Knowledge and Attributability

Cameron Boult

Abstract

A prominent objection to the so-called 'knowledge norm of belief' is that it is too demanding or too strong. I argue that a prominent way of motivating the objection leads to an impasse in the epistemic norms debate. A way out of the impasse becomes available when we take a closer look at some distinctions in the theory of *responsibility*. There are at least three relevant notions of responsibility. I argue that a weaker notion of responsibility – attributability – should be used to motivate the objection. This opens up space to move beyond the impasse.

Introduction

Call a belief epistemically justified if and only if it complies with the 'epistemic norm of belief.' A currently popular view is that the epistemic norm of belief is knowledge. This is the view that it is epistemically permissible to believe that p only if one knows that p. According to proponents of the knowledge norm of belief, one's belief that p is epistemically justified if and only if one *knows* that p.

The view has attractive features. For instance, while a competing 'truth norm' readily captures the idea that epistemically justified belief has something to do with truth, it seems too weak.³ For example, accidentally true beliefs comply with it. If someone merely happens to form a true belief by way of wishful thinking, they do not thereby form an epistemically justified belief.⁴ The knowledge norm also readily captures epistemic justification's connection with truth—since knowledge is factive—but accidentally true beliefs do not comply with it. This is because knowledge comes with additional conditions. For example, it comes with conditions linking cognitive abilities of the agent to the agent's true beliefs. The knowledge norm makes plausible demands on the agent in addition to requiring truth.

Of course, the knowledge norm has detractors. One important worry is that it is in some sense too demanding or too strong (a worry that also applies to the truth norm).⁵ The idea is that knowing that p is not necessary for being justified in believing that p. For example, consider a couple of well-known cases:

GETTIER: Ed comes downstairs in the morning and checks the time by looking at a clock that he knows to be generally very reliable but which happened to stop exactly 24 hours ago. He forms the (true) belief that it's 8am. But he doesn't know that it's 8am.

BIV: Dave has recently been envatted. He is the internal duplicate of an ordinary, epistemically responsible person. He continues to believe he has hands because it looks to him just like he has hands. But he doesn't know that he has hands.

According to proponents of the knowledge norm, these agents violate the norm of belief. As such, their beliefs are not epistemically justified. But GETTIER and BIV are often appealed to as vivid examples of cases in which we feel compelled to call a given belief justified, despite other epistemic shortcomings. Traditional reliabilists (Goldman 1975; Dretske 1971, 1981), recent reliabilist virtue epistemologists (Sosa 2011, 2015; Greco 2010; Pritchard 2012), and proper functionalists (Burge 2003; Graham 2012; Plantinga 1993) all appeal to the idea that agents in such cases have justified (or 'warranted') beliefs. They take this as a kind of datum to account for in their theories.

So, from the perspective of much mainstream epistemology, the knowledge norm of belief is too demanding or too strong. I refer to this worry a lot in this paper. So I'll simply call it (TS). Interestingly, aside from appeals to intuitions about justification, there is no clear sense in the literature of the best way to motivate (TS). Why, exactly, does it seem wrong not to attribute justified beliefs in these cases? As I will explain below, there are a variety of explanations available. But I think, in the background at least, a very general approach involves appealing to what I will call a *blamelessness principle*. Very roughly, the basic idea is that, because agents in GETTIER and BIV are perfectly blameless for believing what they do, the knowledge norm of belief imposes overly strong demands in counting them unjustified.

In this paper, I do two main things. First, I argue that ways of motivating (TS) that appeal to the blamelessness principle lead to an *impasse* between proponents of the knowledge norm and proponents of weaker views. They lead to an impasse when knowledge normers invoke the notion of an *excuse* to explain away a prima facie connection between blamelessness and justified belief (Littlejohn *forthcoming*; Williamson *forthcoming*; see also DeRose 2002; Lackey 2007; Gerken 2011; McGlynn 2014). Second, I argue that a way out of this impasse becomes available when we take a closer look at some distinctions in the theory of *responsibility*. It is very common in epistemology to find responsibility inseparably linked to the notion of being *open to blame responses*. But I think this is a mistake. A closer look at some recent work on moral responsibility reveals as much. Indeed, I will argue that there are at least three different notions of responsibility relevant to present purposes. Following David

Shoemaker (2011), I call one particularly weak notion of responsibility 'attributability-responsibility', or 'attributability' for short. As I will explain, being attributability-responsible comes apart from being open to blame responses. For this reason, attributability is interesting when it comes to motivating (TS). In a nutshell, I will argue that there is good (though defeasible) reason to doubt that a norm has been violated in a given case if the agent in the case is not attributability-responsible for doing what the putative norm prohibits. I will argue that, since cases like BIV and GETTIER are cases in which agents are not attributability responsible for doing what the knowledge norm prohibits (i.e. believing that p when they don't know that p), the knowledge norm is too demanding or too strong.

An upshot of the discussion is that we can avoid an impasse stifling the debate about the knowledge norm of belief (and, as I'll explain briefly below, I think the upshot has more general implications, too). Rather than a story about how agents may violate a norm blamelessly (as proponents of the knowledge norm use excuses to do), we need a story about how agents may do so without being attributability-responsible for doing whatever it is that the putative norm prohibits. A story about excuses simply doesn't help with that. As such, the proposal presents knowledge normers with a new challenge—one that opens up space for proponents of weaker norms of belief to move beyond the impasse.

The plan is as follows. In Section One I sharpen the basic idea that there is a tight connection between norm violation and the appropriateness of criticism or blame. I argue that motivating (TS) with this idea leads to an impasse between knowledge normers and proponents of weaker norms of belief. In Section Two, I lay the foundations of a finergrained conception of responsibility. In Section Three, I show how this leads to a notion of responsibility—attributability—which is particularly promising in the context of motivating (TS). In Sections Four and Five, I defend the claim that attributability is tightly connected to norm violation (and specify what that means). In Section Six, I explain why this poses a problem for the knowledge norm of belief. Since the problem can't be resolved by appealing to excuses, I argue that the proposal (which I call the 'Attributability View') opens up space to move beyond the impasse. Section Seven concludes.

I should emphasize that my basic issue arises not just in the context of the norm of belief. In epistemology it also comes up frequently in discussions about assertion and practical reasoning, for example. I'm interested in the issue as it crops up in all of these discussions. I focus mainly on the norm of belief for ease of exposition. I intend what I say here to have very general implications (and will sometimes use examples from the norm of assertion debate in illustrating certain points).

1. Blamelessness and Norm Violation

It will be helpful to start with a look at some examples in the literature of different ways of motivating (TS). We see appeals to reasonableness, rationality, virtue, proper use of evidence—sometimes all in the same breath. We also see appeals to *blamelessness*. As I'll explain below, I think there is a core of agreement when it comes to the role that blamelessness ought to play in our understanding of (TS), which is why I focus on it.

Here is an instructive quote from Robert Audi, in which we see an appeal to virtue (or skill), proper use of evidence, blamelessness (or non-criticizability), and reasonableness (Audi is discussing a case of vivid hallucination; we can set the differences between this sort of case and both BIV and GETTIER for present purposes):

[G]iven the vivid hallucination, I am in no way at fault for believing what I do, nor do I deserve any criticism (at least on the non-skeptical assumption that we may generally trust our senses in this way). Far from it. I am like a surgeon who skillfully does all that can be expected but loses the patient. There I should feel regret, but not guilt; I should explain, but need not apologize; and when we know what my evidence was, we approve of what I did. We consider it reasonable (Audi 2001, p.23).

And here is a much-cited quote from Cohen's seminal paper, in which he introduces the so-called 'New Evil Demon' problem:⁶

My argument hinges on viewing justification as a normative notion. Intuitively, if S's belief is appropriate to the available evidence, he is not to be held responsible for circumstances beyond his ken (Cohen 1984, p. 282).

There are important differences between what each of these authors is getting at. But I also think they are unified in a theoretically interesting way. The basic unifying idea is that the agents in GETTIER and BIV are blameless for believing what they do. This is (at least part of) what makes them reasonable, rational, or virtuous. The idea continues: there is some kind of important relationship between blamelessness and the question of whether a norm has been violated. The knowledge norm tells us that certain important cases of blameless believing are cases of norm violation. And that seems too demanding or too strong.

A bit more precisely, we can put the idea in the form of a 'blamelessness principle' (B):

(B): There is good but defeasible reason to doubt that a norm N has been violated by an agent S if S is blameless for doing what N prohibits.

We can also find plenty of reference more directly to something like (B) in the literature. For example, in her discussion of DeRose's contextualist approach to the knowledge

norm of assertion, Jennifer Lackey says, 'I shall begin by taking seriously the widely accepted thesis that there is a very intimate connection between being deserving of criticism and violating a norm' (2007, p.603). Amia Srinivasan also notes, 'Many think that as a conceptual matter, there is an alignment between the facts about what we ought to do (the deontic facts) and the facts about what we are responsible or blameworthy for doing (the hypological facts)' (Srinivasan 2015, p.9). These quotes highlight the apparent prominence of (B), but they don't tell us much about how to connect it to (TS). Douven (2006) makes something like an appeal to (B) in order to defend his weaker 'rational credibility' norm against the knowledge norm of assertion:

[O]n the knowledge account the asserter has breached a rule, however good her evidence for what she asserted, given that she asserted something she did not know. And if breaching a rule makes one blameworthy, which typically it does, then, on the knowledge account, the asserter comes out as being blameworthy, contrary to intuition. As a result, it seems that false but intuitively reasonable assertions are perfectly handled by the rational credibility account but pose a problem for the knowledge account (2006, p.477).

Of course, the knowledge normer can try to account for (B) without giving up their view. Indeed, this is exactly what the prominent appeal to *excuses* is designed to do. It is available to proponents of the knowledge norm to agree that the BIV is not blameworthy, yet to maintain the BIV violates the norm of belief. He does so excusably. For example, perhaps the fact that the BIV does precisely what an ordinarily reliable cognizer would do in the situation means the BIV has an excuse, rather than a justification (Williamson *forthcoming*). Or perhaps—contrary to what many seem to find obvious—rationality and reasonableness are the mark of the excusable as opposed to the justified (Littlejohn *forthcoming*, p.15).

This calls for a worked out explanation or theory of what excuses specific agents under what circumstances. And there have been multiple attempts at this recently (Littlejohn *forthcoming*; Williamson *forthcoming*; see also DeRose 2002; Lackey 2007; Gerken 2011). We should then compare this sort of explanation with the presumably simpler explanation that the norm of belief has not been violated. Perhaps the most well known worry about the appeal to excuses is that it seems like an *ad hoc* way of saving the knowledge norm from an otherwise compelling objection (Gerken 2011). But to assess this worry, we'd need to look at the details of some particular theory of excuses. We'd need to assess whether it fits within a well-motivated framework that gives an inprinciple way of drawing the justification-excuse distinction in particular cases. Unfortunately, as Aidan McGlynn (2014) makes clear, this has led to an impasse in the epistemic norms debates:

Do the asserters in our problem cases require excuses? That also seems likely to prove a vexed question, and it is not clear how to answer it. There's a danger that the debate will descend into a mere clash of intuitions at this juncture, with defenders of the knowledge norm citing it as a virtue of their account that it "explains" the need for an excuse in such cases, and the norm's critics claiming this as a clear defect (McGlynn 2014, p. 114).

In other words, it's just not clear what one can say to get beyond this disagreement. It is not clear exactly what we ought to be challenging about our opponent's view, other than that it clashes with our intuitions.

Rather than address that impasse head on, I want to consider a different approach. This approach promises a way out of the impasse—if not because it's right, then because it puts more clearly on the table what each party to the debate needs to defend or reject in order to defend their position. As I've said, something that has not been considered in much detail, but which is surely important, is what sort of notion of responsibility is at play in this general way of motivating (TS) in the first place. Blameworthiness presupposes or implies responsibility. We only blame agents we hold responsible in some way or another. But agents can also be held responsible in a way that does not require being the appropriate subject of blame.

2. Responsibility: A Useful Framework

An influential idea in contemporary moral philosophy is to understand responsibility broadly in terms of being open to a particular range of moral responses. Peter Strawson's seminal 'Freedom and Resentment' (1962) focused on reactive attitudes such as gratitude, guilt, indignation and resentment. Reactive attitudes such as indignation and resentment are precisely what we normally think of as blame-responses. Perhaps the most natural way of explaining what it is to blame someone for Φ -ing is in terms of having or expressing feelings of resentment, indignation, or issuing other forms of sanction. Thus, were Strawson right about the sorts of moral responses that essentially constitute what it is to be morally responsible, the project I am proposing would seem to be a nonstarter. That is, motivating (TS) in terms of the connection between norm violation and some notion of responsibility that does not amount to being the appropriate target of blameresponses wouldn't really make sense. However, there are well-known issues with the Strawsonian account. Perhaps most crucially, it is plausible that agents can hold other agents responsible for Φ-ing without harboring feelings of resentment or indignation. For example, imagine someone like Gandhi holding you morally responsible for some unjust action. Being Gandhi, he fails to feel resentment or indignation. Nevertheless, it seems possible for Gandhi to hold you morally responsible for your action.

What else might we mean by responsibility within this broad framework? Recent work by Tim Scanlon (1998; 2008) and Angela Smith (2005; 2012) continues a broadly

Strawsonian approach, but with a subtler view of the types of moral responses that are relevant to responsibility. It is helpful to articulate the view by starting with the idea that being responsible for an action or attitude implies a special kind of activity. This is why we do not hold each other morally responsible for our heights, and why we do not hold avalanches morally responsible for destroying villages. A key desideratum for any theory of responsibility is accounting for the intuitive idea that responsibility for actions and attitudes implies a special kind of activity. What are sometimes known as 'volitional views', for example, account for this in terms of choice, identification, decision, or voluntary control. An agent is responsible for Φ -ing only if the agent made/is able to make some kind of past, present or future choice concerning their Φ -ing. Rather than accounting for the activity involved in responsible actions and attitudes in these terms, the Smith/Scanlon view focuses on what Smith calls 'rational activity'.

I will explain what Smith means by 'rational activity' right away. But first, one of the primary motivations for the Smith/Scanlon view is that it can account for a comprehensive range of moral responses involved in our moral practices. For example, contrasting the view with the volitional view of responsibility, we can more readily account for our practice of holding each other responsible for attitudes and actions over which it seems we have no immediate voluntary control—like forgetting one's best friend's birthday, or failing to *notice* something. It is a familiar feature of our moral and epistemic practices that people are open to moral and epistemic assessment for their impulses and spontaneous reactions—indeed, we often take a person's spontaneous reactions to reveal *deeper* moral and epistemic features of the person than the considered and controlled behavior they manifest. And while volitional views of responsibility have a hard time accounting for this, the Scanlon/Smith view is well positioned to do so. ¹¹ Let me briefly outline the basic idea.

According to Smith, moral agents have a set of 'evaluative commitments.' Evaluative commitments are not always, or even very often, consciously held propositional beliefs. Rather, they are tendencies to treat certain things as having evaluative significance. As such, they are often things we discover about ourselves through our responses to situations, as opposed to things we consciously choose. To use Smith's example, I may discover in some situation that I care more about being liked by others than I do about standing up for my moral principles' (Smith 2005, p. 252). On the Smith/Scanlon view, our evaluative commitments—'the basic evaluative framework through which we view the world'—are the core of what makes us open to moral assessment for our actions and attitudes. This is because, according to the view, moral assessment of a person's actions or attitudes is a kind of demand that that person 'acknowledge and defend or disavow the judgments that are implicit in her responses to the world around her' (Smith 2005, p.256). Evaluative commitments are susceptible to the sort of critical reflection and requests for justification implicit in our practice of moral assessment. As such, on the Smith/Scanlon view, a person's evaluative commitments

(rather than their conscious choices, or what they identify with, or what they have voluntary control over) are the seat of their status as a responsible agent. But how does this tell us anything about our responsibility for particular actions and attitudes, such as my belief that the U.S. needs more gun control, or my fear of flying?

The idea is that our particular actions and attitudes stand in a *rational relation* to our evaluative commitments. That is to say, most of the time, our actions and attitudes rationally follow from, or fail to follow from, our evaluative commitments (Smith 2005). For example, if I judge that flying is perfectly safe, but do everything in my power to avoid flying when flying is something I need to do, very roughly we might say that my acts of avoiding flying fail to follow from my judgment that flying is perfectly safe. As Smith puts it, the relation is a normative one: given what my judgments about flying are, other things being equal, I should not be avoiding it at all costs. The fact that we usually call this kind of behavior *ir*rational is one way of bringing out the general idea of a rational connection between particular attitudes or actions and an agent's evaluative commitments.

Importantly, we hold each other morally responsible in virtue of substantive critical engagement with our evaluative commitments. For example, take Jim's attitude of hatred towards red haired people. According to the Smith/Scanlon view, Jim's hatred of red haired people is open to moral assessment, not because it is freely chosen, or because it is something he identifies with, or because it is something under his voluntary control (though all of these things may be true); rather, Jim's hatred is open to moral assessment because it stands in a direct rational relation to his evaluative commitments: for example, we can imagine that (whether he knows it or not) Jim thinks that a person's worth is determined by their intelligence, and that a person's intelligence is determined by their hair color. These evaluative judgments are substantively mistaken. Jim is thus open to moral criticism, not just for these evaluative commitments, but for the actions and attitudes that are rationally connected to them.

Summing up, according to the Smith/Scanlon view, we are responsible for our actions and attitudes in virtue of (and only in virtue of) the fact that they bear a rational connection to our evaluative commitments. When someone Φ s, they can in principle be called upon to justify the particular evaluative commitments that their Φ -ing implies. This is how the view accounts for the intuitive idea that responsibility requires a certain kind of activity. It requires the in-principle capacity to reflect upon and justify one's evaluative commitments. For this reason, we can further specify this notion of responsibility as a kind of *answerability* (which will be useful below). It is a view that fits within the Strawsonian picture of thinking of moral responsibility in terms of being open to a certain range of moral responses. It is also a view that contains important theoretical resources that I will expand on below.

3. From Answerability to Attributability

An important recent criticism of the Smith/Scanlon view, due to Shoemaker (2011), opens up useful conceptual territory for present purposes. Shoemaker argues (among other things) that the implication of the view that responsibility is connected to agents' abilities to justify their evaluative commitments makes the view too strong. Shoemaker distinguishes between *answerability* (a notion of responsibility that entails the ability to justify) and mere *attributability* (a notion of responsibility that does not entail the ability to justify). All it takes for an agent to be attributability-responsible for an action or attitude is that it in some sense expresses the agent's evaluative commitments. This can be the case even when an agent cannot reasonably be expected to justify anything.¹³

Defending his claim, Shoemaker appeals to cases of 'non-rational caring'. Imagine that I am a parent who continues to care for my son, even after discovering that he has become a serial killer. Or imagine someone pining for their ex-lover, despite the fact that their ex-lover was physically abusive. Shoemaker suggests that these sorts of cases are examples of caring merely in virtue of standing in a sort of 'mine-ness' relation to something or someone. He claims that requesting reasons in justification for an agent's caring in such cases would be pointless. Discussing the serial killer son example, Shoemaker says:

I may arrive at the consciously held propositional belief that he's a worthless human being, that he's dead to me. And yet when I read of his upcoming execution, I may well up with tears or fall into a depression. 'I still care about him,' I might say. 'There are no reasons to do so—he's an awful man—but it still matters to me what happens to him' (Shoemaker 2011, p.610).

Shoemaker's claim is that he (the subject of the case) is attributability-responsible for these actions and attitudes (welling up with tears; becoming depressed)—despite his inability to give any justifying reasons in support of them. They still stand in a rational relation to his evaluative commitments.

The important thing is that attributability counts as a genuine kind of responsibility. To see why, recall our broader starting point according to which we understand responsibility in terms of being open to a certain range of moral responses. What sort of moral responses underwrite attributability-responsibility? According to Shoemaker, actions and attitudes that we are attributability-responsible for us open us up, minimally, to 'aretaic appraisals'. In other words, they open us up to judgments about the normatively relevant aspects of an agent's *character* in light of the agent's actions or attitudes. We judge each other's characters on the basis of the things we do, think, and feel, where this need not have anything to do with sanctions, requests for reasons, or the reactive attitudes. For example (returning to the serial killer son case), you might think of the father as loyal, or 'generous to a fault' in caring for his son. This need not involve

sanctions, or requests for reasons, or the reactive attitudes. Nevertheless such a judgment constitutes a genuine kind of moral response. Aretaic appraisals are obviously different, for example, from judging someone tall or skinny. They are genuinely moral responses insofar as they have significant practical upshots when deciding who to make plans with, or who to associate with (Shoemaker 2011, p.615). Of course, actions and attitudes that we are attributability-responsible for may open us up to a lot more than mere aretaic assessment—including requests for reasons and more standard kinds of blame-response—but the point here is that they need not.

4. Attributability and Norm Violation

To summarize so far, drawing on Strawson, Scanlon, Smith and Shoemaker, I have argued that there are at least three types of responsibility underwritten by three types of moral response. Schematically, the picture looks like this:

Type of Responsibility	Type of Moral Response
Openness to Blame	Reactive attitudes on the model of sanctions.
Answerability	Requests for reasons in justification of an agent's evaluative commitments.
Attributability	Aretaic assessments, and mere modification of dispositions regarding future interactions.

Before returning to the main context of this paper, I should note that one might question whether the moral responses that underwrite answerability and attributability shouldn't also be classified as a kind of blame-response. I have two things to say about this. First, it seems to me that these latter responses are in some sense too 'weak' to be called blame-responses. I agree with Wallace when he says, 'I think it would indeed be strange to suppose that one might blame another person without feeling an attitude of indignation or resentment toward the person, or that one might blame oneself without feeling guilt' (Wallace 1994, p.75). This strikes me as the commonsense notion of blameworthiness. Perhaps a revisionary account of blame can be motivated and defended, but that leads me to my second point. For those who would insist that these latter responses are a kind of blame-response, the point I am making can be recast in different terminology. That is to say, rather than showing that blameworthiness and responsibility come apart, I can be understood as arguing that a non-revisionary notion of blameworthiness and responsibility come apart, and that the notion of mere attributability is particularly interesting when it comes to motivating (TS). However, in what follows, I will simply put things in the former way.

The basic idea relevant for present purposes is that being open to the ordinary notion of blame-responses is something over and above being responsible, and the notion of attributability is particularly interesting in this respect. To return to our starting point in epistemology, I am going to consider whether there is room to motivate (TS)—the claim that factive epistemic norms are too demanding or too strong—by appealing to the notion of attributability. The idea is that the agents in cases like BIV and GETTIER, for example, are not only not blameworthy for failing to believe only what they know (they are not open to sanctions or reactive attitudes), they are not attributability-responsible for failing to believe only what they know. If it can be shown that there is some kind of tight connection between norm violation and being attributability-responsible, then this is significant. It is significant because, as I will explain, it can take us beyond the important impasse I identified in the epistemic norms debates.

Let's put the idea that there is a tight connection between attributability-responsibility and norm violation in terms of an 'attributability principle' (A):

(A): There is good but defeasible reason to doubt that a norm N has been violated by an agent S if S is not attributability-responsible for doing what N prohibits.

The idea is that attributability, as opposed to blameworthiness, is tightly connected to the notion of a norm violation. Is this a plausible principle? For a start, (A) is not immediately problematized by certain cases that problematize (B). We can see this by considering a kind of case that has come to be regarded as 'the difficult case of psychopaths' in the moral responsibility literature. ¹⁴ Consider a specific example:

PSYCHOPATH: A psychopath cheats an elderly woman out of her life savings.

Ceteris paribus, the psychopath violates some moral norm. The psychopath does something morally wrong in cheating the elderly woman out of her life savings. Notice, then, given some plausible claims about the moral responsibility of psychopaths, (B) gets the wrong results here, while (A) does not. Let me explain.

The psychopath is not the appropriate subject of Strawsonian reactive attitudes and certain kinds of sanctions for cheating the elderly woman out of her life savings. This is not to say that blame would not be understandable in such a case (it is understandable that, for example, the elderly woman would blame the psychopath). The point is that blame would not be reasonable. The reason for this is that such blame-responses imply capacities that the psychopath lacks. These are capacities to recognize the reasons implied by certain demands at play in their relationships with others. As Gary Watson puts it (in a discussion of the case of psychopaths that shares significant similarities with Shoemaker's):

[P]sychopathy (as I read the evidence) precisely involves an incapacity to recognize the interests of others as making any valid claims on them. Consequently, they are disabled from standing in the reciprocal relations or (to use another idiom) from engaging in the mutual recognition that lies at the core of moral life (Watson 2011, p.307).

But, intuitively, the psychopath is morally responsible in some way. This is precisely why the notion of attributability is useful. We can say that the psychopath is responsible in the sense of being attributability-responsible for cheating the elderly woman out of her life savings. That is to say, the psychopath's actions and attitudes are rationally connected to his evaluative commitments (however 'thin' the latter are); they are expressions of his self qua agent. As such the psychopath is open to aretaic assessment. For example, it is appropriate to think of the psychopath as callous, or cold. This is a case, then, in which norm violation and being open to blame-responses (but not attributability) come apart. PSYCHOPATH is a case in which (B) fails. But it not a case in which (A) fails. To be sure, this is a purely negative consideration. It would be helpful if we could say something more positive in support of (A).

The plausibility of the (A) principle is derived from the plausibility of the basic idea that we make normative demands on the actions and attitudes of persons or agents. We do not make normative demands, for example, on inanimate objects or the mechanical proceedings of nature. Consider another case:

AVALANCHE: One day, the weather conditions are just right, and a big avalanche destroys a town in the Rocky Mountains.

Is it appropriate to say that the avalanche violated some moral norm by destroying the town? Obviously not. But why? To put it one way, the avalanche does not manifest the 'special activity' that responsible agents manifest. Of course, we do 'attribute' things to inanimate objects and the proceedings of nature. And we use the language of 'responsibility' to do so. It is perfectly intelligible to say that the avalanche was responsible for the destruction of the ski lodge. Note, however, that it is not plausible to understand this in terms of attributability, as I have outlined that notion. Rather, what is plausible is that the avalanche is causally responsible for the destruction of the ski lodge. ¹⁵

To see how this is relevant for present purposes, consider another case:

DUMMIES: Susan is driving in an obstacle course full of life-like dummies, and her objective is to run all of the dummies over. She has been led to believe (by sources she knows are trustworthy) that it is just a game and there are no real people on the obstacle course. Unbeknownst to Susan, however, a real person has snuck onto the course. In the process of doing her best to succeed in the game, Susan runs the person over.

It seems to me that not only is Susan blameless (she does not seem to be the appropriate subject of reactive attitudes or sanctions) for running a real person over, Susan is not attributability-responsible for running a real person over (where the event is to be understood under that description). The fact that Susan runs a real person over is not rationally related to her evaluative commitments (whatever they are). As evidence for this, notice that we would not adjust our assessment of Susan's character, make decisions about the kinds of plans we would make with her, or become less inclined to associate with her on the basis of this event. These are precisely the sorts of moral responses that underwrite attributability. To put it one way, then, we might say that Susan stands in a relation to her Φ -ing that is similar to the relation that an avalanche stands in to its destruction of a town. And just as we do not think it is appropriate to think of the *avalanche* as violating a moral norm, it is inappropriate to think of *Susan* as violating a moral norm by running a real person over on the obstacle course. The person over on the obstacle course.

Of course, there is at least one important disanalogy between Susan and the avalanche. It has to do with reparation for damages. Susan may be liable for at least some kind of reparation for damages, for example to the family of the deceased. Perhaps minimally she owes them an apology. But the avalanche is of course not liable to make reparations for damages to the town. The fact that Susan may be liable to make reparation for damages reflects on the fact that *she* did it in a sense it seems I want to deny.

There is indeed a sense in which Susan is responsible and the avalanche is not. But, crucially, it is not one that is appropriately linked to the question of norm violation. That is, Susan is not liable to make reparation for damages because she violated some moral norm. Rather, she is liable because something horrible happened and she is at least causally responsible for it. The fact that she ran someone over in this case does not reflect badly on her moral character. However, were she not now to undertake certain actions—such as apologizing, for example—it would reflect badly on her moral character. She is morally responsible for running the person over in a sense that we can call 'forward-looking'.

I think this is broadly what Barbara Harmen has in mind in her defense of Kantian ethics against the objection that the consequences of our actions, and not just the quality of our will, have moral significance:

The basic idea is that the circumstances of failure *call for* a moral response; the agent does not terminate her relation to a context of action once she has adopted an appropriate maxim and prepared herself to act. In this continuity lies not only a plausible response to the objection about actual consequences, but also a welcome widening of the focus of moral judgment from the isolated action to the more natural interval of intention, action, and response (Herman 1993, p.98).

This is surely a matter for further discussion. But the main point is that we can account for the intuition that Susan is responsible for running the person over in a way that does

differ from the way the avalanche is responsible for destroying the town. However, being responsible in the forward-looking sense does not in itself imply that any moral norm has been violated. So, the objection gives us no reason to think that a norm has been violated in the absence of attributability-responsibility. Rather, it implies that something bad, or undesirable has happened, and that a particular agent is linked to the bad or undesirable thing in such a way that certain norms or obligations are triggered.

5. Attributability and Access

Why isn't Susan attributability-responsible for running the real person over? I think the most plausible answer is that attributability—like blameworthiness—implies a certain degree of epistemic access. While many philosophers find it highly plausible that norm violation can occur in ignorance, equally many find it implausible that blameworthiness works the same way. Indeed, appeals to excuses in certain contexts highlight this fact. That is to say, non-culpable *ignorance* is often (though not always) a paradigm excuse or exculpation from blame. The thought here, then, is that attributability-responsibility is enough like being open to blame in this respect. It comes with a kind of condition of accessibility. The fact that there is a real person on the obstacle course is inaccessible to Susan in the relevant sort of way.

It's not clear exactly how the condition of accessibility should be spelled out. For a start, we might say:

Necessarily, an agent is attributability-responsible for Φ -ing only if they are in a position to know that they are Φ -ing.

Note that the condition is not that the agent must know that they are Φ -ing. There is of course a big difference between knowing and being in a position to know. For example, one can fail to know that they are Φ -ing where it would be appropriate to say, 'they should have known.' If one is not even in a position to know, it's far less clear there is any sense to the claim that they should have known. After all, knowing is not something the agent is in a position to do. That said, a good articulation of the accessibility condition will also need to add that an agent can of course fail to be in a position to know when they ought to have been in a position to know. For example, doctors are obligated to read about the latest important findings in their area. Now imagine a GP who doesn't bother, and so ends up making an error that kills a patient. The doctor was not in a position to know that what she was doing was an error (according to the latest findings), but she ought to have been in a position to know. ¹⁹ So perhaps the access condition should go:

Necessarily, an agent is attributability-responsible for Φ -ing only if they are in a position to know that they are Φ -ing (and they have not culpably failed to put themselves in a position to know that they are Φ -ing).

There are no doubt further refinements we could make to such a condition. But the basic idea is that if S is not in a position to know that they are Φ -ing (and they have not culpably failed to put themselves in a position to know), they are not attributability-responsible for Φ -ing. Their Φ -ing does not bear a rational relation to their evaluative commitments. It's hard to see how to argue otherwise. And again, this seems like the best explanation of why Susan is not attributability-responsible for running a real person over.

Since attributability is understood in terms of actions or attitudes that bear a rational relation to the agent's evaluative commitments, it is also natural to ask: What does the accessibility condition on attributability imply about an agent's self-knowledge? That is, what does it imply about the agent's knowledge of their own evaluative commitments? One might worry that, if it involves the idea that agents must know what their evaluative commitments are in order to for their Φ -ing to stand in a rational relation to those commitments (and thus for them to be in the market for counting as attributability-responsible for Φ -ing), agents will fail to count as responsible in a large number of cases in which they intuitively are. After all, our evaluative commitments are sometimes (likely often) opaque to us. For example, consider cases of denial or repression about one's sexual orientation, or one's feelings about their career.

However, there is a clear distinction between an agent's having the right kind of access to the conditions under which they count as Φ -ing, and an agent's having access to what their evaluative commitments are. The accessibility condition on attributability does not imply that agents must be in a position to know what their evaluative commitments are in order to be attributability-responsible for Φ -ing (perhaps something like that is required for *answerability*). All the accessibility condition on attributability says is that the agent must have the right kind of access to the conditions under which they count as Φ -ing, in order for their Φ -ing to stand in a rational relation to the agent's evaluative commitments, or manifest their agential self.

There is also an interesting question about the description under which an action or attitude counts as being something the agent is attributability-responsible for. For example, let's say for sake of argument that Susan is attributability-responsible for running something over. It is plausible that Susan is attributability-responsible for running something over that looks just like a dummy. But it is far less plausible that Susan is attributability-responsible for running a real person over (qua real person). What explains the difference? I think the best explanation is that Susan does not have the right kind of access to the fact that she is running a real person over. Judgments about the descriptions under which actions or attitudes count as actions or attitudes for which an

agent is attributability-responsible are nicely explained by the accessibility condition on attributability.

6. The Attributability View

(A) and the accessibility condition on attributability are two commitments underlying a more plausible way of motivating (TS). Call it the Attributability View.²⁰ We can quickly see the Attributability View in action by returning to the BIV case. The BIV is not in a position to know that he fails to believe only what he knows (for example, when he believes that he has hands). Thus according to the accessibility condition, the BIV's failure to believe only what he knows is not something he is attributability-responsible for. And because the BIV is not attributability-responsible for his failure to believe only what he knows, according to the (A) principle, this undermines the judgment that he violates the epistemic norm of belief. Insofar as proponents of the knowledge norm of belief are committed to claiming that agents like BIV violate the norm of belief, the Attributability View motivates the claim that the knowledge norm of belief is too strong. It motivates (TS).

The Attributability View provides a way of motivating (TS) that is more structured and theoretically motivated than mere appeal to intuition, and is more plausible than (B). Of course, proponents of the knowledge norm will still want to say something in response at this point. But can they appeal to excuses, like before? I don't think so. This is because I have connected responsibility to norm violation in a way that is not about being open to blame responses in the first place. Arguably, this means the idea of an excuse is now out of place in responding to (TS), since the point of an excuse is to exculpate, i.e. to show why blame-responses are not appropriate. Rather than explaining why agents in GETTIER and BIV aren't blameworthy but nevertheless violate a norm, the Attributability View says we need a story about why (A), or the accessibility condition on attributability, or both, is misguided or false. These are commitments that both sides of the debate should be focusing on.

I won't subject these commitments to further scrutiny here. But I will make a couple of points about why I think responding to (TS) won't be as easy as it looks when we take (TS) to be motivated by (B). For starters, the weaker the commitments underlying (TS) are, the harder (TS) should be to explain away. (A) is a weaker principle than (B), at least in the sense that it involves a much weaker notion of responsibility. Once again, (B) turns on considerations of what it takes to be blameworthy. I have argued that this is something over and above what it takes to be attributability-responsible. The minimal condition on an action or attitude being something that an agent is attributability-responsible for is that the action or attitude bears a rational connection to the agent's evaluative commitments. Correspondingly, I think, it is relatively easy to imagine cases of agents violating norms via blameless actions or attitudes. But, as

AVALANCHE and DUMMIES highlight, it is difficult to imagine cases of agents violating norms via actions or attitudes that do not even bear a rational relation to the agent *qua* agent (i.e. to their evaluative commitments).

Another point arises in the context of a related debate. I can only touch briefly on it here. The debate concerns a distinction between what is sometimes called 'factual ignorance' and 'evaluative ignorance', and the relevance this has for judgments about culpability. Consider the following case:

50s FATHER: Suppose Don is your typical television father from the late 50s or early 60s. He loves his children and does what he can to try to keep them happy. He puts money away for his son to go to school. He puts money away for a sailboat for his daughter. Although his daughter has said repeatedly that she wants to go to school, he sees no reason to help her because he sees no reason for women to go to college (Littlejohn 2014, p.144).

Littlejohn (2014) plausibly claims that Don ought to have saved for his daughter's education. And he maintains this is so, even though Don had no evidence that would rationally support the belief that he ought to do so (that is to say, even though he is 'evaluatively ignorant'). Littlejohn and others argue that there is an asymmetry between the appropriateness of blame in cases of evaluative versus factual ignorance. They think, in addition to violating moral norms, agents in cases like 50s FATHER can even be the appropriate subjects of blame—despite being evaluatively ignorant. 'Because Don shouldn't act like a sexist and his sexism isn't an excuse for his wrongdoing, he's culpable for his wrongful behavior' (Littlejohn 2014, p. 144). And we might think that, if this is plausible, it leads to problems for those who think, for example, that the BIV is not merely excused but *justified* in believing that p despite factual ignorance. How can factual ignorance justify while evaluative ignorance sometimes doesn't even excuse?

This is a complex issue worth exploring in much more detail. I will simply claim here that the Attributability View gets the right results in this sort of case, and I think this speaks strongly in its favour. That is, Don is attributability-responsible for failing to save for his daughter's education. This is because his decision not to save for his daughter's education bears a rational connection to his evaluative commitments. So it is consistent with the view to maintain that he violates some moral norm in so doing (the view leaves it open whether or not Don is blameworthy). More generally, the Attributability View will allow that agents can be attributability-responsible for actions arising out of evaluative ignorance, while they are not attributability-responsible for actions arising out of the right kind of factual ignorance.

7. Conclusion

By taking a step back and exploring recent work in the theory of responsibility, I have put forward a novel way of motivating (TS). As I have noted, the aim has not specifically been to claim that (TS) is a fatal blow to the knowledge norm of belief, or factive epistemic norms more generally. Rather, I have argued that a prominent way of motivating (TS) leads to an impasse (the (B) principle); and I have argued that the Attributability View opens up space to move beyond that impasse. When it comes to the project of responding to (TS), the appeal to excuses generally seems to presuppose that (B) is the only way of explaining, or providing a rationale, or backing up (TS). There is a different, more plausible way of backing up (TS). So, in effect we can see that this presupposition is false. While there is an increasing amount of debate about the appeal to excuses in epistemology, the Attributability View reveals a new challenge. What is needed in responding to (TS), rather than a story about how agents can blamelessly violate norms, is a story about how agents can violate a norm by Φ -ing, and yet fail to be attributability-responsible for Φ -ing. 21

Centre for Logic and Analytic Philosophy KU Leuven

References

- Adler, J. (2002). Belief's Own Ethics. Cambridge, MA: MIT Press.
- Audi, R. (2001). 'An internalist theory of normative grounds'. *Philosophical Topics*, 29:19-46.
- Boult, C. (2016a). 'Epistemic normativity and the justification-excuse distinction'. *Synthese*, Online First: DOI 10.1007/s11229-016-1127-8.
- _____(2016b). 'Epistemic conditions on "ought": E=K as a case study'. *Acta Analytica*, Online First: DOI 10.1007/s12136-016-0301-4.
- Brown, J. (2010). 'Knowledge and assertion'. *Philosophy and Phenomenological Research*, 81:549-566.
- Burge, T. (2003). 'Perceptual entitlement'. *Philosophy and Phenomenological Research*, 67:503-548.
- Cohen, S. (1984). 'Justification and Truth'. *Philosophical Studies* 46: 279-95.
- DeRose, K. (2002). 'Assertion, knowledge, and context'. *Philosophical Review*, 111:167-203.
- Douven, I. (2006). 'Assertion, knowledge, and rational credibility'. *Philosophical Review*, 115:449-485.
- Dretske, F. (1981). *Knowledge and the Flow of Information*. Cambridge, MA: MIT Press. (1971). 'Conclusive reasons'. *Australasian Journal of Philosophy*, 49:1-22.

- Gardner, J. (2001). 'Obligations and outcomes in the law of torts'. In Cane, P. and Gardner, J. (eds.), *Relating to Responsibility: Essays for Tony Honore*. Hart Publishing.
- Gerken, M. (2011). 'Warrant and action'. Synthese, 178: 529-47.
- Gibbons, J. (2013). The Norm of Belief. Oxford: OUP.
- Goldman, A. (1975). 'Innate knowledge'. In Stephen Stich (ed.), *Innate Ideas*, Berkeley: University of California Press.
- Graham, P. (2012). 'Epistemic entitlement'. Noûs, 46:449-482.
- Greco, J. (2010). Achieving Knowledge. Cambridge: CUP.
- Hattiagandi, A. (2010). 'The love of truth'. *Studies in the History and Philosophy of Science*, 41:422-432.
- Huemer, M. (2007). 'Moore's Paradox and the Norm of Belief'. In, Nuccetelli, S. and Seay, G. (eds.), *Themes from G.E Moore: New Essays in Epistemology and Ethics*. Clarendon Press.
- Lackey, J. (2014). 'Assertion and expertise'. *Philosophy and Phenomenological Research*, 92:509-517.
- (2007). 'Norms of assertion'. *Nous* 41: 594-626.
- Lehrer, K. & Cohen, S. (1983). 'Justification, truth, and coherence'. *Synthese*, 55:191-207.
- Littlejohn, C. (forthcoming). 'A plea for epistemic excuses'. In, F. Dorsch and J. Dutant (eds.), *The New Evil Demon*. Oxford: OUP.
- (2014). 'The unity of reason'. In *Epistemic Norms*. Oxford: OUP.
- McHugh, C. & Whiting, D. (2014). 'The normativity of belief'. Analysis, 74:698-713.
- Plantinga, A. (1993). Warrant: The Current Debate. New York: OUP.
- Pritchard, D. (2012). 'Anti-luck virtue epistemology'. *The Journal of Philosophy*, 109:247-279.
- Scanlon, T.M. (2008). *Moral Dimensions: Permissibility, Meaning, Blame*. Cambridge, MA: Belknap Press.
- _____(1998). *What We Owe to Each Other*. Cambridge, MA: Harvard University Press.
- Shoemaker, D. (2011). 'Attributability, answerability, and accountability: toward a wider theory of moral responsibility'. *Ethics*,121:602-632
- Simion, M. (2015). 'Assertion: knowledge is enough'. *Synthese*, Online First: DOI 10.1007/s11229-015-0914-y.
- Smith, A. (2012). 'Attributability, answerability, and accountability: in defense of a unified account'. *Ethics*, 122:575-589.
- _____(2008). 'Control, responsibility, and moral assessment'. *Philosophical Studies*, 138:367-392.
- Sosa, E. (2015). Judgement and Agency. Oxford: OUP.
- _____(2011). *Knowing Full Well*. Princeton University Press.

- Srinivasan, A. (2015). 'Normativity without Cartesian privilege'. *Philosophical Issues*, 25:273-299.
- Steup, M. (1999). 'A defense of internalism'. In L. Pojman (ed.) *The Theory of Knowledge: Classical and Contemporary Readings, 2nd Edition*. Wadsworth Publishing.
- Strawson, P. (1962). 'Freedom and resentment'. *Proceedings of the British Academy*, 48:1-25.
- Sutton, J. (2007). Without Justification. Cambridge, MA: MIT Press.
- Talbert, M. (2008). 'Blame and responsiveness to moral reasons: are psychopaths blameworthy?' *Pacific Philosophical Quarterly*, 89:516–535.
- Watson, G. (2011). 'The trouble with psychopaths'. In Wallace, R.J., Kumar, R. and Freeman, S. (eds.) *Reasons and Recognition: Essays on the Philosophy of T.M. Scanlon*, Oxford: OUP.
- Whiting, D. (2013). 'Nothing but the truth: on the norms and aims of belief'. In Timothy Chan (ed.), *The Aim of Belief*. Oxford: OUP.
- Williams, B & Nagel, T. (1976). 'Moral luck'. *Aristotelian Society Supplementary Volume*. 50:115-151.
- Williamson, T. (forthcoming). 'Justifications, excuses, and sceptical scenarios'. In, F. Dorsch and J. Dutant (eds.), *The New Evil Demon*. Oxford: OUP. (2000). *Knowledge and its Limits*. Oxford: OUP.

¹ In this paper, I set to one side the question of whether there is a single epistemic norm of belief, or any norm of belief at all. Those are interesting questions. But I simply presuppose the idea that there is a (single) norm of belief as a helpful way of getting at the issue I am interested in here.

² Proponents of a knowledge norm of belief include Adler (2002), Huemer (2007), Littlejohn (2012), Sutton (2005), and Williamson (2000; 2013).

³ The view that truth is the epistemic norm of belief is not prominently defended. That said, a closely related view—that truth is the *constitutive* norm of belief—is more prominent. For example, see Whiting (2013). While closely related, the constitutive norm of belief is importantly different from the epistemic norm of belief (McHugh and Whiting 2014). The epistemic norm of belief determines whether or not a belief is epistemically justified. The constitutive norm of belief, if there is one, is a norm that determines what it takes for a given mental state to *be* a belief.

⁴ For further criticism of the truth norm of belief, as well as support for the knowledge norm, see Hattiangadi (2010).

⁵ There is also a worry about the sufficiency of knowledge, though this kind of objection has seen more attention in the assertion literature (Brown 2010, Lackey 2014; Simion 2015). In this paper, I focus attention solely on the issue of necessity.

⁶ Sometimes it is claimed that the problem was introduced in Lehrer & Cohen (1983).

⁷ Srinivasan challenges the idea that there is a tight connection between blameworthiness and norm violation. But I take the fact that she dedicates an entire paper to this and related issues to indicate she thinks it is an idea worth discussing.

⁸ Boult (2016a) provides an in-depth analysis of this issue.

⁹ I will speak of the 'Scanlon/Smith view,' but I do not mean to suggest that they have collaborated or explicitly endorse each other's views on responsibility (though this happens to be true of Smith, at least).

Rather, I use this terminology to engage with Shoemaker, who also groups the two together as representatives of *broadly* the same sort of view of responsibility.

As Smith notes, the 'volitional' view is a loose cluster of views with important differences. Roughly, they share a commitment to the idea that choice, identification, decision, or voluntary control is central to responsibility. Smith briefly points out differences between 'prior choice' views, 'endorsement' views, and 'voluntary control' views. See Smith (2005), p. 240).

¹¹ I don't intend these brief remarks on its motivation to settle the question of whether the Smith/Scanlon view is the best view of responsibility available. I do think it is a very plausible view, and it is a prominent view in the moral responsibility literature (one that I think epistemologists would do well to pay more attention to). The idea here is simply to appeal to it as a useful framework. Indeed, I regard its fruitfulness in application to the epistemic norms debates as a mark in its favour.

¹² To quote Smith, 'these judgments, taken together, make up the basic evaluative framework through which we view the world' (Smith 2005, p.251).

¹³ To be clear, Smith also uses the term 'attributability' when talking about responsibility. In Smith's mouth, to say that some action or attitude is attributable to someone is merely to say that it is *theirs* in the minimal sense required for them to be held responsible for it. She uses the notion to pick out a sense of 'responsibility' that is neither positive nor negative, but rather only implies that an agent is a candidate for negative or positive moral assessments. As Shoemaker himself points out, his dispute with Smith at this stage concerns what is necessary for attributability. He thinks it is something weaker than answerability. So, when I refer to Shoemaker's notion of attributability I should be taken as meaning *mere* attributability.

¹⁴ I rely on Watson's (2011) characterization of the psychopathic profile: 'Psychopaths (as such) suffer no psychoses or notable neuroses, or general deficits of intelligence. Their condition is characterized by callous interpersonal relations, lack of significant attachment to other people or institutions (shallow and fleeting 'friendships,' amorous relations, and loyalties), lack of a sense of shame and guilt, lack of self-criticism, refusal to take responsibility for the troubles caused to others or oneself, and lack of sincere commitment to long-term goals. Psychopathy is present from childhood and, by all indications, endures a lifetime. As this description implies, this condition has proven unamenable to psychopharmacological or psychological therapies and treatments' (Watson 2011, p. 308).

¹⁵ Of course, in the context of someone pressing me on this issue, these sorts of differences are exactly what we'll want to know a lot more about. See Gibbons (2013, Ch.6) for an interesting discussion of some of the difficulties in this area.

¹⁶ More on descriptions of events shortly.

¹⁷ This is not to say that the event is not unfortunate and that the person's death is not of moral significance. What (A) says about this case is that there is good (though defeasible) reason to doubt that Susan has really violated some moral norm. If this sounds jarring, note that I don't intend to persuade those who have strong intuitions to the contrary. Those who have strong intuitions to the contrary should think of (A) as a proposed rationale or way of understanding why anyone would ever find (TS) compelling.

¹⁸ Thanks to Clayton Littlejohn for pressing this point.

¹⁹ Thanks to Duncan Pritchard for this example.

²⁰ In Boult (2016b) I appeal to a version of the Attributability View to provide an account of what John Gibbons calls the 'natural reaction' to the idea that a person can be normatively required to Φ when this requirement is in some sense outside the person's first person perspective; I focus specifically on Williamson's E=K thesis as a case study and provide an account of some troubling features of E=K.

²¹ Acknowledgements. Thanks to Fernando Broncano-Berrocal, Allan Hazlett, Chris Kelp, Clayton Littlejohn, Aidan McGlynn, Ram Neta, Duncan Pritchard, Mona Simion, and audiences at the Epistemic Norms Conference, KU Leuven, and the CLAW Seminar, KU Leuven for helpful feedback on earlier drafts of this paper.