

Why Justification Matters

Declan Smithies

Forthcoming in *Epistemic Evaluation: Point and Purpose in Epistemology*,
edited by J. Greco and D. Henderson. Oxford University Press, 2015.

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, since 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.

Knowledge is traditionally analyzed as justified true belief. On this analysis, justification is the property that turns true belief into knowledge. But one 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 (2005) argues that debates about the nature of justification threaten to descend into purely terminological disagreements in which different epistemologists use the term 'justification' to pick out 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 emphasis 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 we need not follow his recommendation to eliminate use of the term 'justification' in epistemology. Instead, we can avoid the threat of purely terminological disagreement by defining the concept of justification in terms of its 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 application 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.²

There are several advantages that result from applying this methodology to the theory of justification. First, it promises to illuminate the *importance* of justification. After all, we do evaluate beliefs as justified or unjustified and these

¹ Compare Chalmers's advice for avoiding verbal disagreement: "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 Haslanger (1999) proposes a closely related methodology for a theory of knowledge aimed at incorporating insights from feminist theory.

evaluations matter to us. 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 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 the practice of critical reflection. There is perhaps some irony in the fact that I draw this hypothesis from the work of William Alston, since 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 makes a belief stable under 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 section three, I defend this version of access internalism against regress and over-intellectualization objections. In section four, I argue that it explains and justifies some internalist intuitions about cases. In section five, I explain why justification is an important dimension of epistemic evaluation. I conclude with some general reflections on the current state of the debate between internalism and externalism in epistemology.

1. Justification and Reflection

The guiding hypothesis of this paper is that the point and purpose of using the concept of justification in epistemic evaluation is tied to the practice of critical reflection. This idea is characteristic of the Cartesian tradition in epistemology, but it finds perhaps its clearest articulation in Alston’s work. He writes:

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 careful to distinguish the activity of *justifying* our beliefs from the property of *being justified*, which is what we reflect upon when we engage in the activity. Moreover, having justified beliefs requires neither engaging in the activity of justifying those beliefs through critical reflection nor having the psychological capacity to do so. Alston's proposal is not that critical reflection is what makes one's beliefs justified, but rather that the significance of justified belief derives from its connection with critical reflection.

To add more detail, the proposal is that the standards for justified belief are defined by reference to the activity of critical reflection. Alston writes:

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)

To a first approximation, Alston's proposal is that a justified belief is one that is *stable under reflection* in the sense that if it were subjected to critical reflection, then it would survive. Alston denies that a justified belief must actually survive the test of reflective scrutiny, so long as it has the potential to survive the test. Justification on

this view is the epistemic property in virtue of which a belief has the potential to survive critical reflection.³

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 instance, a justified belief that is held on the basis of perceptual experience could survive on the same or a sufficiently similar basis after critical reflection although it is now held in a reflective way, rather than an unreflective way.

Second, Alston (1989: 226, n. 45) registers doubts about whether the concept of justification extends 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. Animals and children can have justified beliefs in the absence of any capacity to engage in critical reflection, so long as their beliefs could survive on their actual basis in some idealized counterpart that has the relevant capacities.

Third, Alston's proposal is vulnerable to the objection that when critical reflection is done badly enough, justified beliefs can be abandoned and unjustified beliefs can be retained.⁴ 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 is stable under reflection that is ideally rational, reasonable, or justified.

³ As I learned from Peter Graham, a related idea can be found in Audi's work. Compare his *process-property integration thesis*: "a belief is justified...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).

⁴ See Kornblith (2012) for this objection and an overview of relevant empirical evidence bearing on the unreliability of human reflection.

Finally, Alston's proposal concerns what Roderick Firth (1978) called *doxastic justification* as distinct from *propositional justification*: it is a thesis about the conditions under which one holds a belief in a way that is justified, rather than the conditions under which one has justification for holding a belief regardless of whether one does in fact hold the belief and, if so, on what basis. Having drawn this distinction, we can extend the proposal 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.

The crux of the distinction is that doxastic justification requires that one holds a belief on the right kind of basis, whereas propositional justification does not.

As it stands, the proposal commits a version of the conditional fallacy, since the process of idealization has psychological side effects that impact which beliefs one has justification to hold. If I were to engage in critical reflection, then I would have justification to believe that I was doing so, but 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 engaging in critical reflection, although of course I wouldn't have justification to believe this if I were engaging in critical reflection. In order to avoid this conditional fallacy objection, we need to understand the idealization in a way that brackets its psychological side effects.

Following Michael Smith (1994), we can avoid the problem by invoking a distinction between two models of the relationship between one's actual self and one's ideal self: the *example model* and the *advice model*.⁵ On the example model, one

⁵ There is a structural parallel between my analysis of justification and Smith's (1994: Ch. 5) analysis of normative reasons in terms of hypothetical desires, according to which one has a reason to ϕ if and only if one would desire that one ϕ if one were ideally rational. However, Smith invokes a different kind of idealization

has justification to believe a proposition if and only if one's ideal self would believe it (in the evaluating world). On the advice model, in contrast, one has justification to believe a proposition if and only if one's ideal self would advise that one believes the proposition (in the evaluated world). The conditional fallacy problem arises for the example model, but not the advice model, since one's ideal self would not advise one to follow their example in cases where the idealization affects which propositions one has justification to believe. So which propositions one has justification to believe is determined by the advice, rather than the example, of one's ideal self.

What kind of idealization is involved in the analysis of justification in terms of critical reflection? We can idealize an epistemic agent along many different dimensions, but not all of these idealizations are relevant for understanding the concept of justification. For instance, we cannot assume that ideal critical reflection involves omniscience and infallibility, since this would imply that one has justification to believe all and only truths. Ideal critical reflection must be understood in a way that is consistent with the possibility of ignorance and error. 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.

Critical reflection is a purely reflective activity in the sense that it is a matter of revising one's beliefs in light of one's reflection on the evidence that is currently in one's possession. As such, it can be distinguished from practical or social activities that involve the acquisition of new information through empirical investigation or consulting experts. Given a commitment to epistemic pluralism, we can allow that there may be other epistemic properties that are defined by their connection with these more practical or social critical activities. For instance, one lesson of Gilbert Harman's (1973) assassination case is that one knows a proposition only if it is stable under the acquisition of new information that is easily available in one's practical or social environment. Suppose I read about the assassination of the President in an early edition of the newspaper, but then there is a media conspiracy

from mine, since his ideally rational agent is also omniscient and infallible about all the relevant facts.

later in the day to replace it with a modified edition that retracts the story. In that case, my belief is justified because it is stable under reflection on the information in my possession, but it is not knowledge because it is not stable under acquisition of new information that is available in my practical or social environment. After all, justification is merely one among many conditions that are necessary for a belief to be knowledge.

As I understand the idealization, ideal critical reflection requires neither omniscience nor infallibility, but merely requires engaging in reflection on the limited information in one's possession in a way that is ideally rational, reasonable, or justified.⁶ An ideally rational process of critical reflection need not involve the acquisition of more accurate or complete information about the world, but is rather a matter of reflecting on the information in one's possession, which may be both inaccurate and incomplete, and revising one's beliefs in light of those reflections. Therefore, ideal critical reflection guarantees neither omniscience nor infallibility: it is consistent with massive ignorance and error about empirical matters of fact.

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 ideal conditions are defined as conditions in which one believes whatever one has justification to believe, then the analysis is certainly trivial. On the current proposal, however, ideal conditions are defined as those in which one brings one's beliefs into alignment with one's justified higher-order reflections about which beliefs one has justification to hold. It is a substantive commitment of the proposed analysis that

⁶ Foley (1993) gives a closely related 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, where I count them as unjustified because they would not survive appropriate idealization in one's capacity for critical reflection.

beliefs formed in this way are themselves justified. Moreover, in the next section, I argue that this 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.⁷ Access internalism is the thesis that all the epistemic facts about which beliefs (and other doxastic attitudes) one has justification to hold are *reflectively accessible* to one in the following sense:

Access Internalism: one has justification for some belief if and only if one has justification to believe upon reflection alone that one has justification for that belief.

Access internalism implies that the epistemic facts about which beliefs one has justification to hold are *self-intimating* in the sense that if they obtain, then one has justification upon reflection to believe that they do. And it implies that reflection is *infallible* in the sense that such facts obtain if one has justification upon reflection to believe that they do. If access internalism is true, then ideal critical reflection is incompatible with the possibility of ignorance or error about which beliefs one has justification to hold. My aim in this section is to argue for access internalism by appealing to the analysis of justification defended in the previous section.⁸

Once again, I take my starting point from Alston's work. Alston argues that a belief is justified and so has the potential to survive critical reflection only if its

⁷ See Goldberg (this volume) for a different perspective on the debate between internalism and externalism. Strictly speaking, there is no one such debate, but a cluster of related debates that are often grouped together under the same rubric.

⁸ In Smithies (2012), I argue 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*.'

justifying ground or basis – what he calls a “justifier” – is accessible to the subject upon 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)

And yet 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 so accessible. As a result, one might have access to one’s justifiers without having access to the facts about which beliefs they justify.

The problem with Alston’s internalist externalism is that it undermines his proposal about the connection between justification and critical reflection. On his view, one’s beliefs can be justified without thereby having what it takes to survive an ideal process of critical reflection. Indeed, Alston seems to acknowledge 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 the fact that it justifies my belief is not. In that case, my belief is not stable under ideal critical reflection. After all, the aim of critical reflection is to bring my beliefs into line with my justified higher-order reflections about which beliefs I have justification to hold. If I don't have justification to believe upon reflection that I have justification for my perceptual belief, then it cannot survive an ideal process of critical reflection. So, my belief is stable under ideal critical reflection only if what is accessible to me upon reflection alone includes not only my justifiers, but also the facts about which beliefs they justify.

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 upon reflection alone to believe 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 upon 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 higher-order 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 upon reflection

that one has justification to believe that proposition. Access internalism follows given the analytic 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 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 and for which it reserves the term 'justification'. To accept the existence of such a property, while using the term 'justification' to pick out a different property, is to oppose access internalism on purely terminological grounds. Those who oppose access internalism on substantive grounds must argue that no important epistemic property is accessible in the sense defined. In what follows, I will defend access internalism against various arguments of this kind.

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.⁹ 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,

⁹ 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.

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.¹⁰

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 it implies that 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.

¹⁰ The regress problem is developed in different ways by Alston (1989: Ch. 8), Greco (1990), Bergmann (2006: Ch. 1) and Kornblith (2012: Ch. 1).

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.¹¹ On this view, epistemic justification is a source of epistemic obligations,

¹¹ 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 that my argument for access internalism in section two does not rely on a deontological conception of justification, unlike those criticized by Alston (1989: Ch. 8), Plantinga (1993: Ch. 1), Goldman (1999), and Bergmann (2006: Ch. 4).

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 has justification to believe owing to the corrupting influence of drugs, brainwashing or mental illness.¹² 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

¹² Examples of this kind are described by Feldman and Conee (1985: 17), Alston (1989: 95-6), Pryor (2001: 114-5) and Christensen (2004: 161-2).

ought to do it right. It doesn't matter that in some cases we 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.¹³

The argument equivocates between thick and thin 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

¹³ See Pryor (2001: 115, n. 36) for a related distinction between thick and thin senses of 'obligation'.

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. Epistemic Intuitions

Michael Bergmann (2006: Ch.1) argues that access internalism is motivated by intuitions about cases that cannot be explained except by generating an infinite regress of a vicious kind. 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 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. Some proponents of reliabilism 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 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, the same objection applies to any theory of justification that does not impose the following necessary condition on justified belief:

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.¹⁴ 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 Bergmann's dilemma can be avoided. But what about the second horn?

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

Bergmann claims that any version of access internalism that is weak enough to avoid the regress problem is also too weak to explain internalist intuitions and so is therefore unmotivated. In what follows, I explain why this is mistaken.

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 reflectively justified 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 reflectively justified 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, we understand this connection rather differently, since Bonjour endorses a doxastic version of access internalism, whereas I endorse a propositional version. Therefore, I propose the following alternative to Bonjour's explanation of why Norman's belief is unjustified:

- (1) One has justification to believe that p if and only if one has justification upon reflection to believe that one has justification to believe that p .
- (2) Norman does not have justification to believe upon reflection 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 engages in critical reflection or even whether he has the capacity to do so. Rather, the crucial issue is whether his belief is based in such a way that it has the potential to survive an idealized process of critical reflection. Norman's belief is unjustified because it does not satisfy this condition. After all, Norman has no justification to believe upon reflection that he has justification to believe that the President is in New York City. Even proponents of reliabilism must grant this premise, since he has no reliable way of establishing his own reliability. And, as Bonjour writes in the passage quoted above, "Critical reflection precludes believing things to which one has, to one's knowledge, no reliable means of epistemic access" (1985: 42).

I conclude that Bergmann's dilemma fails because the propositional version of access internalism defended here is weak enough to avoid a vicious kind of infinite regress, while also being strong enough to explain and vindicate the relevant internalist intuitions about cases. But I also want to challenge Bergmann's

assumption that access internalism is motivated primarily by intuitions about cases. 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. On the account proposed here, it is not simply a brute deliverance of intuition that Norman's clairvoyant beliefs are unjustified. It is the conclusion of an argument whose premises are independently motivated by background theoretical considerations about the connection between justification and critical reflection.

This has important methodological implications for the role of intuitions in epistemology. The practice of relying on intuitions about cases as data for an epistemological theory has been the target of much recent criticism in experimental philosophy. For instance, Weinberg, Nichols and Stich (2001) conducted surveys that revealed 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).

One possibility, of course, is that the appearance of disagreement is illusory, since different parties may use the term 'justification' to pick out different epistemic properties. If so, there need be no substantial disagreement about which epistemic properties are instantiated in a given case, as opposed to a merely terminological disagreement about which of these properties deserves the name 'justification'. Indeed, this is exactly Alston's conclusion about the role of conflicting intuitions in the debate between internalism and externalism:

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)

But this passage also suggests a useful strategy for avoiding terminological disputes about whether the word 'justification' applies in a particular case. Instead, we can ask whether the case instantiates properties that are suited to play an important and distinctive role in epistemic evaluation. This is a substantive question, rather than a merely terminological one, and while there is scope for disagreement, this is likely to be much more tractable than brute disagreement about cases.

If this is right, then debates about the nature of justification need not ground out in conflicting intuitions about cases. On the contrary, intuitions about cases need to be vindicated in light of theoretical considerations about the point and purpose of epistemic evaluation. A theory of justification should not merely conform with intuitions about cases, but should also vindicate those intuitions by explaining how they track a property that is suited to play an important role in epistemic evaluation. Insofar as 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 such a role. In other words, a theory of justification should aim for a kind of reflective equilibrium between intuition and theory. This is why the methodology of experimental philosophy is no substitute for the methodology of analytical epistemology, since theoretical considerations must play a role in deciding which intuitions to accept and which to reject.

5. The Value of Reflection

My main aim in this paper is to argue that justification is an important dimension of epistemic evaluation because it captures an epistemic ideal of stability under rational reflection. But the argument so far leaves a crucial question unanswered. Why is this an epistemic ideal worth caring about – that is, why should we care about having beliefs that are stable under rational reflection?

One answer is that rational reflection makes us more reliable and hence that beliefs that are stable under rational reflection are objectively more likely to be true. After all, rational reflection can increase our reliability by weeding out logical fallacies, hasty generalizations, baseless prejudice, and wishful thinking. But there are at least two problems with this answer. The first point is that rational reflection

is not guaranteed to make us more reliable, since it is consistent with massive ignorance and error about the external world. And the second point is that whatever reliability can be achieved through rational reflection can in principle be achieved without it by means of reliable first-order belief-forming mechanisms.

These points might lead one to conclude that stability under reflection has no distinctive epistemic value. Thus, Hilary Kornblith writes:

From an epistemological point of view, we should value reflection to the extent that, and only to the extent that, it contributes to our reliability. Epistemologically speaking, there is no reason to value reflectively arrived at belief in general over unreflective belief. (2012: 34)

In this passage, Kornblith seems to assume a monistic conception of epistemic value on which truth is the only intrinsic epistemic good.¹⁵ On this view, the epistemic value of reflection must be explained instrumentally in terms of its reliability or conduciveness towards truth. On a more pluralistic conception of epistemic value, there are multiple dimensions of epistemic goodness not all of which can be explained in terms of their reliability or truth-conduciveness. I think it would be question begging in the present context to assume that reliability is the only dimension of epistemic value, since many internalist theories of justification deny that its value can be explained in terms of reliability. Even so, the challenge remains to explain why reflection is valuable given that it is not always guaranteed to increase reliability. Are there any benefits that reflection provides that cannot be achieved in any other way?

My proposal is that reflection is important because it is the distinguishing mark of *persons* – that is, subjects who can be held responsible for their beliefs and actions. As I use the concept, personhood is not a biological category, but an evaluative one. Frankfurt puts the point eloquently in the following passage:

¹⁵ See DePaul (2001) for the distinction between monism and pluralism about epistemic value.

The criteria for being a person do not serve primarily to distinguish the members of our own species from the members of other species. Rather, they are designed to capture those attributes which are the subject of our most humane concern with ourselves and the source of what we regard as most important and most problematical in our lives. (1971: 6)

Even if ours is the only species whose members meet the criteria for being persons, there is no reason in principle to suppose that members of other species, such as intelligent aliens, couldn't satisfy them too.

Persons are distinguished from other animals by the fact that they can be held responsible for their beliefs and actions. This is why we regard it as legitimate to make moral and rational demands on them by subjecting them to what Strawson (1962) called *the reactive attitudes*, such as praise and blame, gratitude and resentment, and so on. We don't adopt these attitudes towards other animals. As Kornblith remarks: "When my neighbor's dog runs loose in my garden and destroys the flowers, it is not the dog who is responsible, but my neighbor" (2012: 75).

Many animals have the capacity for justified belief and action, but they cannot be held responsible for these beliefs and actions, so responsibility is not the same as justification: it is a more demanding status. The question that remains is why it is that persons can be held responsible for the justification of their beliefs and actions in a way that other animals cannot. And here I can see no plausible alternative to the traditional Lockean answer that persons are distinguished from other animals by their capacity for reflection.¹⁶

The argument so far takes the form of an inference to the best explanation: the capacity for reflection is suitably correlated with responsibility and is therefore well placed to explain it. But more needs to be said to elucidate the connection. Why should responsibility require any capacity for reflection? The answer is that being

¹⁶ On Locke's (1968: II xxvii 9) definition, a person has not only reason and consciousness, but also reflection and self-consciousness; 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."

responsible for one's beliefs and actions is a matter of being an appropriate target of reactive attitudes whose function to make demands on one to comply with certain normative standards. But this in turn requires that one has some understanding of the normative standards that govern one's beliefs and actions together with some capacity to bring this understanding to bear in regulating one's beliefs and actions. That is to say, it requires the capacity for reflection.

Here I think it helps to reflect on the rationale for our practice of subjecting one another to reactive attitudes in the first place. The point of adopting reactive attitudes is to make demands on one another to comply with normative standards of morality and rationality. We don't adopt reactive attitudes towards other animals because there is no point in doing so: they cannot understand the demands we are thereby placing on them. Animals can be more or less sensitive to the normative standards of justified belief and action, but they cannot understand those normative standards and bring this understanding to bear in regulating their beliefs and actions. That is why they cannot be held responsible for the justification of their beliefs and actions. Being responsible for the justification of one's belief and actions requires not only sensitivity to reasons, but also a capacity for conceptualizing reasons as reasons. But doing this in a fully explicit way requires higher-order reflection upon one's beliefs and other attitudes.

I don't claim any originality for this line of argument: it is drawn more or less directly from Burge's work on the role of the first-person concept in reflection. Here is a representative passage:

A being that lacked the first-person concept could be sensitive to the norms of reason, and might (I am conceding for the sake of the argument) even sensitively shape its attitudes according to a conception of good and bad reasons and reasoning. But the agent would lack full conceptualization of what it is doing. [...] Insofar as *full* intellectual (or any other) responsibility requires the capacity to understand the way norms govern agency and the capacity to acknowledge the responsibility, a being that lacked the first-person concept would not be fully responsible intellectually. It would not

have a fully realized rational agency. Conceptualized self-consciousness seems a necessary condition for fully responsible agency. Using the first-person concept is necessary to being a fully realized person. (1998: 262)

My point is just that this line of argument can be used to answer the challenge we started with. Reflection is valuable not because of its reliability, but because it is the *sine qua non* for being a person who can be held responsible for their beliefs and choices. Since personhood is intrinsically valuable, reflection is too. To explain the value of reflection in terms of reliability alone is to miss the evaluative significance of the distinction between persons and other animals.

6. Concluding Remarks

Many internalist theories of justification emphasize a 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, the connection between justification and responsibility has been widely misunderstood. Bonjour and other proponents of internalism are mistaken insofar as they claim that a justified belief is one that is held in a way that is responsible. As many proponents of externalism have protested, 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 insofar as they claim that we can allow for the possibility of 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 species that preserves the connection with responsible critical reflection and an externalist species that severs the connection.¹⁷

¹⁷ Sosa (1991: 145) draws a 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.

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 maintain that all justified beliefs have the potential to be held on the basis of responsible critical reflection.

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.¹⁸

¹⁸ This paper was presented at the Chambers Conference at the University of Nebraska, Lincoln in September 2010 and 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 helpful comments and discussion.

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. 1998. "Reason and the First Person." In *Knowing Our Own Minds*, edited by C. Wright, B. Smith and C. Macdonald. Oxford University Press.
- 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 *Knowledge, Truth and Duty*, edited by M. Steup. 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 *Values and Morals*, edited by A. Goldman and J. Kim. Dordrecht: Kluwer.
- Foley, Richard. 1993. *Working Without a Net: A Study of Egocentric Epistemology*. Oxford University Press.

- Frankfurt, H. 1971. Freedom of the Will and the Concept of a Person. *Journal of Philosophy* 68.1: 5-20.
- Gettier, Edmund. 1963. Is Justified True Belief Knowledge? *Analysis* 23.6: 121-3.
- Goldman, Alvin. 1979. "What is Justified Belief?" In *Justification and Knowledge*, edited by G. Pappas. 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.
- Harman, Gilbert. 1973. *Thought*. Princeton University Press.
- Haslanger, Sally. 1999. "What Knowledge Is and What It Ought To Be: Feminist Values and Normative Epistemology." *Philosophical Perspectives* 13: 459-80.
- Kornblith, Hilary. 2012. *On Reflection*. Oxford University Press.
- 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.
- Smith, Michael. 1994. *The Moral Problem*. Oxford: Blackwell.
- Smithies, Declan. 2012. "Moore's Paradox and the Accessibility of Justification." *Philosophy and Phenomenological Research* 85 (2): 273-300.
- 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.