Why Justification Matters

Declan Smithies

Forthcoming in *Epistemic Evaluation: Point and Purpose in Epistemology*, edited by J. Greco and D. Henderson. Oxford University Press, 2014.

Justification is one among many dimensions of epistemic evaluation. We evaluate beliefs not only for justification and the lack of it, but also for truth and falsity, reliability and unreliability, knowledge and ignorance, and so on. Moreover, justification comes apart from these other dimensions of epistemic evaluation – thus, justified beliefs fall short of knowledge when they are false or when they are true but unreliable. Accordingly, one of the central tasks for a theory of justification is to explain what justification is and how it differs from these other dimensions of epistemic evaluation.

On the traditional analysis, knowledge is justified true belief. Justification, on this analysis, is the property that turns true belief into knowledge. But the lesson to be learned from Gettier's (1963) counterexamples to the traditional analysis is that there is no unique property that satisfies this description: justification is merely one among many properties that are necessary for a true belief to be knowledge. Hence, the failure of the traditional analysis prompts the need for an alternative account of what sets justification apart from all the other necessary conditions for knowledge.

William Alston argues that debates about the nature of justification threaten to descend into purely terminological disagreements in which various different

epistemologists use the term 'justification' to pick out various different epistemic properties that are necessary for a true belief to be knowledge. The danger, according to Alston, is that "controversies over what it takes for a belief to be justified are no more than a vain beating of the air" (2005: 11). His reaction is to urge that epistemology should broaden its focus from traditional questions about the nature of knowledge and justification to include questions about the nature, importance, and inter-relations among a much wider range of epistemic desiderata.

Alston's insistence on epistemic pluralism is well taken. We should recognize multiple dimensions of epistemic value that play a multiplicity of different roles in our epistemic practices. But this is no reason to ban use of the term 'justification' in epistemology; by parity of reasoning, we would be forced to abandon all our epistemological vocabulary, but this would make further progress impossible to achieve. Instead, we can avoid the threat of terminological disagreement by defining the concept of justification in terms of its distinctive role in our epistemic practices. If we begin by identifying an important role that justification plays in epistemic evaluation, then we can ask what justification must be like in order to play that role. On this approach, what matters is not so much the terminology that we use to pick out an epistemic property, but rather the role that it plays in epistemic evaluation.[1]

The methodology that I am advocating for the theory of justification is an extension of the same methodology that Edward Craig recommends for the theory of knowledge in the passage below:

> We take some prima facie plausible hypothesis about what the concept of knowledge does for us, what its role in our life might be, and then ask what a concept having that role would be like, what conditions would govern its application. (1990: 2)

The general strategy is to begin by considering the point and purpose of using a concept in epistemic evaluation and to use this in constraining a theory of the epistemic property picked out by the concept in question. A constraint of adequacy on a theory of any epistemic property is that it should explain and vindicate the role that our concept of the property plays in epistemic evaluation.

There are several advantages that result from applying this methodology to the theory of justification. First, it casts light on the *importance* of justification. After all, we do evaluate beliefs as justified or unjustified and these evaluations seem to matter. A theory of justification should explain why they matter – that is, why justification is an important dimension of epistemic evaluation. Second, it promises to illuminate the *nature* of justification, since we can ask what justification must be like in order to play its distinctive role in epistemic evaluation. Third, it provides resources for *resolving disagreement* about cases, since we can appeal to the role of justification in adjudicating between conflicting intuitions. And finally, as we noted above, it enables us to avoid *purely terminological debates* about how to use the word 'justification'. What matters is not which terminology we use to pick out an epistemic property, but rather the nature and importance of its role in epistemic evaluation. All of these points will figure in the discussion to follow.

This chapter is guided by the hypothesis that the point and purpose of using the concept of justification in epistemic evaluation is tied to its role in the practice of critical reflection. There is perhaps some irony in the fact that I draw this hypothesis from the work of William Alston, although I develop it in a way that is very much in tension with his own theory of justification. The plan for the paper is as follows. In section one, I use Alston's hypothesis to motivate an analysis of justification as the epistemic property that suits a belief to survive ideal critical reflection. In section two, I use this analysis of justification in arguing for a version of access internalism and against Alston's internalist externalism. In sections three and four, I defend this version of access internalism against objections and I argue that it explains and vindicates some core internalist intuitions about cases. In section five, I conclude with some more general reflections on the debate between internalism and externalism in epistemology.

1. Justification and Critical Reflection

The guiding hypothesis of this paper is that the point and purpose of using the concept of justification in epistemic evaluation is connected with its role in the practice of critical reflection. This idea is central to the Cartesian tradition in epistemology, but it receives perhaps its clearest articulation in the contemporary literature in the following passage from Alston:

> Why is it that we have this concept of *being justified in holding a belief* and why is it important to us? I suggest that the concept was developed, and got

> its hold on us, because of the practice of critical reflection on our beliefs, of challenging their credentials and responding to such challenges – in short, the practice of attempting to *justify* beliefs. (1989: 236)

Critical reflection, as Alston construes it, is the activity that we engage in when we attempt to justify our beliefs by reflecting on what makes them justified. Alston is very careful to distinguish the activity of *justifying* our beliefs in critical reflection from the property of *being justified* which is what we reflect upon when we engage in the activity. Moreover, he insists that the property of being justified does not require engaging in the activity of justifying or even having the psychological capacities needed to do so. Alston's proposal is not that the activity of critical reflection is what makes one's beliefs justified, but rather that the significance of justified belief can be traced to our interest in the activity of critical reflection.

Alston expands on this proposal in the passage below:

> It would be absurd to suggest that in order to be…justified, a belief must actually have been put to the test and emerged victorious. In suggesting that the concept has developed against the background of such a practice the idea is rather that what it is for a belief to be justified is that the belief and its ground be such that it is in a position to pass such a test; that the subject has what it takes to respond successfully to such a challenge. A justified belief is one that *could* survive a critical reflection. (1989: 225)

Alston denies that a justified belief must *actually* survive critical reflection, so long as it has the *potential* to survive critical reflection in the sense that if it were subjected to critical reflection, then it would survive. The proposal, in broad outline, is that justification is the epistemic property in virtue of which a belief has what it takes to survive critical reflection.[2] My aim in the remainder of this section is to develop this proposal and to defend it against objections.

Alston's proposal needs to be qualified in various ways. First, his proposal is vulnerable to the objection that an unjustified belief could survive critical reflection if its basis were to change in the process. To avoid this objection, we can say that a justified belief is one that would survive critical reflection *on its actual basis*. So, for example, a justified belief that is based on perceptual experience could survive on the same basis after critical reflection even if it is now based in a reflective way, rather than an unreflective way.[3]

Second, Alston (1989: 226, n. 45) registers doubts about whether the concept of justification applies to unreflective creatures, such as animals and children. But these doubts can be assuaged since a justified belief has the potential to survive critical reflection in virtue of the basis on which it is held and not in virtue of the subject's reflective capacities. A justified belief is one that would survive on its actual basis if it were subjected to critical reflection by some *idealized counterpart* of the subject with the very same evidence together with the capacity to reflect on it. So, for example, an animal or a child might have a justified belief that could survive critical reflection on its actual basis in an idealized counterpart even if they do not themselves have the capacity to engage in critical reflection.

Third, Alston's proposal is vulnerable to the objection that when critical reflection is performed badly enough, justified beliefs may be abandoned and unjustified beliefs may be retained. To avoid this objection, we need to idealize not only the subject's capacity to engage in critical reflection, but also the way in which this capacity is exercised. In other words, we can say that a justified belief is one that would survive critical reflection that is *sufficiently ideal*. The nature and extent of the relevant idealization is an issue for further discussion below.[4]

Fourth, Alston's proposal is primarily concerned with *doxastic justification*, rather than *propositional justification*, in the sense of Firth (1978). That is to say, Alston is primarily concerned with the conditions for holding a belief in a way that is justified, as opposed to the conditions for having justification for holding a belief independently of whether the belief is held and, if so, on what basis. The crux of the distinction is that doxastic justification requires holding a belief on the right basis, whereas propositional justification does not. With this point in mind, we can extend the proposal from doxastic to propositional justification in the following way:

> A belief is *doxastically justified* if and only if one holds the belief on some basis on which it would be held after ideal critical reflection.

> A belief is *propositionally justified* if and only if one has some basis on which the belief would be held after ideal critical reflection.

In what follows, I will defend this analysis of justification against various objections.

The analysis faces a version of the conditional fallacy objection, since the process of idealization has psychological side-effects that impact which beliefs one has justification to hold. For instance, if I were to engage in critical reflection, then I would have justification to believe that I was doing so, although I don't in fact have justification to believe this, since I'm not engaging in critical reflection right now. Indeed, I have justification to believe that I'm not currently engaged in critical reflection, although of course I wouldn't have justification to believe this if I were engaged in critical reflection. As I will explain, however, we can understand the process of idealization in a way that brackets its psychological side-effects and thereby avoids the conditional fallacy objection.

Critical reflection is an activity that can be applied to its own results. After reflecting on which beliefs one has justification to hold and revising one's beliefs accordingly, one can engage in another iteration of critical reflection on how to revise one's beliefs in light of the critical reflection just undertaken, and so on without end. For the purposes of the analysis, however, we can stipulate that one has justification to believe a proposition if and only if one would believe it after *just one iteration* of critical reflection. This avoids the problematic implication that one has justification to believe whatever one would believe after engaging in critical reflection on the results of critical reflection itself. If the process of idealization is understood in this way, then the conditional fallacy objection can be avoided.

One might push the objection further by arguing that idealization can affect which propositions one has *a priori* justification to believe. For instance, one might argue that, after sufficient idealization, everyone would believe every a priori truth,

although no one actually has justification to believe every a priori truth. What motivates this objection is the assumption that one has a priori justification to believe a proposition only if one has some appropriate mental relation to that proposition, such as understanding an argument or a proof or having an intuition that the proposition is true. On this view, the process of idealization affects which propositions one has a priori justification to believe by affecting one's mental relation to the relevant propositions.

This objection raises fundamental issues about the nature of a priori justification, which I discuss in more detail elsewhere (Smithies, forthcoming). In brief, though, my response is that the process of idealization does not affect which propositions one has a priori justification to believe, but only which a priori justified beliefs one is able to form. Everyone has justification to believe every a priori truth, but it does not follow that everyone is able to form justified beliefs in every a priori truth, since not everyone satisfies the relevant mental conditions. On this view, mental conditions, such as intuition and understanding, are not the source of one's a priori justification, but rather enable one to use a priori justification that one already has in forming a justified belief.

One might protest that this idealized conception of justification loses touch with our ordinary practice of epistemic evaluation. For instance, Frege's belief in Axiom V was justified by ordinary standards, despite being unsuited to survive ideal critical reflection on Russell's set-theoretic paradoxes. In response, though, we can explain our ordinary standards of epistemic evaluation as an approximation towards more ideal standards. For instance, we can give a contextualist account on

which the sentence, 'Frege's belief is justified,' is true when uttered in an ordinary context if and only if it constitutes a sufficient approximation towards the ideal of critical reflection given the human limitations that are salient in that context. In this way, our ordinary standards of epistemic evaluation can be explained by reference to ideal standards that abstract away from our normal human limitations together with further assumptions about the extent of these normal human limitations.[5]

These objections raise more general questions about how to understand the idealization that figures in the analysis. We can idealize an epistemic agent along many different dimensions, but not all of these idealizations are relevant for understanding the notion of justification. For instance, we cannot assume that ideal critical reflection involves omniscience and infallibility about a posteriori truths as well as a priori truths, since this would imply that justification coincides with truth across the board. Ideal critical reflection must be understood in a way that is consistent with the possibility of ignorance and error about empirical matters of fact. This is crucial for capturing a dimension of epistemic evaluation that concerns what one ought to believe given the limitations of one's subjective perspective on an objective world.

As I understand the idealization, ideal critical reflection does not require empirical omniscience or infallibility, but merely requires engaging in the activity of critical reflection in a way that is ideally justified. Critical reflection is a purely *reflective* activity: it is a matter of revising one's beliefs in light of one's reflection on the information that is currently in one's possession. As such, it may be contrasted with more practical or social activities that involve the acquisition of information

through empirical investigation or expert testimony. An ideally justified process of critical reflection need not involve the acquisition of more accurate and complete information about the world, but is rather a matter of reflecting on one's actual information, which may be both inaccurate and incomplete, and revising one's beliefs in light of those reflections. Therefore, ideal critical reflection cannot guarantee empirical omniscience or infallibility; indeed, it is consistent with massive ignorance and error about empirical matters.

Clearly, the analysis contains an element of circularity, since the relevant idealization cannot be explained without using the concept of justification. However, while it follows that the analysis is non-reductive, it does not follow that the analysis is trivial. The general form of the analysis is that one has justification to believe a proposition if and only if one would believe that proposition in *ideal conditions*. If these ideal conditions are simply defined as conditions in which one believes whatever one has justification to believe, then the analysis is trivial. On the current proposal, in contrast, ideal conditions are defined as those in which one brings one's beliefs into alignment with justified higher-order beliefs about which beliefs one has justification to hold. It is a substantive commitment of the proposed analysis that beliefs formed in this way are themselves justified. Moreover, the analysis is illuminating because of the way in which it connects the concept of justification with the activity of critical reflection. In the next section, I go further in arguing that the analysis provides the basis of an argument for access internalism.

2. An Argument for Access Internalism

One branch of the debate between internalism and externalism in epistemology concerns the nature and extent of one's access to epistemic facts about which doxastic attitudes one has justification to hold.[6] Access internalism is the thesis that all the epistemic facts about which beliefs and other doxastic attitudes one has justification to hold are accessible by reflection alone in the following sense:

> **Access Internalism:** one has justification for some doxastic attitude if and only if one has justification to believe on the basis of reflection alone that one has justification for that doxastic attitude.

My aim in this section is to argue for access internalism by appealing to the analysis of justification defended in the previous section.[7]

Once again, I take my inspiration from Alston's work. Alston argues that a belief is justified and so has what it takes to survive critical reflection only if its justifying ground or basis – what he calls a 'justifier' – is accessible to the subject by reflection alone:

> A justified belief is one that *could* survive a critical reflection. But then the justifier must be accessible to the subject. Otherwise the subject would be in no position to cite it as what provides a sufficient indication that the belief is true. (1989: 225)

Nevertheless, Alston rejects access internalism in the sense defined above. Instead, he defends a hybrid view – which he calls *internalist externalism* – on which one must have access to one's justifiers, although one need not have access to the facts in virtue of which they play their justifying role. Alston claims that one's justifiers play their justifying role in virtue of their reliable connections to the external world. On his view, one's justifiers must be accessible, but the facts in virtue of which they justify one's beliefs – namely, their reliable connections to the external world – need not be accessible. As a result, one might have access to one's justifiers without having access to the facts about which beliefs they justify.

My objection to Alston's internalist externalism is that it undermines his claim about the connection between justification and critical reflection. On this view, one's beliefs can be justified without having what it takes to survive ideal critical reflection. Indeed, Alston acknowledges this point in the following passage:

> To illustrate, let's suppose that experiences can function as justifiers, and that they are accessible to us. I can always tell what sensory experiences I am having at a given moment. Even so, if I am unable to tell what belief about the current physical environment is justified by a given sensory episode, I am thereby unable to regulate my perceptual beliefs according as they possess or lack experiential justification. (1989: 221)

Suppose I form a justified belief on the basis of perceptual experience. And suppose my perceptual experience is accessible, but not the fact that it justifies my belief. In

that case, my belief does not have the potential to survive ideal critical reflection. After all, the aim of critical reflection is to bring my beliefs into line with justified reflections about which beliefs I have justification to hold. If don't have justification to believe on the basis of reflection alone that I have justification for my belief, then my belief doesn't cannot withstand ideal critical reflection. So, my belief has what it takes to survive ideal critical reflection only if the fact that I have justification for the belief is accessible to me on the basis of reflection alone.

The upshot of this critique of Alston is that access internalism can be motivated as a consequence of the analysis of justification defended in section one. Here is the argument for access internalism:

(1) One has justification to believe that *p* if and only if one has some basis on which one would believe that *p* after ideal critical reflection.
(2) One has some basis on which one would believe that *p* after ideal critical reflection if and only if one has justification to believe on the basis of reflection alone that one has justification to believe that *p*.
(3) Therefore, one has justification to believe that *p* if and only if one has justification to believe on the basis of reflection alone that one has justification to believe that *p*.

Both premises of this argument were defended in section one. The first premise restates the analysis of justification in terms of ideal critical reflection, while the second premise articulates the way in which the idealization is to be understood.

Critical reflection is a matter of reflecting on which beliefs one has justification to hold and revising one's beliefs accordingly. The aim of the activity is to bring one's beliefs into alignment with one's justified reflections about which beliefs one has justification to hold. Therefore, ideal critical reflection is a matter of believing a proposition if and only if one has justification to believe on the basis of reflection alone that one has justification to believe that proposition. Access internalism follows given the connection between justification and ideal critical reflection.

Opponents of access internalism must reject one or both of these premises. In doing so, however, they must be careful to avoid the risk of purely terminological disagreement. Merely replacing the proposed analysis of justification with an alternative threatens simply to change the subject by using the word 'justification' to pick out a different epistemic property. Access internalism is consistent with a form of epistemic pluralism on which there are many important epistemic properties that play a range of different roles in epistemic evaluation. It is not committed to the accessibility of all of these epistemic properties, including knowledge, reliability, and truth. It is committed only to the existence of one epistemic property that is accessible for which it reserves the term 'justification'. Anyone who accepts the existence of such a property, but uses the term 'justification' to pick out a different property, disagrees with access internalism on purely terminological grounds. Therefore, those who oppose access internalism on substantive grounds must argue that no important epistemic property is accessible in the relevant sense.[8]

This completes my argument for access internalism. If access internalism is true, a theory of justification should explain *why* it is true – that is, it should give an

account of the determinants of justification that explains why justification is accessible. This is a task that I take up in other work.[9]

3. Defending Access Internalism

Access internalism, as I have defined it, is a thesis about propositional justification, rather than doxastic justification: it is not a thesis about which of one's beliefs are justified, but about which propositions one has justification to believe.[10] Justified belief requires not only *having* justification to believe a proposition, but also *using* it – that is, believing the proposition on the basis of one's justification to believe it. Propositional justification is therefore necessary but not sufficient for doxastic justification, since one may have justification to believe a proposition without using it and, as I argue below, without even having any capacity to use it. As I will explain, the distinction between propositional and doxastic versions of access internalism is crucial for avoiding some familiar objections.

The propositional version of access internalism states that one has justification to believe that *p* if and only if one has higher-order justification to believe that one has justification to believe that *p*. The doxastic version, by contrast, states that one's belief that *p* is justified if and only if it is based on a justified higher-order belief that one has justification to believe that *p*. For instance, one of the key premises of Laurence BonJour's (1985: Ch.2) argument against foundationalism is that a belief B is justified if and only if it is held on the basis of a meta-justificatory argument of the following form:

(1) B has feature φ.

(2) Beliefs having feature φ are highly likely to be true.

(3) Therefore, B is highly likely to be true.

According to BonJour, "it is necessary, not merely that a justification along the above lines exist in the abstract, but also that [the subject] himself be in cognitive possession of that justification, that is, that he believe the appropriate premises of forms (1) and (2) and that these beliefs be justified *for him*" (1985: 31).

The doxastic version of access internalism is vulnerable to the charge of over-intellectualization, since it is plausible that some human infants and non-human animals can form justified beliefs, although they do not have the conceptual or reflective abilities to form justified beliefs about what they have justification to believe. Moreover, the doxastic version of access internalism faces a regress problem, since one's first-order beliefs are justified only if they are based on justified second-order beliefs, but these second-order beliefs are justified only if they are based on justified third-order beliefs, and so on ad infinitum. Thus, one has justified beliefs only if one has an infinite hierarchy of increasingly complicated higher-order justified beliefs. But no finite creature can have this kind of infinite hierarchy of increasingly complicated higher-order justified beliefs, so this generates the skeptical conclusion that no finite creature has justified beliefs.[11]

The charge of over-intellectualization does not apply to the propositional version of access internalism, since it does not claim that one's beliefs are justified only if they are based on justified higher-order beliefs. It does generate an infinite

regress of a kind, since it implies that one has first-order justification to believe a proposition only if one has second-order justification to believe that one has first-order justification, and one has second-order justification only if one has third-order justification to believe that one has second-order justification, and so on ad infinitum. However, it does not imply that one must believe any of these higher-order propositions in order to have justified beliefs. As a result, the infinite regress is not vicious, but benign.

It might be objected that while having justification to believe a proposition does not require using it in forming a justified belief, it does at least require that one has the *capacity* to do so. Call this *the capacity principle*:

> **The Capacity Principle:** if one has propositional justification to believe that *p*, then one has the capacity to form a doxastically justified belief that *p*.

If the capacity principle is true, then the problems of over-intellectualization and vicious regress arise for propositional as well as doxastic versions of access internalism. After all, human infants and non-human animals have justified beliefs about the external world, but they lack the capacity to form justified higher-order beliefs about what they have justification to believe. Likewise, normal human adults have justified beliefs, but they lack the capacity to form an infinite hierarchy of higher-order justified beliefs of ever-increasing complexity. In my view, however, there are no compelling reasons to accept the capacity principle; in fact, there are some compelling reasons to reject it.

Many philosophers have found it plausible to suppose that there is an analytic connection between propositional and doxastic justification. For instance, Alvin Goldman proposes the following connection:

> S is *ex ante* [i.e. propositionally] justified in believing that *p* at *t* just in case his total cognitive state at *t* is such that from that state he could come to believe *p* in such a way that this belief would be *ex post* [i.e. doxastically] justified. (1979: 21)

But this analytic connection between propositional and doxastic justification need not be understood in a way that accords the capacity principle with the status of an analytic truth. The general form of the proposal is that one has propositional justification to believe that *p* just in case one would have a doxastically justified belief that *p* in certain conditions. But it cannot be assumed without further argument that these are conditions that one must have the capacity to bring about. For instance, I have argued that one has propositional justification to believe that *p* just in case one would have a doxastically justified belief that *p* after a sufficiently idealized process of critical reflection, but there is certainly no presumption here that the idealization must be constrained by one's actual psychological capacities. The capacity principle is therefore a substantive claim that needs further argument and cannot be regarded merely as a trivial consequence of the analytic connection between propositional and doxastic justification.

An influential line of argument for the capacity principle appeals to a deontological conception of justification combined with an ought-implies-can principle.[12] On this view, epistemic justification is a source of epistemic obligations, which are binding only insofar as one has the psychological capacities required to discharge them. The argument proceeds roughly as follows:

(1) If one has propositional justification to believe that *p*, then one ought to form a doxastically justified belief that *p*.
(2) If one ought to form a doxastically justified belief that *p*, then one can form a doxastically justified belief that *p*.
(3) So, if one has propositional justification to believe that *p*, then one can form a doxastically justified belief that *p*.

However, the conclusion of the argument is subject to intuitive counterexamples in which one is incapable of believing what one what one has justification to believe owing to the corrupting influence of drugs, brainwashing or mental illness.[13] In such cases, one cannot legitimately be blamed for failing to believe what one has justification to believe, since one's limited capacities provide an excuse. But while it is plausible that blameworthiness is constrained by one's psychological limitations, it does not follow that there are corresponding limitations on which propositions one has justification to believe.

If the conclusion is false, then which of the premises should be rejected? Alston (1989: Chs. 4 & 5) rejects premise (1) by abandoning the deontological

conception of justification in favor of an evaluative conception on which justification is a source of epistemic values, rather than epistemic obligations or permissions. In contrast, Feldman (2000) accepts the deontological conception of justification as stated in premise (1), but rejects the 'ought' implies 'can' principle as stated in premise (2). Feldman argues that there are so-called 'role oughts' that apply to anyone who plays a certain role, regardless of how well they are able to play that role – for instance, chefs ought to make delicious food and jugglers ought to keep their balls in the air. Similarly, Feldman argues, there are epistemic 'oughts' that apply to us in virtue of our role as believers: "It is our plight to be believers. We ought to do it right. It doesn't matter that in some cases were are unable to do so" (2000: 676).

Deciding between these two options depends on how we understand the relationship between values and obligations. Are we obliged to achieve evaluative ideals or merely to approximate towards them as closely as we can? Perhaps there is a 'thin' sense in which we are obliged to achieve ideals regardless of whether we are capable of doing so. Thus, Feldman and Conee write, "In any case of a standard for conduct…it is appropriate to speak of 'requirements' or 'obligations' that the standard imposes" (1985: 19). But there seems also to be a 'thick' sense, which is more closely connected with reactive attitudes of praise and blame, in which we are obliged merely to approximate towards ideals to the extent that we are capable of doing so. Obligations in this thicker sense can be reconstructed from epistemic ideals together with further assumptions about our limited capacities.[14]

The argument equivocates between thick and think senses of 'ought': in the thin sense, premise (1) is true, but premise (2) is false, whereas in the thick sense, premise (1) is false, but premise (2) is true. For this reason, I suspect that there is no deep disagreement between Alston and Feldman in their diagnosis of what's wrong with the argument aside from their differing interpretation of the relevant notion of 'ought'. The key point is that however we interpret it, we can block the argument that all evaluative ideals must be humanly attainable.

Moreover, there is good reason to suppose that many of the evaluative ideals that we care about may be humanly unattainable. After all, we may be interested in evaluating the performance of human beings along some dimension whose extremes lie beyond human reach. For this reason, epistemic ideals – like ideals of morality, scientific understanding, and chess – may lie beyond our limited human capacities. The limits of human capacities need not constrain our understanding of epistemic ideals themselves, but only the extent to which we are capable of approximating towards the ideal. As David Christensen puts the point, "Not all evaluation need be circumscribed by the abilities of the evaluated. In epistemology, as in various other arenas, we need not grade on effort" (2004: 162).

4. Intuitions about Cases

Michael Bergmann (2006: Ch.1) argues that access internalism is motivated by intuitions about cases that cannot be explained except in a way that generates a vicious kind of infinite regress. Thus, he argues, access internalism is faced with a dilemma: either it is viciously regressive or it is unmotivated. In response to

Bergmann, I argue that the version of access internalism defended in this paper is strong enough to explain the relevant intuitions but also weak enough to avoid the vicious regress. Therefore, I conclude that Bergmann's dilemma fails.

Bergmann claims that access internalism is motivated primarily by intuitions about cases, such as Laurence BonJour's (1985) case of the clairvoyant, Norman:

> Norman, under certain conditions which usually obtain, is a completely reliable clairvoyant with respect to certain kinds of subject matter. He possesses no evidence or reasons of any kind for or against the general possibility of such a cognitive power or for or against the thesis that he possesses it. One day Norman comes to believe that the President is in New York City, though he has no evidence either for or against this belief. In fact the belief is true and results from his clairvoyant power under circumstances in which it is completely reliable. (1985: 41)

BonJour's intuition is that Norman's belief is unjustified, despite the fact that it is formed on the basis of a clairvoyant power that is reliable in the circumstances. This intuition is widely shared, but not everyone is persuaded. In fact, many of BonJour's opponents simply bite the bullet and insist that Norman's belief is justified. As Bergmann reconstructs the dialectic, however, BonJour confronts these opponents with the following *subject's perspective objection*:

> **The Subject's Perspective Objection:** If the subject holding a belief isn't aware of what that belief has going for it, then she isn't aware of how its status is any different from a stray hunch or an arbitrary conviction. From that we may conclude that from her perspective it is an accident that her belief is true. And that implies that it isn't a justified belief. (2006: 12)

As Bergmann observes, this objection applies to any theory of justification that does not impose the following *strong awareness requirement* on justification:

> **The Strong Awareness Requirement:** S's belief B is justified only if (i) there is something, X, that contributes to the justification of B…and (ii) S is aware (or potentially aware) of X [as contributing to the justification of B]. (2006: 9)

But Bergmann argues that the strong awareness requirement generates a vicious regress. Consider a doxastic version of the strong awareness requirement on which S's belief B is justified only if S has (or potentially has) a justified belief that X contributes to the justification of B.[15] This implies that S's belief B1 is justified only if S has (or potentially has) a justified belief B2 that X1 contributes to the justification of B1, and S has a justified belief B2 only if S has (or potentially has) a justified belief B3 that X2 contributes to the justification of B2, and so on ad infinitum. And yet no finite subject has (or potentially has) an infinite series of increasingly complicated higher-order beliefs. Bergmann concludes that access internalism is faced with a

dilemma: if access internalism imposes the strong awareness requirement, then it is viciously regressive, but if not, then it is unmotivated.

My response to Bergmann's dilemma trades on the distinction between propositional and doxastic versions of access internalism. In the previous section, I argued that the propositional version of access internalism, unlike the doxastic version, generates an infinite regress that is benign, rather than vicious. Hence, the first horn of the dilemma can be avoided. But what about the second horn? Bergmann claims that any version of access internalism weak enough to avoid the regress problem is too weak to explain internalist intuitions and is therefore unmotivated. My response has two parts. First, as I explain below, the propositional version of access internalism is strong enough to explain and vindicate the relevant internalist intuitions about cases. And second, I deny that this version of access internalism is motivated solely by intuitions about cases. Indeed, one of the central goals of this paper has been to motivate access internalism by appealing to more general theoretical considerations about the connection between justification and critical reflection. Taken in tandem, these points have important methodological implications for the role of intuitions in epistemology. On the internalist theory of justification proposed in this paper, it is not simply a brute deliverance of intuition that Norman's clairvoyant beliefs are unjustified. Rather, this is the conclusion of an argument whose premises are independently motivated by theoretical considerations about the connection between justification and critical reflection.

In the following passage, BonJour appeals to the connection between justification and critical reflection in explaining why Norman's belief is unjustified:

> Norman's acceptance of the belief about the President's whereabouts is epistemically irrational and irresponsible, and thereby unjustified, whether or not he believes himself to have clairvoyant power, so long as he has no justification for such a belief. Part of one's epistemic duty is to reflect critically upon one's beliefs, and such critical reflection precludes believing things to which one has, to one's knowledge, no reliable means of epistemic access. (1985: 42)

According to BonJour, Norman's belief is unjustified because he does not fulfill his epistemic duty to subject his belief to critical reflection and he is therefore guilty of epistemic irresponsibility. The suggestion here is that a belief is justified only if it is held on the basis of critical reflection in a way that is epistemically responsible. As Bergmann (2005: 430) comments, "It seems that he [i.e. BonJour] thinks that Norman must take up some doxastic attitude towards [the proposition that he has justification to believe that *p*] in order for his belief that *p* to be justified." BonJour's reasoning can be made explicit in the form of an argument that relies on a doxastic version of access internalism:

(1) One's belief that *p* is justified if and only if it is based on a justified higher-order belief that one has justification to believe that *p*.

(2) Norman's belief that the President is in New York City is not based on a justified higher-order belief that he has justification to believe that the President is in New York City.

(3) Therefore, Norman's belief that the President is in New York City is unjustified.

The problem with this argument is that the doxastic version of access internalism is vulnerable to the charges of over-intellectualization and vicious regress.

Like BonJour, I propose to explain why Norman's belief is unjustified by appealing to the connection between justification and critical reflection. However, I understand this connection rather differently from BonJour, since I endorse a propositional version of access internalism, rather than a doxastic version. And so I want to propose a different argument that Norman's clairvoyant belief is unjustified that avoids the objections to BonJour's argument:

(1) One has justification to believe that *p* if and only if one has higher-order justification to believe that one has justification to believe that *p*.

(2) Norman does not have higher-order justification to believe that he has justification to believe that the President is in New York City.

(3) Therefore, Norman does not have justification to believe that the President is in New York City.

The crucial issue, in my view, is not whether Norman subjects his clairvoyant belief to critical reflection or even whether he has the capacity to do so. Rather, the crucial issue is whether his belief has what it takes to survive an idealized process of critical reflection. Whether or not he reflects on the matter, Norman's belief does not have the potential to survive ideal critical reflection, since he does not have higher-order justification to believe that he has justification to believe that the President is in New York City. Even reliabilist theories of justification must grant this premise, since Norman does not have a reliable way of establishing his own reliability. Therefore, it follows that Norman lacks justification to believe that the President is in New York City and hence his belief is unjustified.

Let me conclude this section with some brief methodological reflections on the role of intuitions about cases in a theory of justification. In my view, a theory of justification should not merely conform with intuitions about cases; it should also vindicate those intuitions by explaining how they track a property that is suited to play an important role in epistemic evaluation. To the extent that this cannot be done, intuitions about cases should be revised in light of more general theoretical considerations about what the property would have to be like in order to play an important role in epistemic evaluation. That is, a theory of justification should aim for reflective equilibrium between intuition and theory, but theory should take priority over intuition, rather than vice versa. As I will explain, this methodology has implications both for resolving substantive disagreements about cases and for avoiding purely terminological disagreements.

The practice of relying on intuitions about cases as data for an epistemological theory has been the target of criticism by recent work in experimental philosophy. For instance, Weinberg, Nichols and Stich (2001) present experimental evidence for cultural and socio-economic variation in epistemic intuitions. These results raise challenging questions about which epistemic intuitions, if any, should carry evidential weight in cases of disagreement. As they pose the challenge, "Why should we privilege our intuitions rather than the intuitions of some other group?" (2001: 435). Certainly, we should not privilege our intuitions just because they are ours! On the other hand, we need not suppose that all intuitions are equal. On the contrary, intuitions about cases can and should be vindicated in light of theoretical considerations about the role of a property in epistemic evaluation. Indeed, my main aim in this section has been to argue that BonJour's intuitions about clairvoyance can be explained and vindicated by the version of access internalism that I motivated in earlier sections of the paper by appealing to the connection between justification and critical reflection. Therefore, epistemological disagreements need not ground out in conflicting intuitions about cases, but can be made more theoretically tractable by reflecting upon which epistemic properties are suited to play which epistemic roles.

One possibility, of course, is that there is only a superficial disagreement, since different parties may use the term 'justification' to pick out different epistemic properties. In that case, there is no substantial disagreement about which epistemic properties are instantiated in a given case, but merely a terminological disagreement about which of these properties deserves the name 'justification'.

Indeed, Alston suspects that debates between internalists and externalists about BonJour's clairvoyance cases may be purely terminological debates of exactly this kind. Thus, he writes:

> Norman exhibits one epistemically important desideratum – a belief formed in a reliable way – and lacks another. We can then discuss what the further implications are of the possession or lack of each of these desiderata. (2005: 55)

Alston's methodological recommendations are highly congenial to the approach that I have adopted in this paper. We can avoid terminological disputes about whether the word 'justification' applies in a particular case by asking instead whether the case instantiates properties that are suited to play an important and distinctive role in epistemic evaluation. As far as I am concerned, what matters is not which terminology we use to pick out an epistemic property, but rather the importance of its role in epistemic evaluation.

5. Why Critical Reflection Matters

My main aim in this paper has been to argue that justification is an important dimension of epistemic evaluation because it corresponds to an ideal for the activity of critical reflection. Nevertheless, the argument so far leaves a crucial question unanswered. Why should we suppose that critical reflection is an important activity

and hence that there is an important dimension of epistemic evaluation that corresponds to ideal performance in that activity?

I do not claim that critical reflection is an important activity because it is guaranteed to make us more reliable. When done well, it can increase our overall reliability by weeding out logical fallacies, hasty generalizations, baseless prejudice, and wishful thinking. When done poorly, though, it can make us less reliable by rationalizing the results of these irrational habits of belief-formation. Even when critical reflection is done perfectly, it is not guaranteed to make us more reliable, since we can imagine circumstances – in the style of BonJour's clairvoyance cases – in which we increase our overall reliability by engaging in irrational habits of belief-formation. Nevertheless, I maintain that that critical reflection is an important and valuable activity even if it is not guaranteed to make us more reliable.[16]

Critical reflection is a component of good reasoning, but I do not claim that critical reflection is a precondition for engaging in reasoning at all. Some human infants and non-human animals engage in reasoning without having any capacity to engage in critical reflection. Indeed, mature humans are unique in having the ability to engage in critical reflection, but not all of our reasoning is carried out under the guidance of critical reflection. Our beliefs are often formed, revised and maintained in a way that is highly automatic and unreflective.

Following Tyler Burge (1996, 1998), I claim that critical reflection is an important form of reasoning because it is the distinguishing mark of being a person. Persons are distinguished from other animals by the fact that they are responsible for their beliefs and choices. We do not regard it as appropriate to adopt reactive

attitudes, such as praise and blame, towards human infants and non-human animals because they are not responsible for their beliefs and choices.[17] Responsibility requires reflective control over one's beliefs and choices in the sense that one can revise them in light of critical reflection. Therefore, persons are distinguished from other animals by the fact that they are responsible for their beliefs and choices in virtue of having the capacity to engage in critical reflection.[18]

Many internalist theories of justification emphasize the connection between justification and responsibility. Thus, BonJour writes, "The idea of avoiding … irresponsibility, of being epistemically responsible in one's believings, is the core of the notion of epistemic justification" (1985: 8). In my view, however, proponents of internalism are mistaken in claiming that a belief is justified if and only if it is held in a way that is responsible. As proponents of externalism rightly protest, this is an over-intellectualization, since responsibility requires a capacity for critical reflection, whereas justified belief does not. On the other hand, proponents of externalism are mistaken in claiming that we can allow for unreflectively justified belief only by severing the connection between justification, responsibility, and critical reflection altogether. A related mistake is to bifurcate the concept of justification into an internalist kind that preserves the connection with responsible critical reflection and an externalist kind that does not.[19]

The internalist theory of justification developed in this paper provides an alternative to the influential view that externalism is mandated by the need to avoid over-intellectualization in epistemology. On the one hand, the internalist mistake can be avoided, since we need not claim that all justified beliefs are actually held on

the basis of responsible critical reflection. On the other hand, the externalist mistake can be avoided, since we can insist that all justified beliefs have the potential to be held on the basis of responsible critical reflection. Therefore, we need not follow Burge in supposing that "epistemic externalism … constitutes an antidote to hyper-intellectualization" (2003: 504).

The debate between internalism and externalism about justification tends to oscillate between two extremes. On the one hand, internalist theories tend to over-intellectualize the requirements for justification in such a way as to rule out the possibility of unreflectively justified belief. On the other hand, externalist theories tend to allow for the possibility of unreflectively justified belief by severing the connection between justification and responsibility altogether. This dialectical situation is exacerbated by the fact that the connection between justification and responsibility disappears from view when we restrict our attention to cases of unreflectively justified belief. After all, the significance of the concept of justified belief emerges only in the context of its role in critical reflection. In order to reach a satisfactory resolution of the debate between internalism and externalism, we need to make sense of the possibility of unreflective justification without losing sight of the role of justification in the practice of critical reflection. My proposal is designed to occupy this elusive middle ground.[20]

References

Alston, William (1989) *Epistemic Justification: Essays in the Theory of Knowledge*, Cornell University Press.

Alston, William (2005) *Beyond 'Justification': Dimensions of Epistemic Evaluation*, Cornell University Press.

Audi, Robert (1988) "Justification, Truth, and Reliability" *Philosophy and Phenomenological Research* 49 (1): 1-29.

Bergmann, Michael (2005) "Defeaters and Higher Level Requirements" *The Philosophical Quarterly* 55: 419-436.

Bergmann, Michael (2006) *Justification Without Awareness: A Defense of Epistemic Externalism*, Oxford University Press.

BonJour, Laurence (1985) *The Structure of Empirical Knowledge*, Harvard University Press.

Burge, Tyler (1996) "Our Entitlement to Self-Knowledge" *Proceedings of the Aristotelian Society* 96: 91-116.

Burge, Tyler (1998) "Reason and the First Person" in (eds.) C. Wright, B. Smith and C. Macdonald, *Knowing Our Own Minds*, Oxford University Press.

Burge, Tyler (2003) "Perceptual Entitlement" *Philosophy and Phenomenological Research* 68 (3): 503-48.

Chalmers, David (2011) "Verbal Disputes and Philosophical Progress" *Philosophical Review* 120 (4): 515-66.

Christensen, David (2004) *Putting Logic In Its Place: Formal Constraints on Rational Belief*, Oxford University Press.

Craig, Edward (1990) *Knowledge and the State of Nature: An Essay in Conceptual Synthesis*, Oxford University Press.

DePaul, Michael (2001) "Value Monism in Epistemology" in (ed.) M. Steup, *Knowledge, Truth and Duty*, Oxford University Press.

Feldman, Richard and Conee, Earl (1985) "Evidentialism" *Philosophical Studies* 48 (1): 15-34.

Feldman, Richard (2000) "The Ethics of Belief" *Philosophy and Phenomenological Research* 60 (3): 667-695.

Firth, Roderick (1978) "Are Epistemic Concepts Reducible to Ethical Concepts?" in (eds.) A. Goldman and J. Kim, *Values and Morals*, Dordrecht: Kluwer.

Foley, Richard (1993) *Working Without a Net: A Study of Egocentric Epistemology*, Oxford University Press.

Gettier, Edmund (1963) Is Justified True Belief Knowledge? *Analysis* 23 (6): 121-3.

Ginet, Carl (1975) *Knowledge, Perception and Memory*, Dordrecht: Reidel.

Goldman, Alvin (1979) "What is Justified Belief?" in (ed.) G. Pappas, *Justification and Knowledge*, Dordrecht: Reidel.

Goldman, Alvin (1999) "Internalism Exposed" *Journal of Philosophy* 96 (6): 271-93.

Greco, John (1990) "Internalism and Epistemically Responsible Belief" *Synthese* 85 (2): 245-77.

Greco, John (2005) "Justification is Not Internal" in eds. M. Steup and E. Sosa, *Contemporary Debates in Epistemology*, Oxford: Blackwell.

Locke, John (1968) *An Essay Concerning Human Understanding*, London: Dent.

Plantinga, Alvin (1993) *Warrant: The Current Debate*, Oxford University Press.

Pryor, James (2001) "Highlights of Recent Epistemology" *British Journal for the Philosophy of Science* 52: 95-124.

Schaffer, Jonathan (2010) "The Debasing Demon" *Analysis* 70 (2): 228-37.

Smithies, Declan (2012a) "A Simple Theory of Introspection" in eds. D. Smithies and D. Stoljar, *Introspection and Consciousness*, Oxford University Press.

Smithies, Declan (2012b) "Moore's Paradox and the Accessibility of Justification" *Philosophy and Phenomenological Research* 85 (2): 273-300.

Smithies, Declan (2012c) "Mentalism and Epistemic Transparency" *Australasian Journal of Philosophy* 90 (4): 723-41.

Smithies, Declan (Forthcoming) "The Role of Experience in A Priori Justification" *Synthese*.

Sosa, Ernest (1991) *Knowledge in Perspective: Selected Essays in Epistemology*, Cambridge University Press.

Strawson, Peter (1962) "Freedom and Resentment" *Proceedings of the British Academy* 48: 1-25.

Weinberg, Jonathan, Nichols, Shaun, and Stich, Stephen (2001) "Normativity and Epistemic Intuitions" *Philosophical Topics* 29 (1-2): 429-60.

Williamson, Timothy (2000) *Knowledge and Its Limits*, Oxford University Press.

[1] David Chalmers proposes a similar strategy for avoiding purely terminological disagreement: "On the picture I favor, instead of asking, 'What is X', one should focus on the roles one wants X to play, and see what can play that role" (2011: 538).

[2] Compare Audi's process-property integration thesis: "a belief is justified (has the property of justifiedness) if and only if it has one or more other, non-normative properties such that (i) in virtue

of them it is justified, and (ii) citing them can, at least in principle, both show that it is justified and (conceptually) *constitute* justifying it" (1988: 6). Thanks to Peter Graham for alerting me to the connections between Alston's and Audi's work on these issues.

[3] Here I am distinguishing between the basis on which a belief is based and the way in which it is based, which may be either reflective or unreflective. Thanks to Brian McLean for discussion.

[4] Foley (1993) gives an analysis of egocentric rationality as invulnerability to self-criticism by one's own deepest epistemic standards. However, Foley's analysis does not allow for idealization in one's epistemic standards, or in one's ability to apply them, but only in the conditions in which they are applied. Therefore, Foley counts some dogmatic and delusional beliefs as rational, whereas I count them as unjustified because they would not survive appropriate idealization in one's capacity for critical reflection.

[5] Compare Christensen's (2004: Ch. 6) discussion of the epistemic ideal of probabilistic coherence and its connection with our ordinary practices of epistemic evaluation. See Smithies (forthcoming) for a more detailed discussion of the nature of epistemic idealization.

[6] See Goldberg (this volume) for a different perspective on the debate between internalism and externalism. Strictly speaking, there is not just one debate, but a cluster of related debates that are often grouped together under the same rubric. For instance, Alston (1989: Ch.8) distinguishes "access internalism" from "perspectival internalism". In Smithies (2012a), I argue that these forms of internalism stand or fall together, but that is beyond the scope of the present chapter.

[7] Smithies (2012b) argues for access internalism on the grounds that it is indispensable for solving an epistemic version of Moore's paradox – that is, explaining the irrationality of believing Moorean conjunctions of the form, '*p* and I don't have justification to believe that *p*.'

[8] Williamson (2000: Ch.4), Greco (1990, 2005) and Bergmann (2006: Ch.1) argue that no important epistemic property is accessible in the relevant sense. I respond to Williamson's argument in Smithies (2012c) and to Greco's and Bergmann's arguments in sections three and four below.

[9] Smithies (2012a) argues that access internalism is explained by a form of perspectival internalism on which the facts about which propositions one has justification to believe are determined by facts about one's introspectively accessible mental states.

[10] One reason for this is that the basing relation is not accessible. Schaffer's (2010) debasing demon could make one's beliefs unjustified by undetectably changing the basis on which they are held without thereby undermining one's justification to believe that they are justified.

[11] The regress problem is developed in various different ways by Alston (1989: Ch.8), Greco (1990) and Bergmann (2006: Ch.1). Bergmann's argument is discussed in section four below.

[12] Alston (1989: Ch.5) considers a closely related line of argument that the deontological conception of justification implies that one has voluntary control over one's beliefs. It is worth noting explicitly that my argument for access internalism in section two does not rely on a deontological conception of justification, unlike those proposed by Ginet (1975) and criticized by Alston (1989: Ch.8), Plantinga (1993: Ch.1), Goldman (1999), and Bergmann (2006: Ch.4).

[13] Examples of this kind are discussed by Feldman and Conee (1985: 17), Alston (1989: 95-6), Pryor (2001: 114-5) and Christensen (2004: 161-2).

[14] See Pryor (2001: 115, fn. 36) for a related distinction between thick and thin senses of 'obligation'.

[15] Bergmann (2006: 14-19) also considers a nondoxastic version of the strong awareness requirement and argues that it too generates a vicious regress.

[16] Here, I am rejecting what DePaul (2001) calls "value monism" in epistemology, which is the view that truth is the only epistemic good. I doubt that the value and importance of justification can be explained non-reductively in terms of its reliability or truth-conduciveness.

[17] Strawson (1962) is the classic discussion of the reactive attitudes and their connection with the concepts of personhood and responsibility.

[18] On Locke's (1689: II xxvii 9) famous definition, being a person requires not only reason and consciousness, but also reflection and self-consciousness; thus, a person is "a thinking intelligent being that has reason and reflection and can think of itself as itself, a thinking intelligent thing, in different times and places."

[19] Here I have in mind Sosa's (1991: 145) distinction between *animal knowledge*, which is "apt" in the sense that it derives from a reliable intellectual virtue or disposition, and *reflective knowledge*, which is "justified" in the sense that it comprises part of a coherent higher-order perspective on one's beliefs. See also Burge's (1996: 94) distinction between *justification* and *entitlement* on which the former is distinguished from the latter on the grounds that the subject must conceptualize it.

[20] This paper was presented at the Chambers Conference at the University of Nebraska, Lincoln in September 2010 and more recently at Aberdeen and St Andrews in June 2013. I am grateful to audiences on those occasions and also to David Chalmers, Stewart Cohen, Trent Dougherty, Jeremy Fantl, Sandy Goldberg, David Henderson, Brent Madison, Brian McLean, Nicholas Silins, Sigrun Svavarsdottir, Daniel Stoljar and an anonymous referee for written comments and discussion.