

The Varieties of Epistemic Evaluation

Karl Schafer

November 24, 2011

This is an essay about two basic epistemological questions: First, what is the relationship between the concept of knowledge and the concept of justification? And, second, should epistemological judgments involving either concept behave in the manner that internalist accounts of knowledge and justification imply they should?

In what follows, I'm going to approach these questions by taking a step back from the more familiar epistemological debates in order to ask some basic questions about the nature and purpose of epistemic judgment. In particular, I want to consider an underexplored way of thinking about the nature of epistemic judgment and use this meta-epistemological perspective to shed light on exactly which aspects of our practice of epistemic judgment deserve to be thought in broadly internalist terms and why. Thus, instead of focusing on the nature of epistemic *properties* like knowledge or justification, I'm going focus here on the nature of epistemic *concepts* and *judgments*, and then use the resulting understanding to *explain* why some epistemic judgments obey the sort of constraint that internalists endorse, and why others do not.

In doing so, I will focus on the manner in which forming an epistemic judgment of a particular sort *rationaly commits* one to be disposed to form certain further beliefs. Thus, the key to my account of these issues is the different ways in which judgments involving different epistemic concepts regulate how we form beliefs, in virtue of these rational commitments. My central claim will be that this variation in regulative role is the reason why some, but not all, epistemic concepts obey the sort of constraint internalists endorse.

In particular, consider the two epistemic concepts that have been the traditional focus of epistemological debate: the concepts of knowledge and justification.¹ I will argue that,

¹Unfortunately, "justification" is ambiguous in at least one important respect. One might use the term "justification" to refer to the sort of warrant that is capable of making a true belief an instance of knowledge. Given the arguments to follow, this sort of warrant or "justification" is - like knowledge itself - best thought

given the rational commitments involved in making a judgment about knowledge, whether we treat someone's *true* beliefs as knowledge will often depend on factors, such as reliability and safety, that do not obey any form of the internalist's constraint. Thus, these considerations will yield the admittedly unsurprising result that our judgments about knowledge behave as they would if an externalist account of knowledge were true.

Just the opposite will turn out to be true of the concept of justification. I will argue that, given the rational commitments involved in making a judgment using this concept, it is only rational to allow our judgments about whether or not someone is justified in believing P to be responsive to factors that are cognitively accessible to the individual in question. This will explain why our judgments about justification, insofar as we are rational, behave in the manner they should, if an internalist account of justification were true. Moreover, I hope this explanation will help us to understand the importance of making epistemic judgments of an internalist sort and, by doing so, go some ways toward justifying this practice.

Explaining these features of our practice of epistemic judgment does not, of course, take us all the way to an explanation of why the *properties* of knowledge and justification have an externalist and internalist character, respectively. For of course, it might be the case that our use of these concepts somehow does not mirror the nature of the properties they refer to. In order to close this gap, we would need to consider in detail how the facts about rational judgment that I will be discussing inform the semantics for these judgments. Unfortunately, this is far too complex a topic to tackle here. So for the most part, I will limit myself to conclusions about how a rational thinker would employ the concepts of knowledge and justification. That having been, at least in the case of "non-defective" concepts, it is natural to think that constraints on the rational use of a concept should, if possible, be respected by the property it refers to. So there is at least a *prima facie* case for extending the conclusions about the concepts of justification and knowledge that I will be arguing for here to the corresponding properties. Nonetheless my focus will be on the idea that attending to the role that different types of epistemic judgments play in regulating how we form beliefs can help us both to understand the often obscure relationship between

of in externalist terms. But if these arguments are correct, there is also second fundamental sense in which someone belief's may evaluated epistemically - one which is associated with the sort of justification that is capable of making one's beliefs (subjectively) rational. In what follows, I reserve the term "justification" for the second of these concepts - a concept which I take to be the source of the internalist intuitions discussed below.

the concepts of knowledge and justification *and* to explain why judgments involving the latter, but not the former, should behave in the manner they would, given the truth of internalism.

1. Internalism and Externalism

Before moving on to discuss these issues, I first need to say something about the relationship between internalism and externalism as I will understand it here. This is crucial because there are a variety of ways of defining the difference between these views in the literature - and both sides of these debates tend to some degree to define these views so that their preferred view appears to be the more modest alternative. Thus, since I will (to some degree) be defending internalism about justification, it will hardly be surprising that I will be offering a fairly modest characterization of internalism here. As a result, the view about justification that I arrive at may be compatible with certain, similarly modest conceptions of externalism.² But nonetheless it seems to me to deserve the label “internalism” insofar as it is responsive to the core commitments of internalism I describe below.

As I will understand it, then, the contrast between these two positions comes down to following point of disagreement. Internalists about some epistemic status S place a constraint on the facts that are relevant to this status - claiming that having the status S supervenes only on facts that are cognitively accessible to the believer in some sense.³ So, for example, internalists about propositional justification will claim that the state of being justified in believing some proposition supervenes only on facts of this sort. Externalists about S deny this principle with respect to S - claiming that questions about S often depend, at least in part, on facts that are inaccessible to the believer being evaluated - e.g. the reliability of the processes by which a given belief was formed.⁴

²For a different way of framing these debates, which makes externalism seem like the more modest position, see John Greco’s contribution to Steup and Sosa (eds.) *Contemporary Debates in Epistemology*. (Oxford: Wiley-Blackwell, 2005).

³A crucial question for any internalist is how this notion of “access” is to be spelt out. For instance, if to have access to some fact is just to know it, then this constraint may not place much of a restriction on the nature of justification. Unfortunately, this way of understanding access does not seem to capture the intuitive reactions to cases that I discuss below. Thus, if it is these intuitions that our talk about access is meant to capture, access to the fact that P will have to involve something more than just knowing that P. Shortly I will have more to say about this, but for now I will rest content with our intuitive grasp of this idea.

⁴As I will discuss later on, in the case of the concept of knowledge, the philosophically interesting form of externalism is one which considers these questions, having factored out the trivial sense in which the

The traditional motivation for internalism about a given type of epistemic evaluation rests on our intuitive responses to particular cases. For example, taking the case of justification as our focus, a traditional externalist account of justification of the sort that identifies justification with formation via a reliable method, seems to be blocked by cases of two sorts:

The Evil Demon: Rene's subjective experience of the world is indistinguishable from our own. But, unlike us, he lives in a world that is controlled by an evil demon who has arranged things so that the beliefs he forms on the basis of his perceptual experiences are systemically unreliable. Is Rene justified in taking his experiences of the world at face value when forming beliefs? To most of us it will seem as though Rene is just as justified in doing so as we are. But the externalist (at least of a traditional sort) seems to be forced to give *different* verdicts about these two cases - saying, for instance, that we are justified in taking our perceptual experiences at face value while Rene is not.⁵

Clairvoyance: Larry's subjective state is indistinguishable from our own, but unlike us he possesses a unique faculty of Clairvoyance that reliably produces true beliefs about some subject matter - beliefs that, when formed, simply pop into his head with no further associated phenomenology. Are these beliefs of Larry's justified in the manner more ordinary beliefs are? Once again, it will seem to many of us that they are not. For it will seem to many of us that we are not justified in believing something when such a belief simply "pops" into our heads in this manner.⁶

In each of these cases it appears as though the link the traditional externalist draws between factivity of knowledge attributions gives them an externalist character.

In addition, it is important to stress that this is only one of many possible ways of drawing the internalist/externalist distinction. For a nice taxonomy of some of these, see Jim Pryor's "Highlights of Recent Epistemology", *British Journal for the Philosophy of Science* 52 (2001). Again this is a fairly modest form of internalism. As noted above, by focusing on this sort of internalism, I am conceding a good deal to traditional critics of the view.

⁵Cases of this sort are, of course, traditionally associated with Descartes. For the application of such examples to externalism see, e.g., Stewart Cohen and Keith Lehrer. "Justification, Truth, and Knowledge." *Synthese* 55 (1983).

⁶See Laurence Bonjour. *The Structure of Empirical Knowledge*. (Cambridge: Harvard University Press, 1985). Of course, the debate about the force of examples of this second sort is ongoing. But there is no doubt that they have provided one of the main impetuses for contemporary discussions of internalism, which makes them worthy of discussion.

between purely external aspects of an epistemic agent's situation and their status as justified or unjustified fails to hold.⁷ And while these cases will not be counterexamples to each and every possible externalist theory of justification, it does appear plausible that they make a larger point about the our judgments concerning justification. For in these cases, it does seem as though the facts we are responsive to in our thinking about justification are constrained by the internalist's basic constraint - or at least something like it.⁸

Thus, these cases seem to support what I will call the Central Idea behind all broadly internalist theories. Namely, they support the following idea (applied here to an arbitrary epistemic status S):

The Central Idea: The state of having some epistemic status S supervenes only on facts that are cognitively accessible (in some sense) to the believer.⁹

⁷For a traditional externalist view that appears to run afoul of these examples, see Alvin Goldman's "What is Justified Belief?" in George S. Pappas (ed.). *Justification and Knowledge*, (Dordrecht: Reidel, 1979).

⁸Externalists have proposed a wide variety of different responses to these cases. Some have suggested that there are multiple notions of justification active in them. As noted above, I very much agree on this point, but this response concedes the crucial point that there is an important notion of justification that behaves as the internalist claims.

Others have suggested that an individual's epistemic status depends on whether their belief-forming methods are reliable, not in their world, but in the *actual* world. (For a recent example of this sort of approach, see Juan Comesana's "The Diagonal and the Demon", *Philosophical Studies* 110: 249-266, 2002.) But while this modification captures the intuitions noted above when we consider the Evil Demon world as a counter-factual possibility, it does not capture them when we consider it as a counter-actual possibility. And if we limit ourselves to considering counter-factual as opposed to counter-actual possibilities, then this view will actually satisfy the definition of internalism provided here. Thus, views such as these represent the sort of convergence between sophisticated forms of internalism and externalism that I noted above. In such cases, whether we call the resulting view internalist or externalist seems to me of secondary importance.

⁹An important issue in making this idea precise is the precise sort of supervenience involved in it. For instance, here are two different ways we might read this principle as applied to the case of justification:

- (i) The state of being justified in possessing some belief supervenes only on facts that are cognitively accessible to the believer, across the space of all *metaphysically* possible worlds.
- (ii) The state of being justified in possessing some belief supervenes only on facts that are cognitively accessible to the believer, across the space of all *epistemically* possible worlds.

The difference between these two readings will not matter much if we focus on the case of perceptual epistemology, which will be my focus here. But it has the potential to matter a great deal when we turn to the implications of internalism for the epistemology of necessary truths. For reasons I discuss in more detail elsewhere, it seems to me that it is the second of these claims that captures the full scope of the internalist's central idea. But I will not argue for this in this paper.

This idea captures the minimal requirement that separates internalist theories from their externalist competitors. But this requirement leaves a great deal unsaid about exactly what form an internalist theory of, say, justification should take. Most importantly, it leaves open the question of exactly what facts should be considered “cognitively accessible” in the sense relevant to the internalist’s Central Idea. And, as is familiar from recent discussions of these issues, this notion may be made precise in any one of a number of ways.

In what follows I will - for the most part - remain agnostic about these different varieties of internalism, and will focus instead on giving an explanation of why certain forms of epistemic judgment - in particular judgments about justification, behave as if the internalist’s Central Idea applied to their subject matter. But I do think that my account, together with certain plausible background ideas, does have important implications for these debates - and I will try to say something about these implications as I proceed. Thus, my focus will be on whether or not our use of epistemic concepts of various sorts are constrained by the internalist’s Central Idea, and, in particular, whether it constrains our use of the concepts of knowledge and justification. In doing so, I hope not only to argue that it does apply to some concepts, but also to explain why epistemic judgment that is so constrained is an activity that I would make sense for creatures like us to engage in.

2. Epistemic Judgment as “Plan-Laden”

The foundation for the account I am interested in comes from what should be a somewhat surprising place: the discussion Allan Gibbard gives of judgments of knowledge as “plan-laden” in his recent book *Thinking How to Live*.¹⁰ There Gibbard outlines a theory of the concept of knowledge that rests upon the idea that attributing knowledge to someone always involves planning to defer in one’s own process of belief formation (and possibly practical reasoning) to the beliefs one treats as knowledge. Thus, for Gibbard, “When I defer, I attribute knowledge. Or, at least, I go some way toward attributing knowledge: my deferral may be tentative and reserved, in which case I am tentative and reserved in my attribution of knowledge.”¹¹

¹⁰Cambridge, MA: Harvard University Press, 2003. Gibbard is building here on some remarks of Simon Blackburn in his *Essays in Quasi-Realism* (Oxford: Oxford University Press, 1993). For a similar line of thought, see Robert Brandom’s discussion of reliabilism in chapter 4 of his *Articulating Reasons* (Cambridge, MA: Harvard University Press, 2000).

¹¹225. This quote is a bit misleading - as it suggests that Gibbard’s point is that whenever we defer, we also attribute knowledge, as opposed to visa versa. But it is relatively clear that Gibbard intends to defend the view I attribute to him, e.g.: “In the preceding chapter .. I treated knowledge as reliability: to regard

Understanding the concept of knowledge in this way involves seeing it as, in Gibbard's terminology, essentially plan-laden. That is, under this account, to judge that someone's belief constitutes knowledge always involves the formation of a plan for future belief-formation and reasoning - namely, a plan to defer to this belief when considering what to believe (and possibly do).¹²

It is important to distinguish the claim that the concept of knowledge is plan-laden in this sense from the claim that whenever I judge that S knows that P, I am implicitly judging *that* I plan to defer this belief. Such a view would imply that part of the content of my ascription of knowledge to S is a claim about my own mental state. This is not terribly plausible, but thinking of the concept of knowledge as plan-laden does not have this implication. To say that the concept of knowledge is plan-laden is to say that whenever I judge that S's beliefs constitute knowledge I *plan* to defer to these beliefs - which is quite different from the claim that whenever I make this judgment, I *judge* that I so plan.

This having been said, it is natural to worry about both of the main aspects of Gibbard's view of these issues. That is, it is natural to wonder whether judgments about knowledge are actually *plan-laden* in Gibbard's sense. And, even if it is accepted that this is the case, it is also natural to wonder whether such judgments involve the sort of plans for *deference* that Gibbard claims they do.

To see why we might be skeptical about the latter claim, consider a case in which a teacher is examining a student in order to evaluate their knowledge of the previous lesson's material. After quizzing the student for several minutes, the teacher judges that the student really does know the material. But despite making this judgment, surely the teacher has no tendency or disposition to defer to the student about these issues. And so surely there is no sense in which the teacher plans to defer in this way - as Gibbard's account suggests he should.

This case indicates a quite general feature of knowledge attributions: namely, that Joe as knowing is to plan to defer, in certain conditions, to judgments like his." (248)

The idea that attributions of knowledge are connected with something like deference is not unique to Gibbard. For an early statement of a similar idea see Austin's comment that "I know that P" functions (at least in part) as a performative with roughly the significance of "I guarantee that P". See J.L. Austin. "Other Minds." *Proceedings of the Aristotelian Society*. Supplementary Vol. 20:148-187 (1946).

¹²Such epistemic plans are quite similar to what Alvin Goldman ("The Internalist Conception of Justification" in *Epistemology: Internalism and Externalism*, ed. Hilary Kornblith (Oxford: Blackwell, 2001)) calls "doxastic decision procedures" (or DDPs). Of course, Goldman introduces this notion to criticize a form of Internalism, but he does not consider the form of Internalism to be defended below. In any case, he seems to have no objection to the idea of a DDP as such.

when someone attributes knowledge to someone they believe to be less knowledgeable than themselves, they will have no tendency to defer to the beliefs of the person to whom they are attributing knowledge. In part, this is the product of the fact that attributions of knowledge play a number of different roles in our overall epistemic practice.¹³ For example, as contemporary virtue epistemologists have stressed, one important function of the attribution of knowledge is to give credit to the knower for the truth of his beliefs.¹⁴ Of course, it is very natural to understand cases like the teacher-student case in terms of just this feature of knowledge attributions.

So one important lesson here is that even if there is a connection between attributing knowledge and deference, this connection does not exhaust the function of knowledge attributions, contrary to what Gibbard sometimes seems to suggest. But cases like the one above also indicate that if there is a connection between knowledge attribution and deference, it cannot be as straightforward as Gibbard's comments suggest. For once again, when we attribute knowledge to someone we take to be less knowledgeable, we will not be inclined to defer to their beliefs.

Fortunately for Gibbard, this does not mean that there is *no* connection between the concept of knowledge and deference. For when someone who has no independent source of knowledge about some issue attributes knowledge to someone else about it, it would be irrational for them not to defer to the knowledgeable individual's beliefs. Thus, while an unrestricted connection between knowledge and deference is implausible, a more restricted variant of this connection remains plausible. In other words, in order to avoid counter-examples like the one above, we need to make the deference that is connected with attributing knowledge *conditional* so that in treating someone's beliefs as knowledge, one plans to defer to those beliefs, *provided that one does not have further potential sources of knowledge available to one*.¹⁵ After all, in cases in which one has several potential sources

¹³Compare Matt Weiner, "The (Mostly Harmless) Inconsistency of Knowledge Ascriptions". *Philosophers' Imprint* 9 (1):1-25 (2009).

¹⁴For example, see Ernest Sosa. *Apt Belief and Reflective Knowledge, Volume 1: A Virtue Epistemology*. (Oxford: Oxford University Press, 2007) and John Greco. *Achieving Knowledge*. (Cambridge: Cambridge University Press, 2010.).

¹⁵For the related view that to say "Smith knows that Q" is to approve of Smith as an informant about Q see Edward Craig. *Knowledge and the State of Nature: An Essay in Conceptual Synthesis*. (Oxford: Clarendon Press, 1990).

As Craig points out in a similar vein, the central case to consider for understanding this aspect of the concept of knowledge is one in which "we as yet have no belief about p, want a true belief about it one way or the other, and seek to get it from someone else". (11)

of knowledge available to one, it is plausible that one should not defer completely to any of these, but should instead form one's beliefs on the basis of an attempt to weigh all of these sources of evidence against one another.

Thus, the attribution of knowledge about some issue to someone can be thought of as recommending deference to them insofar as one does not have any independent source of knowledge about the matter in question. This, of course, is a much weaker connection between the concept of knowledge and deference than the connection between the two that Gibbard presupposes, but, as we will see, it has important implications for our understanding of the concept of knowledge nonetheless.

This, though, leaves unanswered the first, more general worry about Gibbard's account. That is, do judgments about knowledge - or epistemic judgments more broadly - *really* involve planning in the manner this account suggests? It is a familiar theme in contemporary philosophical discussion of these issues that it is simply incoherent to think that we ever *strictly speaking* form plans or intentions for belief formation.¹⁶ And, even if the possibility of such planning is allowed, it does not seem as though the manner in which epistemic judgments regulate our beliefs actually requires the formation of any plans or intentions in a strict sense. Rather, it seems as though such regulation requires only that an epistemic judgment itself be present, and epistemic judgments differ in several important respects from plans or intentions as they are normally conceived. Thus, there seems to be something misleading in Gibbard's discussion of judgments about knowledge as plan-laden - for strictly speaking such judgments do *not* seem to involve planning at all.

If we take planning in a sufficiently strict or narrow sense, I think these comments are correct. But there is something important about the nature of epistemic judgments that we can learn from Gibbard's approach. For we need not understand Gibbard's comments to imply that judgments about knowledge involve planning in a strict sense. Rather, we might take Gibbard to be making a general point about the functional and rational role of judgments involving epistemic normative concepts - namely, that these judgments play a regulative role for the judger's process of belief formation that is *very like* the role that would be played by plans for future belief formation, if we did form such plans. Thus, while

¹⁶For a classic argument of this sort, see Bernard Williams. "Deciding to Believe." in *Problems of the Self* (Cambridge: Cambridge University Press, 1973.) For a more recent variant on this theme, see Pamela Hieronymi. "Believing at Will." forthcoming in *Canadian Journal of Philosophy*.

Ultimately, I'm not sure there is anything *incoherent* in this idea, although there are good reasons to doubt that human beings have much of capacity for epistemic planning in a strict sense.

I do think that it can be misleading to speak of such judgments as plan-laden, I do think that there is an important sense in which the functional and rational role of such judgments *mimics* the functional and rational role of plans for belief formation and deference of the sort Gibbard describes.

For example, a plausible way of capturing at least part of the role that judgments about knowledge play in regulating belief-formation involves noting that judging that S knows that P involves the formation of a mental state that has the following features: (i) All other things being equal, this state rationally commits one to defer to S's belief, provided that one has no independent source of knowledge concerning it. And (ii) as a result, all other things being equal, this state will dispose one to defer to S's belief in this way insofar as one is rational.¹⁷ This, of course, is one important element of the functional and rational role of a Gibbardian plan for deference. Thus, it seems plausible that judging that someone's beliefs constitute knowledge involves entering into a mental state that, in these central respects, plays the functional and rational role of a plan for belief-formation of this sort.¹⁸

What this means is that there is, after all, a weak sense in which we can speak of epistemic normative judgments as "plan-laden". For suppose we say that to w-plan to X under circumstances C is simply to be in a mental state that rationally commits one to X given C in the manner just described. Then we can reasonably speak of epistemic judgments as involving planning *of this weak sort*. For all that this requires is that such judgments involve the functional and rational elements just outlined.

To think of "plan-ladenness" in this way is, of course, to give special weight to the idea that certain judgments or other mental states carry with them certain rational commitments to further mental states or actions. But while there is a great deal of disagreement of the source and nature of such rational commitments, the existence of commitments in the sense relevant to the argument to follow is *relatively* uncontroversial. For example, Gibbard's own account of "plan-ladenness" gives pride of place to primitive relations of disagreement or rational incompatibility between mental states, and this is just another way of capturing the notion of rational commitment I will be interested in here. And more

¹⁷Note that the scope of the set of beliefs in question is determined by our knowledge attribution. Thus, these comments do not raise any (additional) generality problem for the view being outlined.

¹⁸Or at least it will involve the formation of a disposition of this sort in cases in which the believer has a full grasp of the concepts involved and nothing interferes with the manifestation of his grasp of them.

Of course, we will not want to say that such an epistemic judgment shares *every* feature of the associated plan's functional and rational role, only that it shares the features noted above. Otherwise, I might be unclear why the judgment should not count as a plan in the ordinary sense of the world after all.

broadly, it is fairly uncontroversial that a satisfactory account of the nature of judgments involving some concept must explain the manner in which these judgments rationally commit one to certain further judgments under certain circumstances. So there should not be anything terribly worrisome about my reliance on such a notion here, even if my claims about the *particular* rational commitments that associated with the concepts of knowledge and justification are controversial.

Crucially, this way of thinking about epistemic judgments does not require us to give the sort of expressivist account of them that Gibbard does. For this weak sort of w-plan-ladenness is, in fact, compatible with any one of a number of more substantive accounts of the nature of these judgments. For example, it is of course compatible with the sort of account that Gibbard gives of them. But it is also compatible with an account of these judgments that focuses on their conceptual or inferential role - so long as this role, once spelled out, implies that these judgments play the functional and rational role just described. And it is also compatible with a traditional conceptual analysis of epistemic concepts, so long as this analysis delivers the same result. So this way of thinking of epistemic judgments in no way requires that we give a non-cognitivist account of them - although it is, of course, compatible with doing so.

This seems to me to represent the core insight in Gibbard's talk of such judgments as "plan-ladden". But it may well seem to some readers that in so weakening the notion of plan-ladennes we have moved far enough away from our ordinary notion of planning to make such talk potentially misleading at best. So, for the most part, in what follows I will focus on the crucial issue of the rational commitments involved in certain epistemic judgments - although at times I will use the expression "w-plan-laden" to capture this feature of them.

Of course, on any plausible view, knowledge judgments will involve more than just the rational commitments that are involved in being w-plan-laden. For there are also purely descriptive constraints on what counts as an instance of knowledge. So a full account of the concept of knowledge will conceive of this concept as having descriptive elements that are independent of the rational commitments noted above. For example, if we choose to think of the concept of knowledge in terms of its conceptual role, we should accept that this conceptual role involve both straightforwardly descriptive elements and elements that account for the rational commitments involving deference noted above. Thus, the best model for the concept of knowledge will be so-called "thick" normative concepts that involve both descriptive and w-plan-laden constraints on correct usage.

One further point is important in this regard. In conversation, when discussing these ideas, I often encounter the suspicion that acceptance of the idea that epistemic judgment is plan-laden even in this very weak sense involves the acceptance of a form of epistemic voluntarism that many find implausible.¹⁹ From what I have just said, we can see that this is not the case. All that viewing epistemic judgments as plan-laden (in the weak sense I am interested in) requires is that we treat these judgments as having a functional and rational role, one component of which are the sort of rational commitment noted above. And all that this implies is that such judgments can sometimes, in virtue of their functional/rational role, affect how we form beliefs - which is far from any form of belief voluntarism.²⁰

3. Justification and Epistemic Planning

In introducing these ideas, I've followed Gibbard in focusing on the concept of knowledge. But now I want to turn - for the moment - to the concept of justification, since it is there that the most interesting consequences of this approach seem to me to lie.

How then can we make sense of judgments about justification as involving the sorts of rational role that would make them w-plan-laden? In order to do so, we need to give an account of what it is judge that someone's belief that P is justified that generates the result that forming this judgment involves entering into a mental state that mimics the relevant elements of the functional and rational role of a sort of plan for belief formation. But what exactly is the functional/rational role that is relevant here?

The hypothesis I want to explore is that judgments about justification involve a connection of the following sort:

Justification Attribution: Judging that S's belief that P is justified involves entering into a mental state that (all other things being equal) rationally commits one to believe that P *when in S's situation*.²¹

¹⁹To be clear, I don't mean to endorse the anti-voluntarist line, just to indicate that nothing I'm saying commits one to it.

²⁰Similarly, since the understanding of epistemic justification I am interested in does not mention of any notion of "doing our epistemic duty", there is no way of forcing a commitment to epistemic voluntarism on me from this direction.

²¹This will require a bit of modification - see below - but it captures the essential idea I am interested in.

As those familiar with recent debates about rational requirements will be well aware, the "all other things being equal clause" is required to avoid certain counter-examples to the idea that normative judgments generate rational commitments. The precise nature of these connection is, of course, the subject of much dispute - but I will leave these complicated issues to the side here.

As I will go on to explain in a moment, it would be very natural and useful for us to form judgments that play this sort of role in regulating how we form beliefs. And, as we will see later on, if we suppose that judgments about justification play this role, this will explain why these judgments seem to obey some form of the Internalist's Central Idea. Thus, by the end of this paper, I hope it will be plausible that we do in fact engage in a form of epistemic judgment that has this character and that this form of epistemic judgment is the source of internalist intuitions in epistemology.²² In other words, although I hope that there is something intuitive about Justification Attribution, the full support for this principle will come via a sort of philosophical inference to the best explanation.

To explore these ideas, though, we need to see what Justification Attribution involves. If this principle is true, then if one is rational, and all other things are equal, judging that S's belief that P is justified will dispose one to believe that P when in S's situation. Obviously, getting a handle on what exactly contributes to someone's "situation" will be crucial for understanding the nature of this conception of justification. But for now I want to set this question aside for the most part to focus on the general outlines of this sort of account.²³

According to this account, judgments about justification rationally commit one, not to form beliefs in a certain way *in the here and now*, but rather to form beliefs in a certain way *under certain counterfactual conditions*. In this way, when we come to endorse someone's epistemic behavior as justified, this will involve the formation of what might call a hypothetical first-person epistemic w-plan to act as they have acted when in their (first-personal) circumstances.

Counterfactual planning of this general sort is something we engage in all the time - and with good reason. Nothing is more natural or more common than for us to consider someone's situation and to form some hypothetical plan for action for the situation in question. For example, suppose I am reading a biography of Napoleon, and upon finishing

²²By contrast, the question of whether we use the term "justification" to capture this form of epistemic evaluation is relatively unimportant to me here.

²³Still, some features of what someone's situation must involve are clear from the start. In particular, when we consider someone's situation, we will want to hold constant most of the features of their environment and their psychological state. But we *may not* want to hold constant their views about epistemic questions, for we may not want their views on these matters to decide our response to the information that they have. (Ultimately, we will sometimes want to evaluate their response taking this aspect of their state of mind into account - while other times, we will want to set it aside. Thus, there will be a degree of contextual variability with respect to these sorts of question built into our concept of justification.)

a chapter discussing the details of the battle of Waterloo, find myself wondering what I “would have done” had I been in Napoleon’s shoes. When I ask myself this question, I am not really asking myself a descriptive question concerning what someone with my psychology would have done in the situation in question. Rather, I am asking myself what to do in such a situation. Or, in other words, I am forming a hypothetical plan of the sort under discussion.

As was true in the case of the concept of knowledge, this way of thinking about judgments about justification is compatible with any one of a number of more substantive accounts. For example, we might flesh out this picture of judgments about justification by giving an account of the conceptual role of the concept of justification as involving something like the following inference scheme:

S’s belief that P is justified.
I am in S’s situation.
—
P

After all, if this inference rule is one element in the conceptual role of the concept of justification, these judgments will have the functional and rational role laid out above.

In any case, as in the case of the concept of knowledge, I want to remain agnostic about how the picture of such judgments as involving these rational commitments is best fleshed out, since what I say in what follows does not depend on the particular manner in which this idea is developed. What I do want to stress is how intuitive this way of thinking about the concept of epistemic justification is. To see this, let’s consider some simple cases. For example, suppose that I judge that Bob’s beliefs about the weather are justified. Then, in so far as I am rational, I will be disposed to form the same beliefs as Bob when placed in Bob’s situation. Otherwise, I will be disposed, when in this situation, to form beliefs other than those I regard as justified - and this involves an obvious sort of epistemic irrationality.²⁴ Similarly, suppose that I judge that beliefs formed on the basis of

²⁴A slight complication here involves the possibility that I might judge that more than one response to Bob’s situation is justified. Given this judgment, if I am rational, then I will be disposed to respond to Bob’s situation in one of these justified ways. But this need not imply that I am disposed to respond to it exactly as Bob does. Perhaps Bob responds to his situation in one of the ways I deem justified, while I am disposed to respond to it another of them. Such situations are compatible with the spirit of my account, but they require a slight modification to Justification Attribution as stated above:

sense-perception are justified. And suppose that I am not disposed to form beliefs on this basis when confronted with the relevant perceptual experiences. Once again, this involves an obvious sort of epistemic irrationality - for, once again, I am disposed in this case to form beliefs that differ from the beliefs I take to be justified under the circumstances. Thus, part of our intuitive understanding of the concept of justification is the following constraint: if we are rational, we will never judge that some belief is justified in some situation without also being disposed to form this belief in this situation. It is just this connection between judgments about justification and rational dispositions to form new beliefs that Justification Attribution expresses.

4. Justification and Internalism

What I want to do now is to explore the relationship between this way of understanding the concept of epistemic justification and internalism. In doing so, I will proceed in two steps. First, I will explore how thinking of the concept of justification in these terms explains why our judgments about justification should behave in the manner that is consistent with the Internalist's Central Idea. And then, I will explore whether or not this way of thinking of the concept of justification supports any particular form of this Central Idea.

Traditional arguments for internalism have recently come under increasing attack. For example, consider the argument for internalism that Alvin Goldman considers and rejects in his "Internalism Exposed".²⁵ This argument begins with what Goldman calls the "guidance-deontological" conception of justification. Now, while this conception of justification is different from the one I am working with here, there are enough similarities between the two for the comparison to be illuminating. For Goldman, the central idea behind this conception of justification involves the conjunction of two claims. First, that being justified in believing that P is fundamentally a matter of having done one's epistemic duty in so believing (the deontological component). And, second, that it is an individual's "epistemic duty to guide his doxastic attitudes by his evidence, or by whatever factors

*Justification Attribution**: Judging that S's belief that P is the *unique* justified response to his situation involves entering into a mental state that both disposes one to believe that P *when in S's situation* and (all other things being equal) rationally commits one to do so.

Given this modification, such cases are fully compatible with my account. But, for simplicity's sake, I will focus on the earlier formulation of Justification Attribution in what follows.

²⁵In Kornblith, *Epistemology: Internalism and Externalism*. (Cambridge: MIT Press, 2001).

determine the justificational status of a proposition at a given time.”

Taken together, these factors do, Goldman admits, suggest that, “If a person is going to avoid violating his epistemic duty, he must know, or be able to find out, what his duty requires.” But this alone, as he argues at some length, does not imply that internalism is true - for, at this stage of the dialectic, it remains undecided whether it is possible for believers to know (in the relevant sense) that various external facts obtain. If this is possible, then the constraint just derived does not speak in favor of any interesting form of internalism. In order for the internalist to rule this possibility out, he must give some further argument to the effect that the guidance-deontological conception of justification implies that these facts cannot play a justificatory role. And if he can do *this*, then the argument Goldman discusses is simply beside the point.

This is sufficient to doom the argument for internalism that Goldman discusses, and it may also be sufficient to doom many other traditional arguments of this sort. But this does not mean that the internalist lacks any compelling grounds for his position. For beginning with the conception of justification outlined above, it is possible to construct an argument for internalism that does not have the flaws that Goldman discusses.

To understand why this way of understanding judgments about justification supports internalism, we need to briefly discuss one further aspect of the nature of planning in general: A consistent or rational set of plans may distinguish between two sets of circumstances only when it is possible (at least in principle) for the person carrying out the plan to recognize the difference between them.²⁶ Thus, when forming and implementing such plans we must be careful not to plan both to do A in situation C1 and to do NOT A in situation C2, where C1 and C2 are such that we could not distinguish between being in C1 and C2 when in C1 or C2. Otherwise our plans would be fundamentally incoherent. It would be impossible to actually follow a plan that distinguished between situations that were indistinguishable from one another *from the perspective of the person trying to follow the plan*.

Crucially, this does *not* mean that we must always form plans whose antecedents are such that we can *always* tell whether they obtain. For we form such plans all the time. For example, right now I am planning to water the garden if it doesn't rain tonight. But, of

²⁶Gibbard puts this point by saying that a plan must be couched in “recognitional terms”, but this way of speaking obscures the fundamental issue here. For an example of the problems with this way of making this point, see the exchange between Gibbard and John Hawthorne in *Philosophy and Phenomenological Research*(2001).

course, there are conditions under which it will be impossible for me to tell whether it rains tonight. Fortunately, given the above constraint, there need be nothing irrational about such plans. For this constraint does *not* claim that planning to do A given C requires one to *also* plan to A in all circumstances that are indistinguishable from C. Rather, one may form no such further plan and remain fully coherent according to the constraint on rational planning just laid out. All that this constraint rules out is that I could rationally plan to do A under some condition C1 and simultaneously plan to do NOT A under some other condition C2 such that I could not distinguish being in C1 from being in C2.

Thus, the point I am making here is that forming a plan to do A under some condition C1 rationally commits one *not* to plan to do NOT A under an indistinguishable condition C2 - *not* that planning to do A under C1 rationally commits one to also plan to do A under C2 as well. So, for example, if I plan to water the garden if it doesn't rain tonight, this does commit me to *not* planning *not* to water the garden under conditions that are indistinguishable from it not raining, but it need not commit me to actually planning to water the garden under these conditions. For example, it does not commit me to planning to water the garden if an Evil Demon makes it seem to me that it is raining. And it certainly does not imply that I actually do plan to water the garden under these conditions.²⁷

Thus this constraint is fully compatible with the rationality of plans with these sorts of antecedents, and this constraint remains plausible, even after the consideration of such cases. But, of course, even if this constraint applies to planning in a normal sense, this does not show that it applies to judgments about justification as characterized above - for these judgments involve "plans" to believe only in the weak sense that they rationally commit one to form certain beliefs (and so, insofar as one is rational, dispose one to do so). In other words, these judgments involve mere w-planning as opposed to planning in an ordinary sense. Thus, in order to render our discussion here relevant to the case of judgments about justification, we need to show that this constraint on rational planning applies to these weaker cases of "planning" as well.

Fortunately, it is plausible that this is also the case. For what is it about planning in an ordinary sense that makes the above constraint plausible? It is just the fact that planning

²⁷For the same reason, the account of justification I defend here does *not* entail that someone can be justified in believing P only when they are capable of distinguishing worlds in which P is true from worlds in which it is false - a condition that is sometimes referred to as sensitivity. Rather, it only requires that if someone is justified in believing P in one world, they must be justified in believing P in all indistinguishable worlds, be they world in which P is true or worlds in which P is false.

to A (all other things being equal) rationally commits me to A. After all, suppose that I plan both to do A in situation C1 and to do NOT A in situation C2, where C1 and C2 are such that we could not distinguish between being in C1 and C2 when in C1 or C2. What seems irrational about this is precisely that this combination of states will rationally commit me to A'ing in C1 and B'ing in C2, where C1 and C2 are indistinguishable for me. For while this rational commitments will not conflict *all things considered*, they will conflict insofar for subject who is trying to determine which situation he is in. But if this is the reason why the above constraint on rational planning is true, then the same constraint should apply to any states that generate rational commitments of this sort. And so this constraint should apply to judgments about justification, as described above, just as much as it applies to ordinary planning states.

Thus, we seem to be entitled to extend this constraint to apply to all cases of w-planning. And with this constraint in hand, we are in a position to explain why the rational commitments involved in judgments about justification generate a notion of justification whose rational behavior is consistent with the Internalist's Central Idea. For given this constraint, whenever we enter into states that rationally commit us to believing certain things under certain circumstances, we may distinguish between states of affairs only in so far as we could make these distinctions in the circumstances in which these rational commitments would apply. Or, in other words, we can only distinguish between states of affairs, in so far as we could make these distinctions in the circumstances we are w-planning for. But when judging whether S's beliefs are justified, our judgments involve rational commitments to believe certain things in S's situation. So these judgments may distinguish between this situation and others only in so far as we could make these distinctions in S's situation. Thus, some piece of information will be relevant to our evaluation of whether S's beliefs are justified, only if S can distinguish the state of affairs in which this information obtains from the state of affairs in which it does not. In other words, some piece of information is relevant to whether or not S's beliefs are justified only if this piece of information is accessible to S (in the sense just laid out).²⁸

In judging whether S's beliefs are justified, then, we are forming w-plans about what to believe in S's situation. And these w-plans may distinguish between states of affairs only in

²⁸This argument has similarities to some more familiar arguments for internalism. The most similar version I know of is the argument offered by Pollock and Cruz in their *Contemporary Theories of Knowledge, 2nd Edition*. (Lanham: Rowman and Littlefield, 1999). I discuss the relationship between my argument and the one offered by Pollock and Cruz below.

so far as we could make these distinctions in S's situation. To see the implications of this, consider the two "pro-internalist" cases described above. In the case of the Evil Demon, we saw that agents who are fundamentally unreliable nonetheless seem to us to be fully justified in believing what they do. This is just the result that our account would predict. By definition the individuals within the Evil Demon world are unable to distinguish their situation from the one we find ourselves in now. Likewise, we are unable to distinguish our situation from theirs. As such, given the above constraint, if we enter into a state that rationally commits us not to take their perceptual experiences at face value in forming beliefs in their situation, this would make it irrational to also be in a state that rationally commits us to take our actual perceptual experiences at face value. Thus, entering into a state with such a rational commitment would cripple our everyday belief forming practices in the here and now.²⁹

Something similar is true in cases like Clairvoyance. Once again, by stipulation, the Clairvoyant's situation is indistinguishable from the situation of someone who is simply prone to having completely random and unreliable beliefs "pop" into his head. So, if we were to enter into a state that rationally commits us to believe what the Clairvoyant believes when in his situation, we would be irrational to also be in a mental state that rationally commits one not to form beliefs in a similar manner under our actual circumstances. But many of us do exactly this - since many of us judge that it would be irrational for people like us to believe any old belief that "pops" into our head.

Personally, I don't find the answers to these questions obvious. But these are the sorts of questions that push many philosophers towards the claim that the Clairvoyant's "special" beliefs are unjustified. Thus, the conception of justification we are working with fits nicely with the cases that are often taken to fuel internalist treatments of justification. And the point made by our treatment of these cases is a perfectly general one. The constraint on w-planning noted above will guarantee that our w-plans regarding what to believe, when in two different situations, should differ only when we can distinguish being one of these situations from being in the other. Or, in other words, this constraint will rule out, as irrational, being in a set of mental states that collectively commit us to different responses to situations that are internally indistinguishable from one another.

²⁹Here I have been assuming that the reader rejects skepticism about our current state of epistemic justification. The same basic points may be made from a skeptical direction - only then, we will w-plan neither to take our perceptual experiences at face value in the here and now nor to take them at face value in an Evil Demon world.

This will in turn guarantee that if we judge that some belief B is justified under some circumstances C, we are rationally required *not* to judge that B is *unjustified* under any circumstances that are indistinguishable from C. And *visa/versa*, which is enough to generate the result that we should form judgments about justification in a manner that is consistent with one version of the internalists' Central Idea: namely, one that cashes out the relevant sort of access in terms of the believer's ability to distinguish the obtaining of a fact from its non-obtaining:

*The Central Idea**: The state of being justified in possessing some belief supervenes only on facts, the obtaining of which the believer can distinguish from the non-obtaining.

Moreover, it does so in a manner that avoids the problems that have been pointed out for other defenses of internalism of this sort. In particular, remember the complaints Goldman raised against more traditional arguments for internalism. The main problem for the argument that Goldman considers is that it generates an internalist conclusion only if one presupposes that the sort of knowledge required by the "guidance-deontological" conception is itself internalist in nature. And this, of course, is something no externalist will accept. The argument I have been discussing, on the other hand, does not at any point either explicitly or implicitly appeal to anything of this sort. Rather, all it appeals to is (i) a certain characterization of the functional and rational role of judgments about justification and (ii) a basic constraint on when we can rationally enter into mental states that involve certain rational commitments, which rests on our ability to distinguish being in one possible situation from being in another.

Of course, an externalist might, following Goldman, challenge me to explain why this rational constraint should be understood in terms that generate an internalist result. But here there is little room for serious debate. To see this, simply focus your attention on the notion of indistinguishability that drives the internalist's views - namely, the sort of indistinguishability that is at work in cases like *The Evil Demon*. And then ask whether it is rational to be in mental states that generate rational commitments to perform incompatible actions in situations that are indistinguishable from one another *in this sense*. I think it is clear that it is not. After all, if one was in a state of mind that involved such rational commitments, and one did find oneself in one of the situations to which these commitments apply, one would simply be at a loss as to which of the two commitment to comply with. Thus, the notion of "being able to distinguish" that is relevant here is one that delivers a

broadly internalist result.³⁰

Moreover, the understanding of internalism expressed by the Central Idea* is superior to the forms of internalism that Goldman considers in other important respects. For one, since the Central Idea* is a supervenience claim, it does not imply that for every belief there must be some mental state or proposition (“a justifier”) that justifies this belief and to which the believer has access. For instance, it is plausible that many epistemically important logical and general probabilistic facts are never accessible to ordinary believers in any internalistic sense. But because the most important of these facts hold with logical necessity, the supervenience claim expressed by the Central Idea* does not require that believers actually be able to distinguish between situations in which these logical facts obtain and situations in which they do not - for there simply are *no* situations of the latter sort. Thus, even if many of our beliefs are justified only because these logical and probabilistic facts hold, the form of internalism I am arguing for here does not imply that we must be able to distinguish situations in which these facts hold from situations in which they do not.³¹

Thus, the form of internalism the considerations I have been discussing support is superior to many traditional forms of internalism. With this in mind, let me say something about how my argument compares with similar arguments for internalism within the recent philosophical literature. The most similar argument in the literature is the “Refutation of Belief Externalism” offered by John Pollock and Joseph Cruz in their *Contemporary Theories of Knowledge*. In fact, in many respects, my argument may be viewed as an attempt to make more precise the intuitive considerations behind the argument they give. After all, much as I have been doing, Pollock and Cruz focus on the functional role of epistemic judgments and the way in which these judgments regulate the manner in which we form beliefs. And they also attempt to argue that epistemic judgments can play this sort of regulatory role only if they are constrained in the manner the internalist’s Central Idea suggests.

³⁰A focus on this sort of epistemic w-planning also pushes us towards viewing the internalist’s Central Idea in terms of supervenience across all epistemically (as opposed to metaphysically) possible worlds. For it is the epistemically possible scenarios that we are concerned with w-planning for - whether not those possibilities are, as a matter of fact, metaphysically possible. As noted above, this fact is of central importance for the proper internalist account of our beliefs concerning necessary truths.

³¹Here is an area where the precise sort of supervenience involved the internalist’s Central Idea may matter. But, in fact, since the most important basic logical and probabilistic truths hold, not just in all metaphysically possible worlds, but also in all epistemically possible worlds, the difference between the two readings I noted above should not matter much here.

Still, despite these similarities, Pollock and Cruz's argument seems to me considerably less successful than the argument just given. According to Pollock and Cruz, an epistemic norm can guide our belief formation in the manner they consider essential only if the circumstance-types to which it is responsive are ones to which the believer has direct access - where direct access is access unmediated by any belief. But it is not at all obvious why *this* form of "direct access" is required here. After all, our actions are often guided by rules and plans - like my plan to water the garden if it doesn't rain tonight - that involve circumstance-types to which we do not have direct access in Pollock and Cruz's sense.³² So, from Pollock and Cruz's discussion it remains unclear why the role that epistemic judgments play in regulating our belief formation should have internalist consequences.

As I have been arguing, focusing on the nature of rational epistemic w-planning can help us to close this gap. By focusing on the rational constraints on *all* rational planning - including w-planning - we generate the result, discussed above, that it is never rational to w-plan to do A in circumstances C1 and w-plan to do NOT A in circumstances C2, when we could not distinguish between being in C1 and being in C2. And this general constraint on the nature of rational planning explains why the circumstance types that our epistemic w-plans are responsive to must satisfy the internalist's basic constraint, without thereby claiming that there is anything irrational in forming w-plans that have antecedents to which we do not have direct access.

Thus, this argument seems to me to succeed at showing that our judgments about justification should obey some variant of the Internalist's Central Idea, where other similar arguments for this claim have failed. But just what form of this Central Idea is this view of the concept of justification compatible with?

The crucial question here is the following: just when will two situations be "indistinguishable" from one another in the sense we have been discussing? Two ways of answering this question seem to me attractive. One of these begins with the thought that two individual's situations should be taken to be indistinguishable from one another just in case the two individuals in question possess the same *phenomenally conscious mental states*. This is a position that is often labeled Mentalism. But there are a number of at least *prima facie* problems with the Mentalist's answer to these questions. After all, there are many ways in which two individuals' situations may differ from one another, which seem to have justificatory significance, despite not involving any difference in their conscious phenomenology.

³²Ralph Wedgwood makes this point nicely in his discussion of Pollock and Cruz's argument in "Internalism Explained", *Philosophy and Phenomenological Research* 65 (2002), 349-369

For example, there is the relevance that stored beliefs and other phenomenally unconscious states seem to have for questions of justification.³³

Thus, I am not sure that the correct way to understand the relevant sort of indistinguishability is the Mentalistic one just discussed. Rather, it seems to me far more natural to understand this notion, and with it, the internalist's notion of accessibility, in terms of whether an individual's mechanisms for complying with their epistemic judgments have personal-level access to some piece of information. Or, to use a piece of terminology from Ned Block, it seems more natural to understand this notion of accessibility in terms of *access consciousness* (or something like it) as opposed to *phenomenal consciousness*.³⁴ After all, remember that the crucial question here is whether the person in question is capable of distinguishing between different alternatives insofar as they are in the business of complying with the rational commitments that their judgments about justification generate. And while it does seem that such a capacity requires that the distinctions in question be available to the agent's personal level mechanisms for belief formation, it is not at all obvious that it requires that these distinctions be ones that the agent is phenomenally conscious of or that these distinctions be distinctions between possible states of affairs described in purely mental terms.

Of course, it is possible that additional argument can take us these further steps, but here I will be content with establishing a weaker form of "Informational Internalism". After all, it is this sort of internalism, and not the stronger claims of Mentalism, that seems to me to be the intuitive core of internalism about justification. And it is this sort of internalism that is most directly supported by my arguments. Thus, we end up with the following claim about justification:

Informational Internalism: The state of being justified in possessing some belief supervenes only on information to which the believer's personal-level mechanisms for regulating belief formation have access.

This view is weaker than some traditional forms of internalism. So it is reasonable to ask whether it deserves the title of "Internalism"? Ultimately, as noted above, how we answer

³³For a nice discussion of this, see Goldman's "Internalism Exposed". There are, of course, many possible responses to these problems, and I do not want to suggest that any of them are in the end decisive, only that it may be best to explore these issues from a perspective that deemphasizes the role of conscious access.

³⁴"On a Confusion about a Function of Consciousness." *Behavioral and Brain Sciences* 18, 2 (1995): 227-247.

this question may not matter much. But it is worth noting here that this view agrees with the internalist's Central Idea - making sense of this idea in terms that involve an important restriction on the facts that are relevant to justification: namely, a restriction to those facts to which an individual's systems for regulating belief formation have informational access on the personal level. So there is at least an important element of internalism about justification present here.

5. Knowledge and Justification

If this is correct, then we can take the idea that judgments about justification involve certain rational commitments and use this idea to explain why these judgments function in a manner that is consistent with the Internalist's Central Idea. In some ways, this is the main result I hope to establish in this essay. But to make the significance of this result plain, I want to return now to the concept of knowledge and consider why the manner in which this concept involves certain rational commitments of a broadly similar sort does not generate a similar result. This, I hope, should help to make plain why it is that we have *both* an internalistic concept of justification and an externalistic concept of knowledge in our basic "toolbox" for epistemic evaluation.

My discussion of the concept of knowledge will be less ambitious than my discussion of the concept of justification for a number of reasons. First, as noted above, there is some reason to think that knowledge attributions play a variety of roles within our ordinary thought and talk. Thus, we should be cautious about any attempt to provide a simple characterization of the nature of these attributions. And, secondly, the introduction of the concept of knowledge raises a number of difficult issues concerning the relationship between knowledge and justification that I want to remain agnostic about here.

In particular, there is the question of whether we want to approach epistemological questions from a "justification-first" or a "knowledge-first" direction.³⁵ That is, should we take one of these concepts to be prior to the other so that it is possible to understand the latter in terms of the former. Or should we instead take neither of these concepts to have priority over the other - taking a "neither first" view of these issues.

The general account of epistemic judgment and the particular account of judgments about justification I am developing here is compatible with all three of these options. For

³⁵For a "knowledge-first" approach see, for example, John McDowell, "Knowledge and the Internal". *Philosophy and Phenomenological Research* 55.4: 877-93 (1995) and Timothy Williamson, *Knowledge and Its Limits* (Oxford: Oxford University Press, 2000).

example, if we wish to take a traditional, justification-first stance towards these issues by taking justification to be one element in what knowledge requires, we can easily do so via beginning with the above characterization of Justification Attribution and adding further elements to it, including the connection between the concept of knowledge and deference discussed below, to generate a characterization of the nature of knowledge attributions.

Alternatively, if we are attracted to the idea that epistemology should begin with the concept of knowledge and then attempt to understand the concept of justification in terms of it, we may adopt an independent account of the concept of knowledge and then explore how the concept of justification behaves in the context of this account plus the constraint on the concept of justification derived above. The details of this story will depend a good deal on just how we characterize the concept of knowledge, but part of the story is likely to be the following line of thought. Suppose, as is common in knowledge-first circles, that we take knowledge to be our primary aim in forming beliefs. Then presumably we will w-plan to believe P if by doing so we will achieve knowledge. But given the above constraint on rational w-planning, if we form this w-plan, then we must also not w-plan to do otherwise in any situation that is indistinguishable from situations of this first sort. Thus, combining the above account of the concept of justification with an independent account of the concept of knowledge will generate the result that we are justified in believing P whenever we are in a situation that is *indistinguishable* from a situation in which that belief would constitute knowledge.³⁶ And this is just how many prominent proponents of a knowledge-first epistemology have moved from knowledge to notions of subjective rationality or justification.³⁷

In this way, my comments are entirely compatible with a knowledge-first understanding of the proper shape of epistemological theorizing in general. And of course they are also compatible with the third possibility noted above, which takes there to be a much looser connection between the concepts of knowledge and justification so that neither of these concepts takes priority over the other. The internalist conception of justification sketched above is quite flexible in this regard.

³⁶Alternatively, if we assume that knowledge implies subjective rationality, we can argue as follows:

1. If S knows that P, then S is justified in believing P. (Assumption)
2. If S is justified in believing P, and T's situation is subjectively indistinguishable from S's situation, then T is also justified in believing P. (The Internalist's Central Idea as applied to justification)
3. Thus, if S knows that P, and T's situation is subjectively indistinguishable from S's situation, then T is justified in believing P.

³⁷Compare, for example, McDowell's discussion in "Knowledge and the Internal".

What I do take to be plausible, though, is that the concept of knowledge involves elements other than the elements identified above in the case of the concept of justification. And although it is not entirely clear how precisely these elements are best characterized, it is not hard to see why they will make it the case that the concept of knowledge does not satisfy the Internalist's basic constraint - even when we factor out the factivity of knowledge attributions from our discussion.

To see why this is plausible, I want to consider the implications for this issue of Gibbard's claim that attributing knowledge to someone involves the formation of a w-plan to defer to their beliefs under certain circumstances. As noted above, in order for this claim to be plausible, it will have to be qualified so as to apply only to cases in which one does have further independent sources of evidence about the issues in question. But once this qualification is in place, the existence of a connection of this sort appears to me quite plausible. Moreover, considering the implications of this connection, I want to suggest, will help to indicate why the general account of epistemic judgment being developed here helps to explain why it is natural for us to use both internalist and external epistemic concepts in our practice of epistemic judgment.³⁸

In considering this, it is important to distinguish between two different senses in which externalism might be taken to be true about knowledge - one of which is trivial, while the other is not. First of all, we might say that externalism is true of knowledge because facts about knowledge do not obey the internalist's Central Idea - at least as we have been understanding it here. But this is hardly very interesting - since it follows directly from the claim that knowledge is factive, plus the characteristic internalist claim that we do not always have "access" in the relevant sense to all of the facts about which we have beliefs.

That our judgments about knowledge treat the knowledge facts as "external" in this weak sense does in fact follow from the idea that knowledge attributions involve a rational commitment to deference, since it would hardly be rational to defer to someone's beliefs if one took them to be false. But there is also a stronger sense in which the concept of knowledge appears to be externalist in character. For, even if we factor out the role of factivity, it appears that whether or not some belief counts as knowledge will depend on

³⁸As noted above, there are a wide variety of views in the neighborhood of Gibbard's view. And most, if not all, of these will have the implications to be sketched below. Much the same is true for views that connect together knowledge and practical reasoning. For a recent proposal in this *general* area see Jeremy Fantl and Matthew McGrath's "On Pragmatic Encroachment in Epistemology" in *Philosophy and Phenomenological Research* 75.3 558-89 (2007).

factors that are inaccessible to the subject in the sense discussed above.

Once again, this further externalist aspect to the concept of knowledge follows naturally from the connection between knowledge attributions and deference. For example, given this connection, in judging that S knows that P we are, in part, forming a conditional w-plan to defer to S's belief that P under certain circumstances. But suppose then that we are faced with some belief of S's and are trying to decide whether or not this belief is something to defer to in this way. What factors will we feel entitled to appeal in making this decision? Well, we'll appeal to any information we can get our hands on - provided that we would continue to have access to this information in the situation in which our plan for deference is relevant. In particular, there is no reason to think that this information will be limited to information to which the *subject* of our judgment has access in any sense. For example, in considering this issue, we'll want to know as much as possible about the reliability (broadly construed) of the methods by which S came to believe that P - whether or not this is something they have access to.

What this means is that our judgments about whether or not someone knows that P will not in general respect the fundamental constraint Internalism's Central Idea places on epistemic judgment. We will not, when attributing knowledge, limit the information we take to be relevant to whether an individual knows that P to the facts that are accessible to him in anything like the internalist's sense. So, whether we take some individual to have knowledge will depend on *all* the information we as potential deferrers have about their situation, whether this information is accessible to them or not.

6. Conclusion: The Diversity of Epistemic Evaluation

As we have been discussing, there are a variety of different ways in which we can form states that, insofar as we are rational, will serve to regulate how we form beliefs in the future. At least two of these forms of epistemic self-regulation are particularly relevant to epistemological questions. On the one hand, there is the sort of self-regulation that we engage in when we wonder whether or not to defer to someone's beliefs. And, on the other, there is the sort of self-regulation that we engage in when we consider what to believe when placed in certain counterfactual situations.

If what I have been saying is correct, the second of these forms the core of our practice of making judgments about epistemic justification. And it is plausible that something like the first of them forms *part* of our practice of making knowledge attributions. As I have argued, this account of the nature of epistemic judgment provides us with an elegant

explanation of why the concept of justification has an internalist character, while the concept of knowledge does not. And it makes intelligible why our overall practice of epistemic evaluation should have both internalist and externalist components. Thus, if this is correct, then by considering the nature of epistemic self-regulation and its connections to the concepts of knowledge and justification, we can arrive at an understanding of the nature of epistemic evaluation that derives both internalist and externalist features of such evaluation from a common source.