasons (conclusive or prima facie) for thinking that something exemplifies it or exemplifies its negation and (2) what conclusions we can justifiably draw (conclusively or prima facie) from the fact that something exemplifies the concept or exemplifies the negation of the concept. My proposal is that concepts are individuated by their conceptual roles. The essence of a concept is to have the conceptual role that it does. If this is right, the explanation for how there can be such things as prima facie reasons become trivial. Prima

facie reasons are primitive constituents of the conceptual roles that characterize concepts. Prima facie reasons need not somehow have an origin in something deeper about concepts, because there is nothing deeper. In an important sense, there is nothing to concepts over and above their conceptual role. To describe the conceptual role of a concept is to give an analysis of that concept, although not a truth condition analysis.[25]

The epistemological theory of concepts lays to rest the spectre of epistemological relativism. Epistemological relativism is the view that (1) different people could have different epistemic norms that conflict in the sense that they lead to different assessments of the justifiedness of the same belief being held on the same basis, and (2) there is no way to choose between these norms. The epistemological theory of concepts enables us to escape any such relativism. Because concepts are individuated by their conceptual roles, it becomes impossible for people's epistemic norms to differ in a way that makes them conflict with one another. The epistemic norms a person employs in reasoning determine what concepts he is employing because they describe the conceptual roles of his concepts. If two different people reason in accordance with different sets of epistemic norms, all that follows is that they are employing different concepts. Thus it is impossible for two different people to employ different epistemic norms in connection with the same concepts. Their conceptual frameworks are determined by their epistemic norms. Epistemological relativism is logically false.[26]

5. A NATURALISTIC INTERNALISM

The main purpose of this paper is to defend a thoroughgoing internalism. To my mind the most serious objection to existing internalist theories is that they are radically incomplete. Although they may give correct descriptions of some of our epistemic norms, they provide no systematic account of epistemic justification. They do not tell us what epistemic justification is all about and they do not explain why we have the epistemic norms we do. This objection can now be met. Epistemic justification consists of holding beliefs in conformance to correct epistemic norms. But as we have seen, our epistemic norms are constitutive of the concepts we have and hence it is a necessary truth that our actual epistemic norms are correct. Thus we can give an

entirely adequate analysis of epistemic justification as follows:

> A person's belief is justified iff he holds it in conformance to his epistemic norms.

This is a naturalistic analysis of epistemic justification. Reasoning is a natural process. It is something we know how to do. To say that we know how to do it is to say that it is governed by norms. Our epistemic norms are, by definition, the norms that actually govern our reasoning. This is a naturalistic definition of "epistemic norm", and accordingly the above analysis of epistemic justification is noncircular and naturalistic.[27] Of course, I have not proposed an informative logical analysis of the governance process which forms the basis of these definitions, but that should not be expected. This is a natural process that we can observe in operation, and its nature can be clarified by psychological investigations. But it must be emphasized that the only clarification that can be expected here is empirical clarification. We can no more provide an informative logical analysis of the governance process than we can provide an informative logical analysis of electrons or magnetism. These are natural kinds and natural processes that we discover in the world, and their nature is revealed by empirical investigation – not logical analysis.

No doubt some philosophers will be disturbed by the fact that my analysis of epistemic justification does not characterize justified beliefs in terms of a single general property (like reliability) intrinsic to the beliefs, but instead characterizes justified beliefs in terms of the reasoning underlying them. But that is just the way things are. What makes a belief justified is its being supported by reasoning of an approved sort, and there is no reason to think there are general intrinsic properties of beliefs that determine whether that is possible. This is connected with the charge that internalist theories give piecemeal characterizations of epistemic justification. That is only a difficulty if there is something more to be given and hence something is being left out. To clarify this point, let us distinguish between a characterization of epistemic justification in the sense of an analysis of epistemic justification, and a characterization in the sense of an epistemological theory. I gave an analysis above. I will understand an epistemological theory, on the other hand, to be a theory that attempts to describe our epistemic norms. There is nothing piecemeal about my analysis of epistemic justification, but an epistemological theory will

automatically be piecemeal. This is a consequence of the nature of reason-guiding (or more generally, action-guiding) norms. Such norms tell us that under certain circumstances we are permitted to do various things and not permitted to do other things. These norms have to be rather specific because, as we saw above, they must take as input only features of the present circumstances that are directly accessible to our automatic processing systems. This precludes the possibility of the norms appealing to sweeping general features of the circumstances (features like the belief being produced by a reliable process). Compare the norms for bicycle riding. These are going to be very specific, including such things as "If you feel yourself losing momentum then push harder on the pedals" and "If you think you are falling to the right then turn the handlebars to the right". Epistemic norms will be equally specific, telling us things (approximately) like "If something looks red to you and you have no reason for thinking it is not red then you are permitted to believe it is red". There is no more reason to think that we can combine all epistemic norms into one simple general formula than there is for thinking there is a single simple formula governing the use of the pedals, the handlebars, the brakes, etc., in bicycle riding. Action-guiding norms cannot work that way.

It is illuminating to contrast this account of epistemic norms with more conventional internalist formulas. Internalists have been inclined to say instead that our epistemic norms describe the way we *actually reason*. This claim has played an important role in internalist epistemology, because it tells us how to find out what proper epistemic norms are – just examine the way we actually reason.[28] But this is at least misleading. We do not always reason correctly, and what epistemic norms describe is *correct* reasoning. We might similarly be inclined to say that our bike riding norms describe the way we actually ride a bicycle, but even when we know how to ride a bicycle we sometimes make mistakes and fail to conform to our norms – I might be distracted by a pretty girl and lose my balance. Thus we might more accurately say that our bike riding norms describe the way we actually ride a bicycle when we do it correctly. This formulation, however, sounds vacuous. After all, riding a bicycle correctly or reasoning correctly is just conforming to the norms. This creates a real puzzle for traditional accounts of action-guiding norms. The puzzle is resolved by seeing how norms for doing something are connected with

knowing how to do it. The best way to describe the connection between norms and actual behavior is to say, as I did above, that our bike riding norms and our epistemic norms are the norms that *actually guide us* in riding bicycles and reasoning. This is similar, in a very important respect, to the more customary claim that our epistemic norms describe the way we actually reason. In each case, norms are to be elicited from what we actually do and not from some mysterious criterion, separate from our actual behavior, that tells us what we should do. But there is also an important difference between the present formulation and the traditional formulation. The present formulation does not take our reasoning behavior at face value. It recognizes that we can reason incorrectly. That need not confound us in formulating epistemic norms because, by virtue of knowing how to reason, we know how to evaluate reasoning, and so we can recognize correct and incorrect reasoning when we see it (although not necessarily with perfect reliability). This recognition process is part of the internal "non-intellectual" process whereby our norms govern our behavior. The process is non-intellectual in the sense that it does not involve our making any conscious explicit comparison of our behavior with some explicitly formulated paradigm; the process goes on under the surface. But even though we cannot consciously monitor the process, we can make use of the results by noting that under certain circumstances we judge some behavior to be permissible and other behavior to be impermissible. On the basis of these individual (normative) observations we can try to construct a general theory of right reasoning or correct bicycle riding.

This general account of epistemic norms and epistemological theories has important implications for philosophical methodology. Epistemological theories are supposed to give general accounts of "right reasoning" – that is, they purport to describe our epistemic norms. It is a contingent psychological fact that we have the norms we do. Equivalently, it is a contingent psychological fact that we employ the conceptual framework we actually employ. Does this mean that epistemological theories are contingent? This is a rather complicated question. The answer is, "Partly 'yes', and partly 'no'". Part of what we do in epistemology is elicit our actual epistemic norms, and that really is a contingent matter. But our ultimate conclusions are to the effect that particular concepts have conceptual roles of certain sorts. The conceptual role of a concept is an essential property of that

concept, so it seems that our ultimate conclusions are, if true, necessarily true. Let us take this a bit more slowly, looking at each step of what transpires in an epistemological analysis.

We begin with a question like, "How are we justified in forming beliefs about the colors of objects?", i.e., "What are the conceptual roles of color concepts?" We begin our investigation by trying to determine how we actually make such judgments. This is a matter of eliciting the epistemic norms we actually employ. That is a question about human psychology. But this does not mean that the best way to go about answering it is by performing laboratory experiments. To illustrate, consider a simpler case. Typing is an excellent example of something we learn to do automatically. When we learn to type we internalize norms telling us what to do and then we follow those norms automatically. Now suppose we want to describe those norms. Consider the question, "What finger do you use to type a 'w'?" We *could* try to answer that question by designing a laboratory experiment in which we observe people typing "w"s under a wide variety of circumstances, but that would be silly. There is a much easier way to do it. We can *imagine* typing a "w" and observe what we do. Touch typists find themselves using their left ring finger. How can this work as a way of eliciting our norms? After all, we are not just asking what finger a person uses on a particular occasion, and people do not always type correctly. What we want to know is what finger our typing norms prescribe using to type a "w". The reason we can answer this question by performing our thought experiment is that there is an introspectible difference between conforming (or at least trying to conform) to one's internalized norms and not conforming. It is this fact that led us to the discovery of epistemic norms in the first place. We could perversely type the "w" with our right index finger, but if we did we would know that we were not doing it the way we learned to do it.

Now consider the epistemological question, "How do we judge that something is red?", where this is intended to be a question about our epistemic norms. Sometimes we reflect upon actual judgments we observe ourselves making. More often we *imagine* making such judgments under normal circumstance and see what goes on. For example, suppose we are considering the hypothesis that something's looking red to us gives us a prima facie reason for thinking it is red. We imagine being in situations in which things look red to us and note

that if there are no "intervening" considerations we will come to believe that the object is red. This is not just an observation about what actually happens. It is an observation about what we know *to do* in judging colors, i.e., an observation about how our automatic processing system actually guides us in reasoning about colors. Next, suppose it is asked whether, in acquiring justification from this prima facie reason, it suffices to merely believe no defeaters, or if one must instead have the positive belief that there are no true defeaters. We might imagine being in situations in which we believe no defeaters but have not given the matter any thought, and so have no beliefs one way or the other about whether there are any true defeaters. We find that under such circumstances we would judge the object to be red on the basis of its looking red to us, and so we conclude that our epistemic norms permit us to make the inference without having the belief that there are no true defeaters.

This illustrates what goes on in epistemological analysis. Our basic data concern what inferences we would or would not be permitted to make under various circumstances, real or imaginary. This data concerns individual cases and our task as epistemologists is to construct a general theory that accommodates it. Epistemologists have often supposed that our epistemic rules should be, in some sense, self-evident.[29] I have been arguing that many of the individual bits of data on which our epistemological theory is founded will, in a certain sense, be self-evident (more accurately, introspectible). By virtue of knowing how to reason we know how to tell right reasoning when we see it, and that provides us with our data. But that does not guarantee that it will be easy to construct theories describing our epistemic norms or that such theories will be obviously right once we have them. One complication both in the use of thought experiments and in interpreting our data is that because our automatic processing system operates in a non-intellectual way without any conscious monitoring, it need not be obvious to us what makes a particular belief justified even when it is evident to us that it is justified. Our data consists in the fact that various beliefs *are* justified – not *why* they are justified. This can be illustrated by reflecting upon the fact that we have a much better account of perceptual knowledge than we do of many other kinds of knowledge. Internalist theories of perceptual knowledge have been worked out in fair detail and seem reasonably convincing.[30] According to such theories, our being "appeared to" in various ways provides us

with prima facie justification for holding beliefs about our physical surroundings. Being appeared to in a certain way is a nondoxastic psychological state of a sort of which we can often become consciously aware by introspection. But it is important to realize that introspection can give us no direct reason to think that we are typically in such a state in a normal case of perception. Introspection is incompatible with normal perception. If you consider introspectively the way things appear to you then your perceptual situation is no longer normal. The claim that our beliefs in normal perception are based upon our being appeared to in various ways is a contingent psychological theory and cannot be regarded as a self-evident philosophical datum. Nevertheless, we regard it as a well established psychological fact, and so have no misgivings about assuming it in constructing an account of our epistemic norms.

Contrast epistemological theories of perceptual knowledge with those of a priori knowledge. We have no very good theories of a priori knowledge despite the fact that we have no difficulty telling which beliefs are justified and which are not when we are doing mathematics or logic. In other words, we know how to proceed in a priori reasoning, and hence we have the same kind of basic data as in the case of perception – we can recognize some beliefs as justified and others as not. What we lack in the case of a priori knowledge is a psychological account of what is going on when we have justified beliefs. This illustrates both the way in which our basic epistemological data are self-evident and the importance of contingent non-self-evident psychological facts in the construction of epistemological theories. In an important sense, describing our actual epistemic norms is part of psychology. This does not mean that it is best carried out in the laboratory, but the results of standard psychological investigations can be relevant.

The contingent enterprise of describing our actual epistemic norms is not all there is to epistemology. From a description of our epistemic norms, we want to draw conclusions about the conceptual roles of various concepts, and that is a matter of conceptual analysis. But conceptual analysis is supposed to provide us with necessary truths. How is it possible to derive necessary truths from contingent psychological generalizations? In order to answer this question, note first that true statements about the necessary properties of things need not be necessarily true. To take a well worn example, nine is the number of planets, and nine is necessarily such that it is odd, so it

follows that the number of planets is necessarily such that it is odd; but the latter is only contingently true. This is because the necessity involved is de re rather than de dicto. Similarly, a statement describing the necessary properties of a concept must refer to the concept in some way, and if the mode of reference is only contingently a way of referring to that particular concept, then even though the property ascribed to the concept is an essential property of the concept, the resulting statement will be contingent. Applying this to epistemology, in describing epistemic norms we are describing essential properties of concepts, but this does not mean that our epistemological pronouncements are themselves necessary truths. It depends upon how we are thinking of the concepts. For example, we might be thinking of the concept *red* under some description like, "what is ordinarily expressed by the word 'red' in English". The meaning of an English word is a contingent matter, and so the claim that the concept *red*, so conceived, has such-and-such a conceptual role, will be a contingent claim about necessary properties of concepts.

Although conceptual analyses need not be expressed by necessary truths, there will be necessary truths lurking in the wings. We can think of concepts in various ways that are only contingently ways of thinking of those concepts, but it is also possible to think of a concept in a "direct" fashion that is necessarily a way of thinking of that particular concept. This is more obvious in the case of propositions than in the case of concepts. When we think about a proposition, perhaps judging it to be true or false or necessary, we rarely think about it under some contingent description it just happens to satisfy. For example, I *might* think of a proposition under the description "The first proposition entertained by Bertrand Russell on the morning of April 3, 1921", but I will not ordinarily be able to ascribe truth or falsity to the proposition so conceived unless I know what proposition it is. To know what proposition it is is to be able to think of it in another way, "in terms of its content", and know that the two propositions are the same. Similarly, I might think of a concept as "Immanuel Kant's favorite concept" but that will not help me in judging that Holly exemplifies that concept unless I know what concept it is, and knowing what concept it is involves being able to think of it in another, noncontingent, way.[31] In logical contexts, when we think of concepts and propositions we usually think of them in this direct fashion.[32]

If you think about a concept in this direct fashion and you ascribe

an essential property to it, then your belief is necessarily true. A conceptual analysis describes essential properties of concepts, so if the conceptual analysis is expressed by a proposition that is about the concept directly then that proposition is necessarily true. Thus conceptual analyses do generate necessary truths. But they are not a priori truths. The analyses describe the conceptual roles of concepts, and our knowledge of those conceptual roles is derived from the discovery of contingent psychological generalizations regarding what epistemic norms we employ in reasoning. Thus the ultimate issue of epistemology is necessary a posteriori conceptual analyses.

NOTES

* I want to thank Alvin Goldman, Roderick Wiltshire, Stewart Cohen, Hilary Kornblith, and Frederick Schmitt for helpful comments on earlier versions of this paper.

[1] I will say more below about what makes a state "internal". Paradigmatic internalist theories are those of Chisholm (1977) and (1981), and my (1974) and (1979).

[2] Alvin Goldman (1981) seems to be one of the few externalists who is clear on this distinction. He distinguishes between two senses of "epistemic justification" (see section three) and adopts belief externalism with regard to one and norm externalism with regard to the other.

[3] Two good examples of reliabilist theories are Alvin Goldman (1980) and Frederick Schmitt (1984). There are also reliabilists who focus on knowledge rather than justification. Examples are David Armstrong (1973) and Fred Dretske (1981).

[4] Sosa (1981) (21ff) raises this objection.

[5] We can also make "third person evaluations" of our own past behavior, but that is different from what I am calling "first person uses" of norms.

[6] Many philosophers appear to adopt the intellectualist model, although it is doubtful that any of them would seriously defend it. For example, Goldman (1981) (30ff) appears to assume such an account of epistemic norms. The intellectualist model pervades Hilary Kornblith's (1983) discussion. Unfortunately, it is also prominent in my own discussion in Pollock (1979).

[7] This point has been made several times. I made it in Pollock (1974), and Van Cleve (1979) made it again. Despite this, I do not think that internalists have generally appreciated its significance. (At least, I did not.)

[8] There has been a lot of recent work in psychology concerning human irrationality. Psychologists have shown that in certain kinds of epistemic situations people have an almost overpowering tendency to reason incorrectly. (Much of the psychological material can be found in Kahneman, Slovic, and Tversky (1982) and Nisbett and Ross (1980).) It might be tempting to conclude from this that, contrary to what I am claiming, people do not know how to reason. The short way with this charge is to note that if we did not know how to reason correctly in these cases, we would be unable to discover that people reason incorrectly. To say that we know how to reason is to invoke a competence/performance distinction. It in no way precludes our making mistakes. It does not even

preclude our almost always making mistakes in specific kinds of reasoning. All it requires is that we can, in principle, discover the errors of our ways and correct them. (This is pretty much the same as the assessment offered by Jonathan Cohen (1981). See also the critique of Goldman (1986).)

[9] It might be insisted that this is at least sometimes a misleading way of talking – if our norms for doing *X* tell us to do *Y* whenever we *think* it is the case that *C*, we might better describe our norms as telling us to do *Y* when it *is* the case that *C*. I do not care if one chooses to talk that way, but it must be realized that it has the consequence that although the reformulated norm says to do *Y* when it is the case that *C*, knowing how to do *X* will really only result in our doing *Y* when we *think* it is the case that *C*. This will be important. (And, of course, norms appealing to internal states other than beliefs could not be reformulated in this manner anyway.)

[10] It would also be a wholly implausible theory. We do not invariably have beliefs about the reliability of our inferences whenever we make them, and if norms *requiring* us to have such beliefs also require those beliefs to be justified then they lead to an infinite regress.

[11] One defender of the simple rule is Henry Kyburg (1974). In Pollock (1983) I raised some rather technical objections to basing epistemology on probability in accordance with the simple rule. The present considerations constitute a much deeper objection to any attempt to base epistemology on probability.

[12] We do not ordinarily have any beliefs at all about the probabilities of what we believe. Furthermore, even if we did they would presumably not render our beliefs justified unless the probability beliefs were themselves justified, so we would be threatened by an infinite regress.

[13] Lehrer (1974) p. 201.

[14] This is not altered by the fact that Lehrer's use of "prob" is idiosyncratic, prob(*h*) being what *S believes* is the "objective chance" of *h* being true.

[15] I also distinguished between positive and negative coherence theories. Negative coherence theories allege that a belief is automatically justified unless one has a reason for rejecting it. The observation that epistemic norms must regulate our reasoning while it is in progress rather than merely evaluating it after the fact amounts to a rejection of negative coherence theories, so I will only consider positive coherence theories here.

[16] Pollock (1974) (58ff) and (1979) (99f).

[17] This objection is raised by Michael Williams (1977). It is also pressed by Laurence Bonjour (1978) (10ff). Sosa (1981) mentions the objection, but dismisses it.

[18] The *by*-relation is what Alvin Goldman (1976) calls *level-generation*.

[19] Alternatively, we may have the same norms but your physical skills make you better able to conform to them.

[20] They may serve as recommendations in an indirect fashion by conveying to a person that there are relevant facts of which he is not apprised.

[21] Goldman (1980), 28–29.

[22] Historical Reliabilism is the reliabilist theory he propounds in Goldman (1980).

[23] For example, see Bertrand Russell, (1912), 63.

[24] This was apparently first noted by Lewis Carroll.

[25] This view of concepts is reminiscent of the verification theories of the logical positivists. I first defended a theory of this sort in 'What is an Epistemological Problem?', and in more detail in *Knowledge and Justification*, although in those

publications I talked about "justification conditions" rather than conceptual roles, and used the term a bit more narrowly. This view of concepts is also related to the somewhat cruder views expressed by Michael Dummett (1975) and (1976) and Hilary Putnam (1979) and (1984).

[26] The conclusion that if different people employ different epistemic norms then they employ different concepts may seem puzzling because it appears to make it inexplicable how such people could communicate with each other. But two points should be made here. First, I doubt that there really is any variation in epistemic norms from person to person. I suspect that epistemic norms are species-specific. But even if that conjecture is false, it need create no difficulty for communication. I have argued at length in two recent books that concepts play no direct role in communication. (My entire theory of language is developed in Pollock (1982). A briefer sketch of the theory can be found in Chapter Two of Pollock (1984).) The reader who is concerned with this question should consult those books.

[27] In Pollock (1979), I described an epistemic agent as being subjectively justified in holding a belief iff he is justified relative to what he *takes* to be the correct epistemic norms, and I said that he is objectively justified in holding a belief iff he is justified relative to what *are* the correct norms. In light of the necessary correctness of our actual epistemic norms, that distinction makes no sense.

[28] Chisholm (1977) endorsed this under the label "critical cognitivism", and I endorsed it (in Pollock (1974)) and called it "descriptivism".

[29] This is what Sosa (1981) calls "methodism".

[30] For example, see my own theory in Pollock (1974).

[31] I argued in Pollock (1984) (chapter two) that our being able to think of propositions and concepts in such a direct fashion is responsible for there being modal operators of necessity and possibility rather than just modal predicates.

[32] This is why ordinary language objections to philosophers' use of terms like "directly aware" seemed beside the point. Philosophers are using these terms to express concepts they are thinking about directly. They are not thinking about those concepts under contingent descriptions like "the concept ordinarily expressed by the phrase 'directly aware' in English", and accordingly the objection that the way they are using the term "directly aware" does not conform to ordinary usage is simply irrelevant.

REFERENCES

Armstrong, David: 1973, *Belief, Truth, and Knowledge*, Cambridge University Press, Cambridge.

Bonjour, Laurence: 1978, 'Can Empirical Knowledge Have a Foundation?', *American Philosophical Quarterly* **15**, 1–14.

Chisholm, Roderick: 1977, *Theory of Knowledge, 2nd ed.*, Prentice Hall, Englewood Cliffs.

Chisholm, Roderick: 1981, 'A Version of Foundationalism', *Midwest Studies in Philosophy* **5**, 543–564.

Cohen, L. Jonathan: 1981, 'Can Human Irrationality be Experimentally Demonstrated?', *The Behavioral and Brain Science* **4**, 317–370.

Dretske, Fred: 1981, *Knowledge and the Flow of Information*, MIT, Cambridge.

Dummett, Michael: 1975, 'What is a Theory of Meaning?', in Samuel Guttenplan (ed.), *Mind and Language*, Oxford University Press, New York.
Dummett, Michael: 1976, 'What is a Theory of Meaning? (II)', in Gareth Evans and John McDowell (eds.), *Truth and Meaning*, Oxford University Press, New York.
Goldman, Alvin: 1976, *A Theory of Human Action*, Princeton University Press, Princeton.
Goldman, Alvin: 1980, 'What is Justified Belief?' in George Pappas (ed.), *Justification and Knowledge*, D. Reidel, Dordrecht.
Goldman, Alvin: 1981, 'The Internalist Conception of Justification', *Midwest Studies in Philosophy* **5**, 27–52.
Goldman, Alvin: 1986, *Epistemology and Cognition*, Harvard University Press, Cambridge.
Kahneman, Daniel, Paul Slovic, and Amos Tversky: 1982, *Judgment Under Uncertainty: Heuristics and Biases*, Cambridge University Press, Cambridge.
Kornblith, Hilary: 1983, 'Justified Belief and Epistemically Responsible Action', *Philosophical Review* **92**, 33–48.
Kyburg, Henry, Jr.: 1974, *The Logical Foundations of Statistical Inference*, D. Reidel, Dordrecht.
Lehrer, Keith: 1974, *Knowledge*, Oxford University Press, Oxford.
Nisbett, R. E., and L. Ross: 1980, *Human Inference: Strategies and Shortcomings of Social Judgment*, Prentice Hall, Englewood Cliffs.
Pollock, John: 1968, 'What is an Epistemological Problem?', *American Philosophical Quarterly* **5**, 183–190.
Pollock, John: 1974, *Knowledge and Justification*, Princeton University Press, Princeton.
Pollock, John: 1979, 'A Plethora of Epistemological Theories', in George Pappas (ed.), *Justification and Knowledge*, D. Reidel, Dordrecht.
Pollock, John: 1982, *Language and Thought*, Princeton University Press, Princeton.
Pollock, John: 1983, 'Epistemology and Probability', *Synthese* **55**, 231–252.
Pollock, John: 1984, *Foundations of Philosophical Semantics*, Princeton University Press, Princeton.
Putnam, Hilary: 1979, *Meaning and the Moral Sciences*, Harvard University Press, Cambridge.
Putnam, Hilary: 1984, *Reason, Truth, and History*, Harvard University Press, Cambridge.
Russell, Bertrand: 1912, *Problems of Philosophy*, Oxford University Press, Oxford.
Schmitt, Frederick: 1984, 'Reliability, Objectivity and the Background of Justification', *Australasian Journal of Philosophy* **62**, 1–15.
Sosa, Ernest: 1981, 'The Raft and the Pyramid: Coherence Versus Foundations in the Theory of Knowledge', *Midwest Studies in Philosophy* **5**, 3–26.
Van Cleve, James: 1979, 'Foundationalism, Epistemic Principles, and the Cartesian Circle', *Philosophical Review* 55–91.
Williams, Michael: 1977, *Groundless Belief*, Yale University Press, New Haven.

Department of Philosophy
University of Arizona
Tucson, AZ 85721
U.S.A.