

# Moore's Paradox and the Accessibility of Justification

DECLAN SMITHIES

*The Ohio State University*

## 1. The Accessibility of Justification

One dimension of the debate between internalism and externalism in epistemology concerns the nature and extent of one's epistemic access to facts about justification. For current purposes, internalism can be roughly stated as the view that one always has a special kind of epistemic access to facts about which propositions one has justification to believe, whereas externalism is the denial of internalism. Many internalists regard the accessibility of justification as such an essential part of its nature that denying it is tantamount to changing the subject. On this view, nothing deserves the name 'justification' unless it is accessible in the relevant way. But what exactly does it mean to say that justification is accessible? And why should anyone suppose that justification is accessible?

My aim in this paper is to argue that facts about which propositions one has justification to believe are accessible in the following sense:

### **The Accessibility Thesis:**

- **Positive:** one has justification to believe that  $p$  iff one has justification to believe that one has justification to believe that  $p$  ( $Jp \leftrightarrow JJp$ ).
- **Negative:** one lacks justification to believe that  $p$  iff one has justification to believe that one lacks justification to believe that  $p$  ( $\sim Jp \leftrightarrow J\sim Jp$ ).

The accessibility thesis consists of two biconditionals—positive and negative—each of which comprises a self-intimation thesis and an infallibility thesis:

### The Self-Intimation Thesis:

- **Positive:** if one has justification to believe that  $p$ , then one has justification to believe that one has justification to believe that  $p$  ( $Jp \rightarrow JJp$ ).
- **Negative:** if one lacks justification to believe that  $p$ , then one has justification to believe that one lacks justification to believe that  $p$  ( $\sim Jp \rightarrow J\sim Jp$ ).

### The Infallibility Thesis:

- **Positive:** if one has justification to believe that one has justification to believe that  $p$ , then one has justification to believe that  $p$  ( $JJp \rightarrow Jp$ ).
- **Negative:** if one has justification to believe that one lacks justification to believe that  $p$ , then one lacks justification to believe that  $p$  ( $J\sim Jp \rightarrow \sim Jp$ ).

The facts about which propositions one has justification to believe are *self-intimating* in the sense that if they obtain, then one has meta-level justification to believe that they obtain. Moreover, one's meta-level justification is *infallible* in the sense that if one has meta-level justification to believe that those facts obtain, then they do in fact obtain. So, whether or not one has justification to believe a proposition, one always has meta-level justification to form a belief on the matter, since the facts in question are self-intimating. Moreover, whichever belief one has meta-level justification to form, it is guaranteed to be true, since one's meta-justification is infallible. In that sense, one always has epistemic access to the facts about whether or not one has justification to believe any given proposition.<sup>1</sup>

The concept of justification that figures in the accessibility thesis cannot be defined in more basic terms, but it can be elucidated in terms of related notions, such as reasons and rationality. To say that one has justification to believe a proposition is to say that it is rational or reasonable for one to believe it, which is a matter that depends in turn on the strength of one's reasons to believe it. It is reasonable for one to believe a proposition if and only if one has reasons to believe it, which

---

<sup>1</sup> Williamson (2000: Ch.4) argues that no conditions are luminous in the sense that one is in a position to know that they obtain if and only if they obtain. See Smithies (forthcoming, a) for discussion of Williamson's anti-luminosity argument and its bearing on the accessibility thesis defended here.

are not outweighed or defeated by stronger reasons not to believe it, but to disbelieve it or to withhold belief instead. So, to say that one has justification to believe a proposition is to say that, all things considered, belief is the most reasonable doxastic attitude for one to adopt towards the proposition in question.<sup>2</sup>

For the purposes of this paper, I will operate with the tripartite distinction between doxastic attitudes of belief, disbelief and withholding belief. To a first approximation, believing a proposition is a disposition to judge that it is true, while disbelieving a proposition is a disposition to judge that it is false, and withholding belief in a proposition is a disposition to judge that it is an open question whether it is true or false. Thus, withholding belief in a proposition is not merely the absence of a doxastic attitude towards a proposition, but is rather a *sui generis* attitude of open-mindedness with respect to the truth-value of the proposition in question.

I will assume that justification is *exhaustive* in the sense that, for any proposition, one has justification either to believe, to disbelieve, or to withhold belief in that proposition. I will also assume that justification is *exclusive* in the sense that if one has justification to adopt one of these doxastic attitudes towards a proposition, then one does not have justification to adopt any of the other doxastic attitudes towards the same proposition. Exhaustiveness and exclusiveness entail *uniqueness*, which is the thesis that, for any proposition, there is a unique doxastic attitude—that is, either belief, disbelief, or withholding—such that one has justification to adopt that attitude towards the proposition in question.<sup>3</sup>

The overall plan for the paper is as follows. In section two, I respond to some standard objections to the accessibility thesis by arguing that it should be formulated as an epistemic thesis, rather than a doxastic thesis. In section three, I argue that the accessibility of justification solves an epistemic version of Moore's paradox and in section four, I argue more ambitiously that this epistemic version of Moore's paradox cannot be solved except by appeal to the accessibility of justification. In section five, I use this epistemic version of Moore's paradox to vindicate the intuitions prompted by BonJour's (1985) clairvoyance cases and to challenge Goldman's (1986) reliabilist treatment of these

---

<sup>2</sup> This is a threshold notion of justification, rather than a graded notion. I hope to discuss analogues of the accessibility thesis for graded notions of justification elsewhere.

<sup>3</sup> White (2005) argues for a version of the uniqueness thesis that is formulated in terms of subjective probabilities, while Feldman (2007) defends a version that is formulated in terms of the tripartite distinction between belief, disbelief and withholding.

cases. I conclude in section six by considering how the accessibility of justification might be explained and how it imposes substantial constraints on a theory of the determinants of justification.

## 2. Epistemic and Doxastic Accessibility

As I have formulated it, the accessibility thesis is not a doxastic thesis about which of one's beliefs are justified, but is rather an epistemic thesis about which propositions one has justification to believe.<sup>4</sup> Justified belief requires not only *having* justification to believe a proposition, but also *using* it in believing that proposition on the basis of one's justification to believe it. By contrast, having justification to believe a proposition does not require using it in forming a justified belief. As we shall see, this distinction between epistemic and doxastic versions of the accessibility thesis is crucial for avoiding the problems of over-intellectualization and vicious regress.

On the *epistemic* version of the accessibility thesis, one has justification to believe a proposition if and only if one has meta-level justification to believe that one does. On the *doxastic* version, by contrast, one's belief is justified if and only if it is held on the basis of an appropriate meta-level justification. 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  $\varphi$ .
- (2) Beliefs having feature  $\varphi$  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 accessibility thesis faces an over-intellectualization objection, since it implies that one has justified beliefs only if one has psychological and epistemic concepts and the capacity to use them in meta-justificatory reasoning. Moreover, it is subject to a regress

---

<sup>4</sup> Following Firth (1978), this distinction is usually drawn using the terminology of *propositional* versus *doxastic* justification, but see also Goldman's (1979) distinction between *ex ante* and *ex post* justification, which is discussed below.

objection, since it implies that one has justified beliefs only if one has an infinite hierarchy of increasingly complicated meta-justified beliefs.<sup>5</sup> However, these objections can be avoided if we formulate the accessibility thesis in epistemic, rather than doxastic, terms. The epistemic version implies that one has justification to believe a proposition only if one has an appropriate meta-justification—indeed, an infinite regress of meta-justifications—but it does not imply that one’s justified beliefs are *based* on these meta-justifications or even that one has the doxastic capacities, including conceptual capacities, required to base one’s beliefs in this way. It therefore avoids the over-intellectualization objection and it generates an infinite regress of meta-justifications that is not vicious, but benign.

It may be objected that while *having* justification to believe a proposition does not require *using* it, it does at least require having the *capacity* to use it in forming a justified belief. If so, then the epistemic version of the accessibility thesis implies that one has justification to believe a proposition only if one has the capacity to believe an infinite hierarchy of increasingly complicated meta-justificatory propositions, which has obvious skeptical consequences. However, I will argue that there is in fact no compelling basis for the assumption that the limits on one’s doxastic capacities impose corresponding limits on which propositions one has justification to believe. Compare Richard Feldman and Earl Conee, who write:

There is no basis for the premise that what is epistemically justified must be restricted to feasible doxastic alternatives.... Suppose that there were occasions when forming the attitude that best fits a person’s evidence was beyond normal cognitive limits. This would still be the attitude justified by the person’s evidence. If the person had normal abilities, then he would be in the unfortunate position of being unable to do what is justified. (1985: 19)

In my view, there are no doxastic limits on which propositions one has justification to believe, but only which justified beliefs one has a capacity to form.

Why might one suppose that there are doxastic limits on which propositions one has justification to believe? Some epistemologists simply define having justification in terms of the capacity to use it: thus, one has justification to believe a proposition if and only if one

---

<sup>5</sup> Bergmann (2006: Ch.1) provides a detailed presentation of the regress argument against internalism, although he formulates the internalist access requirement in terms of the notion of “awareness” and does not consider the formulation in terms of propositional justification proposed here. See Smithies (forthcoming, b) for a more detailed discussion of Bergmann’s argument.

has a capacity to use it in forming a justified belief. For instance, consider Alvin Goldman's definition of the distinction between *ex ante* and *ex post* justification:

Person S is *ex ante* justified in believing *p* at *t* if and only if there is a reliable belief-forming operation available to S which is such that if S applied that operation to his total cognitive state at *t*, S would believe *p* at *t*-plus-delta (for a suitably small delta) and that belief would be *ex post* justified. (1979: 21)

An immediate consequence of this definition is that limits on one's doxastic capacities impose corresponding limits on which propositions one has justification to believe. Arguably, however, this reverses the correct direction of explanation: one has a capacity to form a justified belief if and only if one has justification to believe the proposition in question and, in addition, one has the doxastic capacities required to use one's justification in forming a justified belief. This leaves it open whether or not one always has the doxastic capacities required for using one's justification to believe a proposition. Moreover, this is not a question that should be closed by definition.

Perhaps the most influential line of argument for doxastic limits appeals to a *deontological* conception of justification as a source of epistemic obligations, which are binding only insofar as one has the psychological capacities required to discharge them.<sup>6</sup> The argument proceeds roughly as follows:

- (1) If one has justification to believe a proposition, then one ought to believe it on the basis of one's justification.
- (2) If one ought to believe a proposition on the basis of one's justification, then one can believe it on the basis of one's justification.
- (3) So, if one has justification to believe a proposition, then one can believe it on the basis of one's justification.

However, the conclusion of this argument is subject to well known counterexamples in which drugs, brainwashing or mental illness destroy

---

<sup>6</sup> Christensen (2004: Ch. 6) discusses this argument in the context of a useful discussion of the role of idealization in probabilistic models of rationality. See Alston (1988) for a related line of argument that the deontological conception of justification implies that one has voluntary control over one's beliefs.

one's doxastic capacity to believe what one has justification to believe.<sup>7</sup> In such a case, one's doxastic limitations excuse one from blame for one's epistemic failings, but blamelessness is not sufficient for justification. It is certainly plausible that there are doxastic limits on blameworthiness, but it does not follow that there are any corresponding doxastic limits on which propositions one has justification to believe.

What are the options for responding to this argument? One option is to deny the second premise by arguing that 'ought' does not always imply 'can'. Thus, Richard Feldman (2000) argues that there are so-called 'role oughts' which apply to anyone who plays a certain role, regardless of how well they are capable of playing that role—thus, 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 we are unable to do so." (2000: 676)

Another option is to deny the first premise of the argument by rejecting the deontological conception of justification in favour of an evaluative conception. Thus, William Alston (1985) argues that justification is a source of epistemic values or evaluative ideals, rather than obligations. Deciding between these options depends on how we understand the relationship between values and obligations. Are we obliged to achieve evaluative ideals or merely to approximate them as closely as we can? There may be an attenuated sense in which we ought to achieve ideals regardless of whether we are capable of doing so.<sup>8</sup> But there seems to be a more robust sense, which is more closely connected to assessments of praise and blame, in which we are obliged merely to approximate towards ideals to the extent that we are capable of doing so.<sup>9</sup> In this sense, epistemic obligations can be reconstructed from epistemic ideals together with further assumptions about the nature and extent of our contingent psychological limitations.

If justification is a source of evaluative ideals, then there is no reason to assume that there are doxastic limits on which propositions one has justification to believe. On this view, which propositions one has justification to believe corresponds to an evaluative ideal, which abstracts away from contingent facts about the limits on one's doxastic

---

<sup>7</sup> Related examples are discussed by Feldman and Conee (1985: 17), Alston (1985: 67–8), Pryor (2001: 114–5) and Christensen (2004: 161–2).

<sup>8</sup> Thus, Feldman and Conee claim: "In any case of a standard for conduct ... it is appropriate to speak of 'requirements' or 'obligations' that the standard imposes." (1985: 19)

<sup>9</sup> Pryor (2001: 115, fn. 36) draws a related distinction between thick and thin notions of 'obligation'.

capacities. Roughly, the propositions that one has justification to believe are just those propositions that one would believe if one were to be idealized in relevant respects. However, ideals need not be humanly attainable. Epistemic ideals—like ideals of morality, scientific understanding, and chess—may lie beyond our limited human capacities. Therefore, the limits on one’s doxastic capacities do not constrain the epistemic ideal, but only the extent to which one is capable of approximating towards the ideal. As David Christensen (2004: 162) 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.”

One might object that this idealized conception of justification loses touch with our ordinary practices of epistemic evaluation. To illustrate, Frege’s belief in Axiom V was justified by ordinary standards, but it was not ideally justified, since it was not suited to survive ideal critical reflection on the set-theoretical paradoxes later discovered by Russell. And yet we can recognize the distinction between ordinary and idealized conceptions of justification while maintaining that there is nevertheless an explanatory connection between them. Being justified by ordinary standards is a matter of meeting some contextually determined threshold on a scale that is defined by reference to the ideal—in other words, it is a matter of approximating towards the ideal to a sufficiently high degree. On this view, we can reconstruct everyday notions of justification by reference to an idealized notion of justification that abstracts away from our contingent psychological limitations together with further assumptions about the nature and extent of those limitations.

So far, I have articulated an epistemic version of the accessibility thesis and I have defended it against some preliminary objections, but I have not yet given any positive arguments in its favour. In the remaining sections of this paper, I will argue for the accessibility of justification on the grounds that anyone who denies it is thereby faced with an epistemic version of Moore’s paradox.

### 3. Moore’s Paradox

As G. E. Moore famously noticed, there is something interestingly defective involved in asserting conjunctions of the following forms:

- (1)  $p$  and I do not believe that  $p$ .
- (2) I believe that  $p$  and it is not the case that  $p$ .



The American journalist, H. L. Mencken, gives an amusing example: “All men are frauds. The only difference between them is that some admit it. I myself deny it.” There is something manifestly fraudulent about a man who asserts that all men are frauds and then proceeds to assert that he denies it. Indeed, he comes dangerously close to contradicting himself, but like any good fraud, he is hard to pin down. What he asserts is not a contradiction of the form, ‘ $p$  and it is not the case that  $p$ ,’ but something closer to a Moorean conjunction of the form, ‘ $p$  and *I believe* that it is not the case that  $p$ .’

What exactly is wrong with asserting Moorean conjunctions? On the one hand, asserting Moorean conjunctions seems absurd or self-defeating in much the same way as asserting contradictions. On the other hand, Moorean conjunctions are not contradictions. Given that I am neither omniscient nor infallible, it is quite possible that I do not believe that  $p$  when  $p$  is true or that I believe that  $p$  when  $p$  is false. Moorean conjunctions can be true, but they cannot be asserted without absurdity. The problem of explaining why this is so has come to be known as *Moore’s paradox*.<sup>10</sup>

As Sydney Shoemaker (1996) observes, a solution to Moore’s paradox should explain why believing a Moorean conjunction is just as bad as asserting it. The absurdity is much the same whether or not one’s belief is given linguistic expression through the speech act of assertion. Since Moore’s paradox is not a purely linguistic phenomenon, it cannot be solved by appealing to purely linguistic norms, such as Gricean norms of conversational implicature. A more promising strategy is to appeal to more fundamental epistemic norms governing belief together with derivative norms linking assertion and belief. If we can explain what is wrong with believing Moorean conjunctions, then we can explain what is wrong with asserting Moorean conjunctions by invoking Shoemaker’s principle, which states: “What can be (coherently) believed constrains what can be (coherently) asserted.” (1996: 76)

A plausible account of Moorean absurdity emerges once we shift our focus from assertion to belief. Intuitively, believing Moorean conjunctions is *irrational*. Moreover, the irrationality of believing Moorean conjunctions does not derive from believing one conjunct or the other, but from the conflict between believing both conjuncts simultaneously. It is irrational to believe that  $p$  while disbelieving that one believes that  $p$  and it is irrational to believe that one believes that  $p$  while disbelieving that  $p$ . Indeed, we can take this one step further: it is irrational to believe that  $p$  while withholding belief that one believes that  $p$  or to

---

<sup>10</sup> See the introduction to Green and Williams (2007) for a history and overview of Moore’s paradox.

believe that one believes that  $p$  while withholding belief that  $p$ .<sup>11</sup> This form of irrationality is harder to convey by means of a Moorean conjunction, since there is no speech act that naturally functions to express withholding in the way that assertion functions to express belief. But if an assertion of the form, 'It is an open question whether  $p$ ,' can be used to express withholding, then we can add the following Moorean conjunctions:

- (3)  $p$  and it is an open question whether I believe that  $p$ .
- (4) I believe that  $p$  and it is an open question whether  $p$ .

Any solution to Moore's paradox should explain why (3) and (4) are defective in much the same way as (1) and (2) above.

To sum up, believing a proposition rationally commits one to believing that one believes it and believing that one believes a proposition rationally commits one to believing it. In other words, it is irrational to believe either (i) the proposition that  $p$  or (ii) the proposition that one believes that  $p$ , while disbelieving or withholding belief in the other. This leads Shoemaker to conclude that if one is fully rational, then one believes that  $p$  if and only if one believes that one believes that  $p$ . Thus, he writes: "in a rational person belief that  $p$  brings with it belief that one believes that  $p$ " and conversely, "a rational person who believes that she believes that  $p$  thereby believes that  $p$ ." (1995: 225–6)

Moore noticed that there are epistemic counterparts of (1) and (2), which are formulated in terms of knowledge, rather than belief:

- (5)  $p$  and I do not know that  $p$ .
- (6) I know that  $p$  and it is not the case that  $p$ .

Timothy Williamson (2000: Ch.11) has argued that the epistemic version of Moore's paradox is explained by the knowledge rule, which governs the permissibility of assertion, and its inner analogue, belief:

- The knowledge rule: "One must: assert [or believe]  $p$  only if one knows  $p$ ." (2000: 243)

Williamson argues that one cannot know either (5) or (6) and so, by the knowledge rule, one must not believe or assert them. The argument

---

<sup>11</sup> Versions of this point are made by Shoemaker (1995: 215) and Chalmers and Hájek (2007: 171).

relies on the premise that one knows a conjunction only if one knows each of its conjuncts. If one knows the first conjunct of (5), then its second conjunct is false and so one cannot know it, since knowledge is factive. Similarly, if one knows the second conjunct of (6), then its first conjunct is false and so one cannot know it, since knowledge is factive. However, the knowledge rule fails to explain what is wrong with believing or asserting epistemic counterparts of (3) and (4):

(7)  $p$  and it is an open question whether I know that  $p$ .

(8) I know that  $p$  and it is an open question whether  $p$ .

And yet any adequate solution to the epistemic version of Moore's paradox should explain why (7) and (8) are defective in much the same way as (5) and (6) above.<sup>12</sup>

As before, the irrationality of believing these Moorean conjunctions stems from the conflict between believing both conjuncts simultaneously. It is irrational to believe either (i) the proposition that  $p$ , or (ii) the proposition that one knows that  $p$ , while disbelieving or withholding belief in the other. Thus, we might extend Shoemaker's conclusion by claiming that if one is fully rational, then one believes that  $p$  if and only if one believes that one knows that  $p$ . In other words, believing a proposition rationally commits one to believing that one knows it, and vice versa. As Michael Huemer writes, "If one believes that  $p$ , one is thereby *rationally committed* to taking one's belief to be knowledge." (2007: 145)

The Moorean conjunctions we have considered so far can be reformulated in terms of justification, rather than knowledge or belief:

(9)  $p$  and I do not have justification to believe that  $p$ .

(10) I have justification to believe that  $p$  and it is not the case that  $p$ .

(11)  $p$  and it is an open question whether or not I have justification to believe that  $p$ .

(12) I have justification to believe that  $p$  and it is an open question whether or not  $p$ .

---

<sup>12</sup> Sosa (2007: 115–6) makes related points in his discussion of the problem of the criterion.

These conjunctions are defective in much the same way as those already considered. In particular, these conjunctions can hardly be any more acceptable than those formulated in terms of knowledge, since (5) – (8) entail (9) – (12) on the assumption that one knows a proposition only if one has justification to believe it.

As before, the irrationality of believing these Moorean conjunctions stems from the conflict between believing both conjuncts simultaneously. It is irrational to believe either (i) the proposition that  $p$ , or (ii) the proposition that one has justification to believe that  $p$ , while disbelieving or withholding belief in the other. In other words, believing a proposition rationally commits one to believing that one has justification to believe it and vice versa. In this case, the extended version of Shoemaker's conclusion is that if one is fully rational, then one believes that  $p$  if and only if one believes that one has justification to believe that  $p$ .

What explains this fact about rationality? If justification is accessible, then we have a ready explanation—namely, that one has justification to believe a proposition if and only if one has justification to believe that one does. Given exclusiveness, if one has justification to believe a proposition, then one lacks justification either to disbelieve or to withhold belief in the proposition that one has justification to believe it. By the same token, if one has justification to believe that one has justification to believe a proposition, then one lacks justification either to disbelieve it or to withhold belief. So, if the accessibility thesis is true, then it is always irrational to believe the Moorean conjunctions above, since one cannot have justification to believe each of their conjuncts simultaneously.<sup>13</sup>

The accessibility of justification therefore provides the resources for solving an epistemic version of Moore's paradox. In light of this, one might argue for the accessibility thesis on the basis of an inference to the best explanation. The problem with this strategy is that it requires not only showing that the accessibility thesis explains the data, but also that it does so better than all of the alternatives. We have already seen that some explanatory strategies fail because they are insufficiently general, but I cannot attempt an exhaustive survey here. Instead, my strategy in what follows is to mount a deductive argument that the accessibility thesis is indispensable for explaining the relevant kind of

---

<sup>13</sup> Here I assume that one has justification to believe a conjunction only if one has justification to believe its conjuncts. Moreover I assume that one has justification to disbelieve a proposition iff one has justification to believe it is false and to withhold belief iff one has justification to believe that it is an open question whether it is true or false.

Moorean irrationality and hence that anyone who denies it is unable to solve the epistemic version of Moore's paradox.

#### 4. An Argument from Moore's Paradox

The overall strategy is to argue by *reductio*. Anyone who denies the accessibility thesis is thereby committed to the possibility of having justification for certain combinations of doxastic attitudes. However, these doxastic attitudes are not rationally co-tenable, since they generate a distinctively Moorean kind of irrationality. Moreover, we cannot explain this Moorean irrationality on the assumption that one has justification for the doxastic attitudes in question. But if we assume that one cannot have justification for doxastic attitudes that are not rationally co-tenable, then our assumptions yield a contradiction. We should therefore conclude that the accessibility thesis is true. The argument can be summarized as follows:

- (1) The accessibility thesis is false.
- (2) If the accessibility of justification is false, then one can have simultaneous justification for certain doxastic attitudes.
- (3) These doxastic attitudes are not rationally co-tenable, since they give rise to an epistemic version of Moore's paradox.
- (4) If these doxastic attitudes are not rationally co-tenable, then one cannot have simultaneous justification for these doxastic attitudes.
- (5) So, the accessibility thesis is true (by *reductio*).

The accessibility thesis comprises four conditionals—namely, positive and negative versions of self-intimation and infallibility—so, in order to flesh out this argument, we should consider how the argument applies to each conditional in turn.

##### 1. Negative Infallibility ( $J\sim Jp \rightarrow \sim Jp$ )

If negative infallibility is false, then there is a possible case in which one has justification to believe that one lacks justification to believe that  $p$ , although one has justification to believe that  $p$ . However, it is irrational to believe that  $p$ , while also believing that one lacks justification to believe that  $p$ , since this is to be committed to the Moorean conjunction:

- $p$  and I do not have justification to believe that  $p$ .

## 2. *Positive Self-Intimation* ( $Jp \rightarrow JJp$ )

If positive self-intimation is false, then there is a possible case in which one has justification to believe that  $p$ , although one lacks justification to believe that one has justification to believe that  $p$ . By exhaustiveness, it follows that one has justification either to disbelieve or to withhold belief in the proposition that one has justification to believe that  $p$ . However, it is irrational to believe that  $p$  while disbelieving or withholding belief that one has justification to believe that  $p$ , since this is to be committed to one of the following pair of Moorean conjunctions:

- $p$  and I do not have justification to believe that  $p$ .
- $p$  and it is an open question whether I have justification to believe that  $p$ .

## 3. *Positive Infallibility* ( $JJp \rightarrow Jp$ )

If positive infallibility is false, then there is a possible case in which one has justification to believe that one has justification to believe that  $p$ , although one lacks justification to believe that  $p$ . By exhaustiveness, it follows that one has justification to disbelieve or to withhold belief that  $p$ . However, it is irrational to believe that one has justification to believe that  $p$ , while disbelieving or withholding belief that  $p$ , since this is to be committed to one of the following pair of Moorean conjunctions:

- I have justification to believe that  $p$  and it is not the case that  $p$ .
- I have justification to believe that  $p$  and it is an open question whether  $p$ .

## 4. *Negative Self-Intimation* ( $\sim Jp \rightarrow J\sim Jp$ )

If negative self-intimation is false, then there is a possible case in which one lacks justification to believe that  $p$ , although one lacks justification to believe that one does. By exhaustiveness, it follows that one has justification to disbelieve or to withhold belief that  $p$  and also that one has justification to believe or to withhold belief that one lacks justification to believe that  $p$ . However, it is irrational to disbelieve or withhold belief that  $p$  while also disbelieving or withholding belief that one has justification to disbelieve or withhold belief that  $p$ , since this is to be committed to one of the following Moorean conjunctions:

- I have justification to believe that  $p$  and it is not the case that  $p$ .
- I have justification to believe that  $p$  and it is an open question whether  $p$ .
- It is an open question whether I have justification to believe that  $p$  and it is not the case that  $p$ .
- It is an open question whether I have justification to believe that  $p$  and it is an open question whether  $p$ .

All of these Moorean conjunctions were encountered in the previous section except the very last pair. However, these are no more acceptable than the others. Just as it seems irrational to believe a proposition while withholding on whether one has justification to believe it, so it seems irrational to disbelieve a proposition while withholding on whether one has justification to disbelieve it or to withhold on a proposition while withholding on whether one has justification to withhold. Each of these combinations of doxastic attitudes involves a similar kind of rational instability.

What are the options for responding to this argument? Since its premises entail a contradiction, the only question is which premise should be rejected. I claim that we should reject the first premise, but why not reject another premise instead?

The only option for denying the second premise is to deny exhaustiveness, which states that, for any proposition, one has justification either to believe it, to disbelieve it, or to withhold belief. But I maintain that one cannot have justification to adopt any doxastic attitude towards a proposition without having justification to adopt one of these. After all, these are coarse-grained doxastic attitudes, which are multiply realized by more fine-grained doxastic attitudes, so it is plausible that any distribution of credence realizes at least one of these coarse-grained attitudes.<sup>14</sup> Following Gilbert Harman (1986), one might argue that one has pragmatic justification to avoid cluttering up one's mind by not adopting any doxastic attitude at all towards certain propositions. However, even if this is true, it does not follow that one lacks epistemic justification to adopt some doxastic attitude towards every proposition in the sense that there is a fact of the matter about which

---

<sup>14</sup> For instance, withholding is consistent with many different ways of distributing one's credence, including a sharp credence of 0.5 or a fuzzy credence that is spread out between the interval between 0 and 1. Similar points apply to believing and disbelieving.

of these doxastic attitudes is the most epistemically reasonable one to adopt.

To deny the third premise is to deny that there is any irrationality involved in believing or asserting the Moorean conjunctions in question. However, this is not easy to swallow. How can it be rational to believe a proposition while disbelieving or withholding belief in the proposition that one has justification to believe it? One might claim, in a Wittgensteinian spirit, that not all beliefs stand in need of justification, such as basic beliefs in the law of non-contradiction and the uniformity of nature. Presumably, however, these beliefs are not unreasonable. Recall my stipulation that what it means to say one has justification to believe a proposition is just to say that it is reasonable or rational for one to believe it. By definition, then, it is unreasonable to believe a proposition unless one has justification to believe it. A more plausible Wittgensteinian position is that some basic beliefs are justified by default, which stands in contrast to the claim that they are not justified at all.<sup>15</sup>

To deny the fourth premise is to claim that sometimes one has justification to adopt a combination of doxastic attitudes, which are not rationally co-tenable. But if one has justification to adopt the doxastic attitudes in question, then what explains why they are not rationally co-tenable? The most promising strategy here is to invoke *defeaters*: perhaps one has justification to adopt certain doxastic attitudes, although one's justification is defeated in the very act of forming the attitudes in question.<sup>16</sup>

By way of illustration, let us consider a case in which one has justification to believe a Moorean conjunction owing to recognition of one's own irrationality. Suppose that one has justification to believe that one's team will lose the game, while recognizing that one cannot quite bring oneself to believe this. In that case, one has justification to believe the Moorean conjunction: my team will lose and I do not believe that my team will lose. However, one cannot justifiably believe this Moorean conjunction, since if one believes the first conjunct, then one's justification to believe the second conjunct is thereby defeated. Hence, one's justification to believe the Moorean conjunction is "finkish" in the sense that it is defeated in the very act of forming the belief in question.<sup>17</sup>

---

<sup>15</sup> Wright (2004) develops a position of this kind, although he uses the terminology of 'entitlement' rather than 'justification' to mark out warrants that are had by default.

<sup>16</sup> Many thanks to Ralph Wedgwood for urging me to consider this strategy.

<sup>17</sup> The allusion is to Martin's (1994) finkish dispositions, which are destroyed whenever their manifestation conditions obtain.



Opponents of the accessibility thesis might try to exploit this loophole in explaining Moorean irrationality, while avoiding any commitment to the accessibility thesis. On this view, one may have justification to believe a proposition, while also having justification to disbelieve or to withhold belief in the proposition that one has justification to believe it. However, one's justification for this combination of doxastic attitudes is "finkish" in the sense that if one disbelieves or withholds belief in the proposition that one has justification to believe the proposition in question, then one's justification to believe it is thereby defeated. If this is correct, then we can explain why these doxastic attitudes are not rationally co-tenable while maintaining that one can have justification for these attitudes simultaneously.

Michael Bergmann (2005) invokes defeaters in attempting to explain away the appeal of higher-level requirements on justification. According to Bergmann, it is an analytic truth that a belief *b* is justified only if it has no defeater, since a defeater is defined as any mental state that causes an otherwise justified belief to be unjustified:

- *d* is a defeater at *t* for *S*'s belief *b* iff (i) *d* is an experience or propositional attitude or combination thereof; (ii) *S* comes to have *d* at *t*; (iii) as a result of *S*'s coming to have *d* at *t*, *b* ceases to be justified. (2005: 422)

Bergmann argues that all *believed* defeaters are *actual* defeaters: in other words, if one *believes* that one's belief *b* is unjustified, then it *is* unjustified. Moreover, he argues that if one does not believe that *b* is unjustified, but merely *withholds* belief in the proposition that *b* is justified, then one's belief *b* is unjustified. Therefore, if one adopts any doxastic attitude other than belief towards the proposition that one's belief *b* is justified, then *b* is unjustified. Nevertheless, he argues that there is no higher-level requirement for justification, since it may be that *b* is justified even if one does not adopt any doxastic attitude at all towards the proposition that *b* is justified.

In response to Bergmann, I think it is crucial to draw a distinction between *propositional* and *doxastic* defeaters. A doxastic defeater is a mental state that undermines doxastic justification—that is, it makes it the case that one's beliefs are unjustified. A propositional defeater, on the other hand, is a mental state that undermines propositional justification—that is, it makes it the case that one does not possess justification to believe a proposition. The crucial point is that not all doxastic defeaters are propositional defeaters: in particular, unjustified doxastic

attitudes can function as doxastic defeaters, but not propositional defeaters.<sup>18</sup>

To illustrate, suppose that one has a justified belief that  $p$ , but then one forms unjustified beliefs that if  $p$ , then  $q$ , and not- $q$ , while retaining one's belief that  $p$ . One's belief that  $p$  is now unjustified, since it is irrationally insensitive to the presence of inconsistent beliefs, but it does not follow that one lacks all things considered justification to believe that  $p$ . After all, the most reasonable course of action would be to retain one's belief that  $p$ , while abandoning one's other beliefs. The presence of unjustified beliefs can undermine the doxastic justification of one's belief that  $p$ , but it does not thereby defeat one's propositional justification to believe that  $p$ . Analogously, suppose that one has a justified belief that  $p$ , but then one forms an unjustified higher-order attitude of disbelieving or withholding belief that one has justification to believe that  $p$ . One's belief that  $p$  is now unjustified, since it is irrationally insensitive to the presence of rationally conflicting higher-order attitudes, but it does not follow that one lacks all things considered justification to believe that  $p$ . After all, the most reasonable course of action would be to retain one's belief that  $p$ , while abandoning one's conflicting higher-order attitudes. The presence of unjustified higher-order attitudes can undermine the doxastic justification of one's belief that  $p$ , but it does not thereby defeat one's propositional justification to believe that  $p$ .

James Pryor (2004) draws a related distinction between which propositions one has justification to believe and which propositions one is rationally committed to believing by beliefs that one already has.<sup>19</sup> If one's beliefs are unjustified, then one may be rationally committed to believing propositions that one has no justification to believe. Moreover, one may have justification to believe a proposition, although one is rationally obstructed from using it in forming a justified belief, since one has unjustified beliefs that rationally commit one to disbelieve or to withhold belief in that proposition, rather than believing it. In my

---

<sup>18</sup> Bergmann (2006: 159–60) claims that “there is not much at stake” between the view that what gets defeated are beliefs and the view that what gets defeated are reasons or justifications for beliefs. But this overlooks the important fact that the doxastic justification of one's beliefs can be defeated, while the propositional justification for one's beliefs remains undefeated.

<sup>19</sup> As Pryor notes, Broome (1999) draws a distinction between reasons and normative requirements, which parallels his distinction between justification and rational commitments. In Broome's terms, one is *normatively required* to believe that  $q$  if one believes that  $p$  and if  $p$ , then  $q$ , but it does not follow that one has any *reason* to believe that  $q$ . As Broome develops the point, normative requirements take wide scope and so they don't detach, i.e. there is no valid argument from  $O(Bp \rightarrow Bq)$  and  $Bp$  to  $OBq$ .

terminology, that is a case in which one has a doxastic defeater, but no propositional defeater.

Why isn't it enough for Bergmann's purposes to claim that disbelieving or withholding belief that one has justification to believe that  $p$  is a doxastic defeater, but not a propositional defeater? The answer is that this does not *explain* why the relevant attitudes are not rationally cotenable. All sides agree that it is irrational to believe that  $p$  while disbelieving or withholding belief that one has justification to believe that  $p$ . In other words, disbelieving or withholding belief in the proposition that one has justification to believe that  $p$  *rationally commits* one to not believing that  $p$ . But what explains the existence of this rational commitment? Why is it that disbelieving or withholding belief in the proposition that one has justification to believe that  $p$  rationally commits one to not believing that  $p$ ?

Following Pryor, I claim that rational commitments are best explained in terms of more fundamental facts about the structure of propositional justification:

I think we can understand rational commitments like this. Take a belief the subject happens to have, e.g., his belief in P. Consider what would be the epistemic effects of his *having (decisive) justification* for that belief. ... If one of the effects is that the subject has decisive justification to believe Q, then his belief in P counts as rationally committing him to the belief in Q—regardless of whether he *really does* have any justification to believe P. (2004: 364)

For instance, there is a rational commitment to believe that  $q$  if one believes that  $p$  and if  $p$ , then  $q$ : in other words, it is irrational to believe that  $p$  and if  $p$ , then  $q$ , while disbelieving or withholding belief that  $q$ . What explains this rational commitment is the fact that one has justification to believe that  $q$  if one has justification to believe that  $p$  and if  $p$ , then  $q$ . Similarly, there is a rational commitment to believe that  $p$  if and only if one believes that one has justification to believe that  $p$ : in other words, it is irrational to believe either the proposition that  $p$ , or the proposition that one has justification to believe that  $p$ , while disbelieving or withholding belief in the other. But what explains this rational commitment? The answer is that justification is accessible in the sense that one has justification to believe that  $p$  if and only if one has justification to believe that one does.

On this view, one cannot explain the rational commitments that are violated in believing Moorean conjunctions without invoking the accessibility thesis. Hence, there is a challenge for opponents of the accessibility thesis to give some alternative explanation of the relevant kind of Moorean irrationality. In the absence of any satisfactory alter-

native, I draw the provisional conclusion that the accessibility thesis is indispensable for explaining the Moorean irrationality in question.

### 5. Reliabilism and Clairvoyance

Having used an epistemic version of Moore's paradox in arguing for the accessibility of justification, I want now to consider how this bears on the classic debate between Laurence Bonjour (1985) and Alvin Goldman (1986) over the viability of an externalist version of reliabilism.

Let us begin with *strong reliabilism*—that is, the view that reliability is both necessary and sufficient for justification. A doxastic process is *reliable* if and only if it is disposed to yield a sufficiently high ratio of true to false beliefs in sufficiently similar circumstances. According to strong reliabilism, one's belief is justified if and only if it is based on a reliable doxastic process and, similarly, one has justification to believe a proposition if and only if one is disposed to believe that proposition on the basis of a reliable doxastic process.

Strong reliabilism is incompatible with the accessibility of justification, since one does not always have epistemic access to the reliability or unreliability of one's doxastic processes. This point is illustrated by many of the intuitive counterexamples to strong reliabilism, including Bonjour's example of the clairvoyant subject, Maud:

Maud believes herself to have the power of clairvoyance, though she has no reasons for this belief. She maintains her belief despite being inundated by embarrassed friends and relatives with massive quantities of apparently cogent scientific evidence that no such power is possible. One day Maud comes to believe, for no apparent reason, that that the President is in New York City, and she maintains this belief despite the lack of any independent evidence, appealing to her alleged clairvoyant power. Now in fact the President is in New York City, and Maud does, under the conditions then satisfied, have completely reliable clairvoyant power. Moreover, her belief about the President did result from the operation of that power. (1985: 40)

If strong reliabilism is true, then Maud has justification to believe that the President is in New York City on the basis of her reliable clairvoyant power. However, she has scientific evidence against the possibility of reliable clairvoyance and so she has justification to believe that her clairvoyant belief is unreliably formed and so lacks justification. Therefore, Maud has justification to believe a proposition, while also having meta-level justification to believe that she lacks justification to believe it, which violates the principles of positive self-intimation and negative infallibility.

Intuitively, Maud's clairvoyant beliefs are unjustified, despite the fact that they are reliably formed. At any rate, this is BonJour's intuitive reaction to the case. But what if one disagrees or simply bites the bullet? Is there anything more to be said in defence of this verdict besides an appeal to the brute deliverances of intuition? At this point, the resources of this paper can be brought into play. On the one hand, if we assume that justification is accessible, then we can argue from the premise that Maud has meta-level justification to believe that she lacks justification for her clairvoyant beliefs to the conclusion that those clairvoyant beliefs are unjustified. On the other hand, if we assume that justification is not accessible, then we are faced with an epistemic version of Moore's paradox. Suppose that Maud has justification to believe that the President is in New York City, although she also has justification to believe that she does not have justification to believe this. In that case, she has justification to believe the Moorean conjunction: the President is in New York City, but I do not have justification to believe that the President is in New York City. Clearly, however, it is irrational to believe this Moorean conjunction. Therefore, proponents of strong reliabilism who deny the accessibility of justification are faced with an epistemic version of Moore's paradox.

Alvin Goldman (1986) pursues a different approach. Rather than biting the bullet on BonJour's example, he accepts BonJour's verdict and tries to accommodate it within a reliabilist theory of justification. However, Goldman opts for a *weak* version of reliabilism on which reliability is necessary but not sufficient for justification.<sup>20</sup> Goldman's proposal is that a belief is justified if and only if it is reliably formed and it also satisfies a *no-undermining* condition, which states that a belief is unjustified if one believes, or if one has justification to believe, that one lacks justification for the belief in question:

Suppose S's belief in p is permitted by a right J-rule system, but S believes that it is not so permitted. Alternatively, suppose S's belief in p is permitted by a right J-rule system and S is *justified* in believing that it is not so permitted (whether or not he actually believes that it is not permitted). In either case, it is counterintuitive to regard S's belief in p as justified. In these cases S's belief, or justification for believing, that the belief is not permitted *undermines* its permittedness. That is, although the belief *is* permitted, this does not intuitively suffice for its justifiedness. (1986: 62)

---

<sup>20</sup> Goldman (1986: 113) argues in response to the new evil demon case that justification does not require reliability in the actual world, but reliability in normal worlds, which are worlds that are consistent with our general beliefs about the actual world. However, this response raises issues that are orthogonal to my present concerns and so I will leave it aside.

Goldman's no-undermining condition entails the negative infallibility thesis, which states that if one has justification to believe that one *lacks* justification to believe that  $p$ , then one lacks justification to believe that  $p$  ( $J\sim Jp \rightarrow \sim Jp$ ). Moreover, if one accepts the negative infallibility thesis, then it is also very natural to accept the positive infallibility thesis, which states that if one has justification to believe that one *has* justification to believe that  $p$ , then one *has* justification to believe that  $p$  ( $JJp \rightarrow Jp$ ). Otherwise, what accounts for the asymmetry? Nevertheless, Goldman's position is distinctively externalist insofar as he rejects both positive and negative versions of the self-intimation thesis. Recently, BonJour (2003: 32) has expressed doubts about "whether accepting part of the internalist requirement for justification while rejecting the other part does not amount to an untenable halfway house." In what follows, I will attempt to substantiate these doubts by explaining why Goldman's position fails to provide a stable and well motivated form of externalism.

The motivation for the no-undermining condition is that it generates the intuitive verdict that Maud's clairvoyant beliefs are unjustified. Since she has justification to believe that her clairvoyant beliefs are unreliable and so unjustified, this undermines any justification that would otherwise have accrued in virtue of their *de facto* reliability. However, this response does not extend to BonJour's example of the clairvoyant, Norman, who is unlike Maud in that he does not have justification to believe that his clairvoyant beliefs are either reliable or unreliable:

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)

Norman has no evidence either for or against the hypothesis that he has a reliable clairvoyant power. So, unlike Maud, he has justification to withhold belief about whether or not he has justification for his clairvoyant beliefs. Nevertheless, there is a strong case to be made that Norman lacks justification for his clairvoyant beliefs. If we assume that justification is accessible, then we can argue from the premise that Norman lacks meta-level justification to believe that he has justification for his clairvoyant beliefs to the conclusion that his clairvoyant beliefs are

unjustified. And if we assume that justification is not accessible, then we are faced with an epistemic version of Moore's paradox. Suppose that Norman has justification to believe that the President is in New York City, although he lacks justification to believe that he does. In that case, he has justification to believe the Moorean conjunction: the President is in New York City, but it is an open question whether I have justification to believe that the President is in New York City. Clearly, however, it is irrational to believe this Moorean conjunction. Therefore, if Goldman denies the accessibility of justification, then he is faced with an epistemic version of Moore's paradox.

Goldman's response is to dispute BonJour's description of the case:

BonJour describes this case as one in which Norman possesses no evidence or reasons of any kind for or against the general possibility of clairvoyance, or for or against the thesis that he possesses it. But it is hard to envisage this description holding. Norman *ought* to reason along the following lines: 'If I had a clairvoyant power, I would surely find *some* evidence for this. I would find myself believing things in otherwise inexplicable ways, and when these things were checked by otherwise reliable processes, they would usually check out positively. Since I lack any such signs, I apparently do not possess reliable clairvoyant processes.' Since Norman ought to reason in this way, he is *ex ante* justified in believing that he does not possess reliable clairvoyant processes. This undermines his belief... (1986: 112)

On the face of it, Goldman's strategy seems unpromising. As BonJour describes the case, Norman's clairvoyant power is domain specific in the sense that it relates only to certain kinds of subject matter, such as the whereabouts of the President. Presumably, then, we can set up the case in such a way that the opportunity to use this clairvoyant power has simply never arisen before. In that case, Norman will not have had any opportunity to reason along the lines that Goldman proposes, by reflecting on the previous operations of his clairvoyant power and checking its deliverances for coherence with his other beliefs. In other words, we can stipulate that this is a case of *justified agnosticism*, in which Norman has justification to believe that the President is in New York City, although he lacks justification to form a belief one way or the other about whether or not he has justification to believe this ( $Jp \ \& \ \sim Jp \ \& \ \sim Jp$ ).

Now, Goldman is faced with a dilemma. If it is possible to stipulate a case of justified agnosticism, then Goldman's position delivers a counter-intuitive verdict on the case and he is thereby confronted with an epistemic version of Moore's paradox. Goldman must therefore deny that it is possible to stipulate any such case of justified agnosticism. In other words, he is committed to a ban on justified agnosticism, which says that if one *lacks* justification to believe that one *has* justification to believe

that  $p$ , then one *has* justification to believe that one *lacks* justification to believe that  $p$  ( $\sim JJp \rightarrow J\sim Jp$ ). In that case, however, his attempt to rescue an externalist version of reliabilism collapses into a form of internalism. If we contrapose negative infallibility ( $Jp \rightarrow \sim J\sim Jp$ ) and the ban on justified agnosticism ( $\sim J\sim Jp \rightarrow JJp$ ), then we can derive positive self-intimation ( $Jp \rightarrow JJp$ ). Similarly, if we contrapose positive infallibility ( $\sim Jp \rightarrow \sim JJp$ ), then from the ban on justified agnosticism ( $\sim JJp \rightarrow J\sim Jp$ ), we can derive negative self-intimation ( $\sim Jp \rightarrow J\sim Jp$ ).

To sum up, if Goldman denies the ban on justified agnosticism, then his weak form of reliabilism generates counterintuitive verdicts in cases of justified agnosticism and confronts him with an epistemic version of Moore's paradox. In that case, his position is unmotivated, since it faces all the same problems as strong reliabilism, which it was designed to avoid. On the other hand, if he accepts the ban on justified agnosticism, then he must either reject the infallibility thesis, in which case his position collapses into strong reliabilism, or he must accept the self-intimation thesis, in which case it collapses into internalism. Either way, Goldman's weak reliabilism fails to locate a stable and well motivated intermediate position between internalism on the one hand and strong reliabilism on the other.

Strong and weak reliabilism are certainly not the only options available for proponents of an externalist theory of justification.<sup>21</sup> Nevertheless, the arguments of this paper have more general repercussions, since they establish that any externalist theory is faced with an epistemic version of Moore's paradox. The more general moral to be drawn from this paper is that the epistemic version of Moore's paradox cannot be avoided except by taking on an internalist commitment to the accessibility of justification.

## 6. Explaining Accessibility

I want to conclude by raising a question that emerges from the arguments of this paper, which requires a more extended discussion elsewhere. The main aim of this paper has been to argue for the accessibility of justification on the grounds that it is indispensable for solving an epistemic version of Moore's paradox. However, it is one thing to argue that justification is accessible, but it is another thing to explain what *makes it the case* that justification is accessible. If justification is accessible, then part of the job description for a theory of justification is to give an account of the determinants of justification, which explains why justification is accessible.

---

<sup>21</sup> For a range of other externalist options, see Bergmann (2006), Goldman (1993), Sosa (2007), and Williamson (2000).



After all, epistemic facts about which propositions one has justification to believe are not brute facts. They *supervene* on non-epistemic facts in the sense that there cannot be epistemic differences between situations unless there are some non-epistemic differences between them. Moreover, they are *determined* by non-epistemic facts in the sense that there cannot be epistemic differences between situations unless they are explained by corresponding non-epistemic differences between them. The *determinants* of justification, then, are the non-epistemic facts that determine the epistemic facts about which propositions one has justification to believe.

One of the main tasks for a theory of justification is to specify the determinants of justification. According to reliabilism, for instance, the determinants of justification are non-epistemic facts about the reliability of one's doxastic processes. However, the accessibility of justification imposes substantial constraints on an account of the determinants of justification, since justification is accessible only if the determinants of justification are themselves accessible. For instance, facts about the reliability of one's doxastic mechanisms are not accessible in the sense that one has justification to believe that they are reliable if and only if they are in fact reliable. So, if justification is accessible, then facts about the reliability of one's doxastic mechanisms cannot be among the determinants of justification. But then what *are* the determinants of justification? Which facts, if any, are accessible in the sense that one has justification to believe that they obtain if and only if they obtain?

Here is a sketch of an answer, which I develop in more detail elsewhere.<sup>22</sup> The determinants of justification are non-epistemic facts about one's mental states, which are introspectively accessible in the sense that one has introspective justification to believe that those mental facts obtain if and only if they obtain. The argument for this claim is by inference to the best (indeed, the only) explanation. If the determinants of justification are introspectively accessible mental states and if the determination relation is accessible by a priori reflection, then the facts about which propositions one has justification to believe are accessible by means of a combination of introspection and a priori reflection. If not, then it is quite obscure what else could explain the accessibility of justification. Therefore, the accessibility of justification provides the basis of an argument for a version of mentalism, on which the determinants of justification are facts about one's introspectively accessible mental states. Of course, it is a further question *which* of one's mental states are introspectively accessible in the relevant sense.

---

<sup>22</sup> See Smithies (forthcoming, c).

According to John McDowell (1995) and Timothy Williamson (2000: Ch.9), the determinants of justification include *factive* mental states, such as seeing that  $p$ , hearing that  $p$  and remembering that  $p$ , which are determinate ways of knowing that  $p$ . However, these factive mental states are not introspectively accessible in the sense that one has introspective justification to believe that those mental states obtain if and only if they obtain. For instance, it is implausible to suppose that one has introspective justification to believe that one sees that  $p$  if and only if one sees that  $p$ . One might argue that if one sees that  $p$ , then one has introspective justification to believe that one sees that  $p$ . But suppose one suffers an illusion or hallucination in which one does not see that  $p$ , but merely seems to see that  $p$ . It seems incredible to suppose that, in such a case, one has introspective justification to believe that one does not see that  $p$ , but merely seems to see that  $p$ . After all, this would be to deprive skeptical scenarios of their intuitive power. But if factive mental states are not introspectively accessible, then according to the line of argument sketched above, they cannot be among the determinants of justification.

Various philosophers, including Conee and Feldman (2001) and Ralph Wedgwood (2002), have argued for a non-factive version of mentalism on which one's non-factive mental states determine which propositions one has justification to believe. However, these philosophers have eschewed any commitment to the accessibility of justification. Nevertheless, if the accessibility of justification can be independently motivated by the kinds of arguments given in this paper, then it provides the basis for a new and perhaps even more compelling line of argument for a non-factive version of mentalism. The argument, in brief, is that justification is accessible only if it is determined by one's non-factive mental states.

This paper began with a rough characterization of internalism as the thesis that one has a special kind of epistemic access to facts about which propositions one has justification to believe. Following Alex Byrne (2005), however, we can distinguish the claim that one's epistemic access is *privileged*, in the sense that it is better than other modes of epistemic access, from the claim that it is *peculiar*, in the sense that it is different from other modes of epistemic access. My main aim in this paper has been to argue that our epistemic access to facts about justification is privileged in the sense that it is infallible and self-intimating. In this section, however, I have suggested that our epistemic access is peculiar in the sense that it has its source in our introspective access to the mental states that are the determinants of justification. Moreover, I have suggested that our privileged access to facts about justification is

best explained by the hypothesis that it is peculiar in just this way. I hope to develop this suggestion in more detail elsewhere.<sup>23</sup>

### References

- Alston, William (1985) "Concepts of Epistemic Justification" *Monist* 68: 57–89.
- (1988) "The Deontological Conception of Epistemic Justification" *Philosophical Perspectives* 2: 257–99.
- Bergmann, Michael (2005) "Defeaters and Higher Level Requirements" *The Philosophical Quarterly* 55: 419–436.
- (2006) *Justification Without Awareness: A Defense of Epistemic Externalism*, Oxford University Press.
- BonJour, Laurence (1985) *The Structure of Empirical Knowledge*, Harvard University Press.
- BonJour, Laurence and Sosa, Ernest (2003) *Epistemic Justification: Internalism vs. Externalism, Foundations vs. Virtues*, Blackwell Publishing.
- Broome, John (1999) "Normative Requirements" *Ratio* 12: 398–419.
- Byrne, Alex (2005) "Introspection" *Philosophical Topics* 33 (1): 79–104.
- Chalmers, David and Hájek, Alan (2007) "Ramsey + Moore = God" *Analysis* 67: 170–2.
- Christensen, David (2004) *Putting Logic In Its Place: Formal Constraints on Rational Belief*, Oxford University Press.
- Conee, Earl and Feldman, Richard (2001) "Internalism Defended" *American Philosophical Quarterly* 38 (1): 1–18.
- Feldman, Richard and Conee, Earl (1985) "Evidentialism" *Philosophical Studies* 48 (1): 15–34.
- (2000) "The Ethics of Belief" *Philosophy and Phenomenological Research* 60 (3): 667–695.
- (2007) "Reasonable Religious Disagreements" in (ed.) L. Anthony, *Philosophers without Gods: Meditations on Atheism and the Secular*, Oxford University Press.
- Firth, Roderick (1978) "Are Epistemic Concepts Reducible to Ethical Concepts?" in (eds.) A. Goldman and J. Kim, *Values and Morals*, Dordrecht: Kluwer.
- Goldman, Alvin (1979) "What is Justified Belief?" in (ed.) G. Pappas, *Justification and Knowledge*, Dordrecht: Reidel.

---

<sup>23</sup> Earlier versions of this paper were presented in 2008 and 2009 at the Australian National University, the Ohio State University, the Australasian Association of Philosophy, and the Bled Conference on Epistemic Value and Virtue. This paper has been much improved by questions from the audience on these occasions and by helpful feedback from Patrick Greenough, Alan Hájek, Ram Neta, George Pappas, Nicholas Silins, Ralph Wedgwood, John Williams and an anonymous referee.

- (1986) *Epistemology and Cognition*, Harvard University Press.
- (1993) “Epistemic Folkways and Scientific Epistemology” *Philosophical Issues* 3: 271–85.
- Green, Mitchell and Williams, John (2007) *Moore’s Paradox: New Essays on Belief, Rationality, and the First Person*, Oxford University Press.
- Harman, Gilbert (1986) *Change in View: Principles of Reasoning*, MIT Press.
- Huemer, Michael (2007) “Moore’s Paradox and the Norm of Belief” in (eds.) S. Nuccetelli and G. Seay, *Themes from G. E. Moore: New Essays in Epistemology and Ethics*, Oxford University Press.
- Martin, Charles (1994) “Dispositions and Conditionals” *Philosophical Quarterly* 44: 1–8.
- McDowell, John (1995) “Knowledge and the Internal” *Philosophy and Phenomenological Research* 55: 877–93.
- Pryor, James (2001) “Highlights of Recent Epistemology” *British Journal for the Philosophy of Science* 52: 95–124.
- (2004) “What’s Wrong With Moore’s Argument?” *Philosophical Issues* 14: 349–78.
- Shoemaker, Sydney (1995) “Moore’s Paradox and Self-Knowledge” *Philosophical Studies* 77: 211–28.
- REPRINTED IN SHOEMAKER (1996) *The first person perspective and other essays*, Cambridge University Press.
- Smithies, Declan (forthcoming, a) “Mentalism and Epistemic Transparency”
- (forthcoming, b) “Why Care About Justification?”
- (forthcoming, c) “A Simple Theory of Introspection” in (eds.) D. Smithies and D. Stoljar, *Introspection and Consciousness*, Oxford University Press.
- Sosa, Ernest (2007) *A Virtue Epistemology: Apt Belief and Reflective Knowledge*, Clarendon Press : Oxford.
- Wedgwood, Ralph (2002) “Internalism Explained” *Philosophy and Phenomenological Research* 65: 349–369.
- White, Roger (2005) “Epistemic Permissiveness” *Philosophical Perspectives* 19 (1): 445–59.
- Williamson, Timothy (2000) *Knowledge and Its Limits*, Oxford University Press.
- Wright, Crispin (2004) “Warrant for Nothing (and Foundations for Free)?” *Aristotelian Society Supplementary Volume* 78 (1): 167–212.