



What do We Want from a Theory of Epistemic Blame?

Adam Piovarchy

To cite this article: Adam Piovarchy (2020): What do We Want from a Theory of Epistemic Blame?, Australasian Journal of Philosophy, DOI: [10.1080/00048402.2020.1833353](https://doi.org/10.1080/00048402.2020.1833353)

To link to this article: <https://doi.org/10.1080/00048402.2020.1833353>



Published online: 26 Oct 2020.



Submit your article to this journal [↗](#)



Article views: 101



View related articles [↗](#)



View Crossmark data [↗](#)



What do We Want from a Theory of Epistemic Blame?

Adam Piovarchy 

University of Sydney

ABSTRACT

This paper identifies a number of questions that any plausible theory of epistemic blame ought to answer. What is epistemic blame? When is someone an appropriate target of epistemic blame? And what justifies engaging in epistemic blame? I argue that a number of problems arise when we try to answer these questions by using existing conceptions of moral blame. I then consider and reject Brown's [2020] belief-desire model of epistemic blame. Finally, I argue that an agency-cultivation model of moral responsibility is not only able to help us to develop a plausible theory of epistemic blame; it is particularly well-placed to do so.

ARTICLE HISTORY Received 29 March 2020; Revised 11 September 2020

KEYWORDS epistemic blame; blameworthiness; responsibility; culpability

1. Introduction

Philosophers examining the intersection between blameworthiness and belief have focused on whether agents can be morally blameworthy for their beliefs. It is generally accepted that agents can have moral obligations to follow epistemic norms, and can be blameworthy for violating those norms without an excuse. Recently, attention has shifted to examining whether agents can be *epistemically* blameworthy for their beliefs or reasoning—that is, blameworthy for violating epistemic norms *qua* epistemic norms, where this blameworthiness is neither reducible to moral blameworthiness, nor indirectly justified by background moral considerations [Rettler 2018; Brown 2020, forthcoming].

There are two main ways in which we could try to construct an account regarding the nature of, and justification for, epistemic blame. The first is to treat it as an independent phenomenon, considering thought experiments and paradigmatic cases to look for conditions that make agents epistemically blameworthy, and from there try to understand the nature of epistemic blame. While it may be tempting to build a new theory from the ground up, if our conception of epistemic blame is too distinct from our conception of moral blame then there is a risk that there won't be any shared features that warrant considering each to be a species of blame.

I think that it would be wiser to take stock of the literature on moral blame and blameworthiness. Understanding the ways in which epistemic blame is similar to moral blame is helpful for investigating whether the two share some foundation. Having a clear articulation of their differences can also ward off worries that instances

of what we identify as epistemic blameworthiness are actually reducible to moral blameworthiness. Additionally, examining what we know about moral blame can help us to identify and avoid objections that might apply to any proposed account of epistemic blame, and to provide some guidance regarding what a theory of epistemic blame should look like.

The aim of this paper is to identify what we want from a theory of epistemic blame, while keeping an eye on our theories of moral blame. The paper engages in some ground-clearing regarding recent proposals, and identifies some standards that future proposals ought to meet if we are to avoid scepticism about epistemic blame, by proceeding as follows. First, I identify a number of questions that any theory of epistemic blame ought to answer, noting a number of *prima facie* similarities between moral blame and epistemic blame. I then consider some problems that arise when we try to use models of moral blame to understand what epistemic blame is, when epistemic blame is warranted, and what justifies engaging in epistemic blame. Notably, it seems like our existing conceptions of moral blame involve distinctly moral components that cannot be modified or adapted for our theory of epistemic blame. I then consider Brown's [2020] recent belief-desire account of epistemic blame, and argue that it faces a few problems. In order to show that a theory that avoids the above problems is possible, I close by arguing that Vargas's [2013] agency-cultivation model of responsibility is particularly well-suited to give us what we want from a theory of epistemic blame.

2. Some Minimal Requirements

Examples of epistemic blame and blameworthiness seem common enough in everyday life. A student affirms the consequent, even though we just spent hours explaining how this is a fallacy, stirring an impulse to raise our voice. A teammate makes a foolish move against an opponent in a game of cards, losing the game and garnering a swift rebuke. Relatives quote pseudoscience despite having previously been told about fake news and misleading sources, which escalates into angry yelling across the dinner table.

While it is uncontroversial enough to point to such cases and to describe such agents as epistemically blameworthy, one worry is that these are actually instances of something else. Consider someone with racist beliefs that are the result of uncritically accepting bad arguments from unreliable sources. Such a person seems epistemically blameworthy, but they are also morally blameworthy. And a sceptic about epistemic blameworthiness might say that the above examples of epistemic blameworthiness are actually instances of moral blameworthiness. We think that people have a moral duty to inform themselves of certain facts about the world, particularly when there are high moral costs to getting the answer wrong [Graham 2015; Vanderheiden 2016; Goldberg 2017]. Additionally, epistemic vices often have some relation to moral vices: for example, selfishness can cause one to spend less time attending to information that would inform them of the needs of others.

There are a few reasons to think that blaming agents for epistemic norm violations is not necessarily moral. We often think it appropriate to blame agents in such cases even if no moral considerations against violating epistemic norms are present, or even if the moral considerations against violating epistemic norms are outweighed by other moral considerations in favour of violating them. It also often seems the

case, phenomenologically speaking, that, when we epistemically blame someone for violating a particular norm, it really is for violating an epistemic norm, for ‘not getting it’, or for holding contradictory beliefs, rather than for violating some moral principle. Finally, numerous philosophers have presented arguments featuring agents whom they take to be blameworthy for violating epistemic norms *qua* epistemic norms [Sutton 2007; Williamson 2007; Littlejohn forthcoming].

What has been written on epistemic blame has tended to focus on showing that, when we blame agents for believing badly, this is not actually an instance of moral blame [Goldberg 2017; Brown forthcoming]. I won’t repeat such arguments here. I will simply proceed under the assumption that we do engage in epistemic blame *qua* blame for violations of irreducibly epistemic norms. Readers need not worry that this could be unjustified (and thus epistemically blameworthy!), because my aim in this paper is not to try to establish anything that follows from the assumption that epistemic blame and blameworthiness exist. Instead, it is to identify a number of questions that any account of epistemic blame ought to be able to answer, and to see whether we can answer them by using existing theories of moral blame, something that epistemic blame sceptics and realists alike should find worthwhile.

To have a plausible theory of epistemic blame, it seems that we ought to have substantive answers to these three questions:

- (1) *The conceptual question.* What is epistemic blame?
- (2) *The warrant question.* When is someone an appropriate target of epistemic blame?
- (3) *The justification question.* What justifies engaging in epistemic blame?

While a satisfying theory might also need to answer additional questions, these are fairly minimal requirements. Failure to answer any of these questions would be cause for doubt that epistemic blame is ever justified, worthy of philosophical investigation, or something genuinely distinct from moral blame or anger.

One reason why we need an answer to all three questions is that our answer to each is likely to have implications for the others. For example, if epistemic blame is thought of as something very unpleasant and as strongly against people’s interests, then we’re likely to want a stronger justificatory story for why we should keep it around. Another reason is that it is possible to answer some questions while also maintaining that no one is ever epistemically blameworthy. Thus, even if instances of epistemic blame are common, without a unified account it might turn out that epistemic blame is never justified. By analogy, while it is common to think of moral blame as a kind of sanction, some philosophers argue that agents are only deserving of moral blame-*qua*-sanction if their wrongdoing is an exercise of their libertarian free will. But one can believe this while also believing that, since determinism is true, no one in fact is an appropriate target of moral blame.

A natural place to look for answers to these questions is within theories of moral blame, not only because epistemic and moral blame share a similar name, but because they also share a number of *prima facie* similarities.

- First, both epistemic blame and moral blame can be fitting or unfitting, in that they can correctly or incorrectly represent their target. Someone who has not violated a moral obligation is typically not a fitting target of moral

blame, and someone who has not committed any error in their reasoning is typically not a fitting target of epistemic blame.

- Second, agents who violate moral or epistemic norms can be excused for doing so. Small children and the mentally disabled are typically not at fault when they violate moral or epistemic norms, and thus are not appropriate targets of epistemic blame.
- Third, the grounds for blaming agents seem to track certain failures of sensitivity to reasons. Moral blame is often targeted towards agents who failed to see or to act on certain moral reasons that they should have seen or acted upon. Likewise, epistemic blame seems to be targeted towards agents who failed to recognise or believe in accordance with certain epistemic reasons that they ought to have.
- Fourth, neither moral blame nor epistemic blame is reducible to mere negative evaluations or disappointment. It seems perfectly coherent to say ‘I’m disappointed in what he did, and he should be blamed for his wrongdoing [or bad reasoning], but I don’t blame him for it.’
- Fifth, neither moral blame nor epistemic blame is reducible to mere anger or frustration (although expressions of anger can be expressions of blame). While we become angry at the weather or at being stuck in traffic, we don’t morally blame the weather or the traffic. And while we might get frustrated at someone’s failure to see our point of view, this doesn’t always amount to epistemically blaming them. We might experience such feelings in response to agents whom we know are excused, for instance.
- Sixth, neither moral blame nor epistemic blame needs to be expressed. Both can be felt privately, or directed towards the dead. One can morally blame their boss for being arrogant, and can epistemically blame their boss for having poor business acumen, but fail to express these for fear of being fired.

These similarities offer some considerations for thinking that epistemic blame and moral blame are related in some important way. But there is another important reason for thinking that our conceptions of epistemic blame and moral blame ought to connect. A conception of epistemic blame that is too distinct from our conception of moral blame risks leaving us without any justification for thinking of the former as a species of blame at all, as opposed to mere anger or frustration, which we sometimes experience in response to frustrating events generally.

In the next section, I will argue that attempting to find a connection between moral and epistemic blame is more difficult than it might seem at first glance, because moral blame seems to involve distinctly moral properties that cannot be simply transplanted to our account of epistemic blame. This leaves the epistemic blame theorist with something of a dilemma. While failing to provide any connection between moral blame and epistemic blame risks failing to provide an account of epistemic *blame*, constructing a theory of epistemic blame by using our conception of moral blame risks failing to provide an account of truly *epistemic* blame.

3. Moral Blame, Epistemic Blame

To see if we can make any progress answering the conceptual, warrant, and justification questions for our theory of epistemic blame, let’s consider some theories of moral blame.

A number of philosophers writing on moral responsibility take blame to be a kind of sanction, of which culpable wrongdoers are deserving (for example, Levy [2011], Pereboom [2014], and Nelkin [2016]). It has been noted that blame carries with it a kind of ‘sting’ or ‘opprobrium’, and that being targeted with blaming-attitudes like resentment and indignation is generally unpleasant. Philosophers who support the sanction model of blame typically maintain that agents are blameworthy if they committed wrongdoing while possessing a certain kind of control over their actions (for example, reasons-responsiveness, or the capacity to avoid wrongdoing). And what justifies blaming culpable wrongdoers who possessed the relevant kind of control over their actions at the time of wrongdoing is that they deserve to be blamed.

This sanction model of blame seems able to answer the warrant question in a way that our account of epistemic blameworthiness could adopt. We could say that agents are epistemically blameworthy if they violate an epistemic norm that they had the capacity to follow, or if their believing badly issued from a sufficiently reasons-responsive mechanism. But problems arise when we try to use the moral blame-as-sanction model to answer the conceptual question, because these accounts don’t identify blame with just any sanction, but rather with hostile reactive attitudes such as resentment and indignation. And this is a worry because resentment and indignation don’t ever seem like appropriate responses to epistemic blameworthiness, no matter how egregious is someone’s believing.

One could perhaps retreat to a weaker conception of deserved sanction, given that being thought of as a bad reasoner, or that being criticised for violating epistemic norms, are unpleasant responses that don’t seem unfair or inappropriate. This response leads to three problems. The first is that if mere negative evaluations constitute a sanction, this would entail that believing that someone is blameworthy simply is one’s blaming them, making the thought ‘He’s blameworthy, but I don’t blame him’ contradictory. The second is that this approach needs to explain why some sanctioning responses (for example, resentment, indignation, and retributive anger) are appropriate responses to moral blameworthiness but not epistemic blameworthiness, while others (for example, criticism, disappointment, and non-retributive anger) remain appropriate responses to either. Finally, these weaker responses also seem warranted, and not unfair, ways to respond to agents who are commonly taken to be excused for their wrongdoing, which is partly why blame-as-sanction proponents don’t think of such responses as the relevant blaming sanctions. Children, the mentally disabled, and agents who had poor formative circumstances can be cruel or callous, or can manifest any number of bad qualities. Yet we can fittingly evaluate them to be cruel, callous, and the like, and can criticise them or react with disappointment, even while maintaining that these agents are excused for their wrongdoing and are not blameworthy. The worry is that, if we move to this weaker sense of sanction when considering epistemic blame, we risk ruling out the possibility that agents could ever be excused for violating an epistemic norm, or risk saying that excused agents are fitting targets of epistemic blame after all.

Given that there seems to be something odd about thinking that someone deserves to be blamed, while also maintaining that nothing is meant by this in any moral sense, it is worth pausing to consider some considerations lurking in the background of our theories of moral blame and moral responsibility. A key feature with which philosophers theorising about moral blame have been concerned, and that conceptions of

epistemic blame seem unlikely to adopt, is that moral blame is the paradigmatic means by which we hold each other to account for failing to comply with the moral *demands* that we make of one another [Watson 1996; Darwall 2006]. In particular, our dignity, moral status, or membership in the moral community imbue us with an authority to demand that others treat us and others in various ways, and blame is our way of holding agents who fail to do this accountable to us.

In contrast, it seems hard to think of epistemic norms as demands in a similar sense. Demands typically have a demander, and I'm not sure that I have the authority to legitimately demand that my students not affirm the consequent. Any compulsion that we feel to comply with epistemic norms seems to come from our faculty of reason, not from other agents.¹ Likewise, when we try to compel other people to follow epistemic norms, we are typically trying to encourage them *to see reason*. We are not trying to make them do as we say because of what we say, or because of some feature about us, as we might do when we demand that they remove their foot from our painful gouty toe. It doesn't seem like we can demand that other agents reason appropriately, independently of any moral obligations to do so, or can hold them to account when they fail.

Some philosophers argue for a different conception of moral responsibility altogether, taking it to be primarily concerned with holding each other answerable, not accountable. Roughly, to say that an agent is morally responsible for an action, trait, or belief is to say that the agent is an appropriate target of requests or demands for justification for that thing, and is eligible for, or open to, certain kinds of responses based on their answer [Hieronymi 2004; Smith 2013; Rettler 2018]. This seems able to give us an account of epistemic responsibility, because our beliefs and reasoning are the sort of thing for which we can be asked to provide justification.² But showing that blame in particular is an appropriate response needs a supplementary story.

According to Smith [2013] and Hieronymi [2004], blame is a form of *protest*, which aims to

register the fact that the person wronged did not deserve such treatment by challenging the moral claim implicit in the wrongdoer's action; second, to prompt moral recognition and acknowledgment of this fact on the part of the wrongdoer and/or others in the moral community. [Smith 2013: 43]

These philosophers are Attributionists, arguing that agents are blameworthy for wrongdoing when that wrong expresses or reflects an objectionable evaluative judgment that is attributable to the agent in some way. It initially seems like this model of moral blame could help us to build our theory of epistemic blame. We could say that agents are epistemically blameworthy when their bad reasoning is attributable to them, and that blame seeks to register the fact that someone has violated an

¹ However, it is also worth noting that—as an anonymous reviewer points out—for many Kantians moral principles are confronted by us as imperatives issued by reason, and so do not necessarily require a particular demander. I'll set aside this possibility and restrict my attention to the contemporary conceptions within the moral responsibility literature.

² One objection to this conception is that demands for justification are not apt responses to wrongdoing because we don't demand excuses or justifications. Instead, blaming-judgments presuppose that no excuse or justification is available. In so far as we think that blame has some important connection with demands, it seems that this is because the agent was subject to demands prior to their wrongdoing. See Nelkin [2015].

epistemic norm, and to prompt recognition of this fact by the agent and the rest of the community.

Setting aside the earlier worry that we still need an explanation for why more hostile responses can be apt when considering moral violations, but inapt as responses to epistemic violations, there are a few reasons to think that using this model for our theory of epistemic blame would be stretching the Attributionist's conception of 'protest'. The first regards what it is against which we are protesting. An important feature of these accounts is that, when we blame, we are forcefully repudiating the attitudes that the wrong expressed—namely, that the wrongdoer has the right to treat others in objectionable ways. This matters, because, if we were blaming only the wrong itself, we would lack any means of distinguishing between blameworthiness of wrongs committed by culpable agents and wrongs committed by excused agents. According to Attributionists, young children and the mentally disabled are excused precisely because their lack of awareness of what they are doing, or of the significance of their actions, means that their wrongdoing is not attributable to them in the right way, or that it fails to express an objectionable attitude.

The problem for the epistemic blame theorist is that it doesn't seem that agents who believe badly express anything stronger than 'This inference is warranted.' Not only is this the kind of judgment that young children and the mentally disabled can express; it also doesn't seem particularly objectionable, in the sense of being the kind of thing that calls for repudiation and disavowal. When I blame a student for affirming the consequent, it doesn't seem like I'm standing up for informal logic, truth, or reason. And it doesn't seem like the student should now publicly disavow their previous inference, or solemnly commit to never making such an error again, although an acknowledgment of error may be fitting.

Another worry is that, according to Smith, moral protest also implicitly seeks a moral reply—namely, an apology that disavows the wrong, and that retracts the threatening claim that the wrong expressed. But genuine apologies seem out of place in cases of epistemic blameworthiness, and this needs to be explained. While we do often say 'I'm sorry' after reasoning badly, we need to distinguish between apologies and mere expressions of regret or condolences. We might say 'I'm sorry' upon hearing that one of your relatives has died, but such utterances are not truly apologies. That we don't apologise for being epistemically blameworthy also seems related to the fact that, as a number of philosophers have argued, the first-person analogue of blame is guilt [Carlsson 2017], and epistemic guilt doesn't seem like something that we experience. I might experience regret at having made such an epistemically foolish decision, but if there was never anything of moral significance at stake then I won't feel guilty about it. I especially don't feel the need to *confess*, which is guilt's central action tendency [Tierney forthcoming]. And, because we don't think of epistemic requirements as obligations *to* other agents, we generally don't think that violations of those requirements call for responses to those agents either, such as by making amends.

4. Brown's Account of Epistemic Blame

One account of epistemic blame that has made some progress towards answering our questions comes from Brown [2020], who answers the conceptual question by adopting and modifying Sher's [2006] account of moral blame. To understand the problems facing this proposal, we first need to understand his account.

Sher takes moral blame to consist in a disposition towards a certain set of affective and behavioural responses, which contingently stem from a belief that someone has acted badly without excuse, and a desire that they not have acted badly. A benefit of this account is that it seems able to explain why moral blaming reactions can take such diverse forms. By treating blame as a belief-desire pair associated with a set of dispositions, Sher can account for why yelling, privately seething, and dispassionately blaming can each constitute blame despite their having so little in common.

Sher provides an original answer to the justification question. He takes moral blame to be justified by the very same considerations that justify our commitment to morality. Because of morality's prescriptive, universal, and overriding nature, my commitment to morality will not only produce a desire that my actions be right; I will also desire that other agents' actions (past, present, and future) be right. But the desire that a wrongdoer not have violated a moral principle simply *is* the desire that one experiences when one blames. And so this desire and its associated dispositions are jointly a way of caring about and being committed to morality. If we were to forswear blaming responses from our lives, we would be forswearing an important part of our commitment to morality, too.

Brown [2020] argues that this model can easily be modified so as to create an account of epistemic blame. According to Brown, epistemic blame consists in the belief that the person has believed or reasoned badly, and the desire that they had appreciated the relevant epistemic reason they had flouted, with this belief-desire pair giving rise to a set of affective and behavioural dispositions associated with blame. While this acts as a sufficiently intuitive answer to the conceptual question, trying to use Sher's model to build a minimal theory of epistemic blame faces two challenges.

First, we need an answer to the justification question, and it doesn't seem possible to formulate an analogous story to the one that Sher provides regarding moral blame. In investigating what makes the belief-desire pair in blame appropriate, Sher notes that the desire can't be rendered appropriate simply because it is an instance of desiring the good and disliking the bad. Otherwise, our account would lack any distinction between moral blame and 'blame' directed at other bad things like shingles or earthquakes. Sher's answer is that the desire that someone not have violated a moral principle is unlike other desires or other commitments that we might adopt, because it stems from a commitment to morality. He reasons that, because of morality's overriding, prescriptive *action*-guiding nature, a commitment to morality conceptually entails a *desire* (or, at least, something capable of motivating action) that we and other agents also act morally, with this desire then providing the dispositions that we associate with blame. If someone did not experience this desire, they would not truly be committed to morality.

The problem for Brown is that it is less clear whether epistemic reasons or principles can generate a similarly privileged kind of desire. It doesn't seem that full acceptance of an epistemic principle is conceptually bound up with having any desire that others follow that principle. Because epistemic principles are belief-guiding, my acceptance of an epistemic principle seems to commit me only to believing that I and others who are similarly situated have reason to believe in accordance with that principle. It seems perfectly coherent to maintain that I can be rationally committed to an epistemic principle without having any desire that others in fact follow that principle. This means that, although the desire that others follow epistemic norms reflects a

commitment to those epistemic norms, we lack any justification for experiencing that desire's associated dispositions.

Even if a connection between the sources of epistemic normativity and the relevant desire could be established, an objection to Sher's account remains in place. He took the dispositions stemming from the belief-desire pair to reflect a commitment to morality. In providing an appealing account of the value of blame, he thereby also provides an appealing argument for answering the conceptual question in a specific way. Because a commitment to morality entails a specific desire, and because it makes sense to react in particular ways when this desire is frustrated, we have grounds for taking those dispositions to constitute blame, and for taking these dispositions to be valuable. But a number of philosophers have noted that, because the dispositions associated with the belief-desire pair are only *contingently* associated, Sher does not succeed in showing that the dispositions, or their expression, have any value at all [Hieronymi 2008; Watson 2008; Franklin 2013]. At best, he succeeds in showing that a commitment to morality leads to a desire that other people not violate moral principles. Additionally, this leaves it open to his opponents to argue that other responses to wrongdoing, such as disappointment or sadness, are either more valuable ways to express a commitment to morality, or more characteristic of blame. In short, Sher fails to give an account of why the dispositions that he identifies should be thought of as valuable or as constituting blame. Similarly, in order to answer the justification question, Brown's account needs to explain why these dispositions or their expression can ever be valuable.

The primary aim of this paper has been to identify what we want from a theory of epistemic blame, and to outline why a number of tempting proposals are not feasible. In providing this ground-clearing, I've effectively presented grounds for being sceptical about its ever being appropriate to respond with *blame* to violations of epistemic norms *qua* epistemic norms. We started out by noting that we seem to experience blame-like responses to epistemic-norm violators, and that such violators seem to be genuinely blameworthy. But if we cannot identify what feature these responses share with moral blame, or show that they are distinguishable from responses that we experience towards frustrating events generally in terms of their character, warrant, and justification, then perhaps we should simply conclude that our initial judgments were mistaken.

In what remains of this paper, I would like to ward off this scepticism by demonstrating that it is possible to develop an account that answers our questions. We can make some progress by reconsidering one of the features that moral and epistemic blame have in common: they both target failures to respond to reasons.

5. The Agency-Cultivation Model of Blame

One account of responsibility that gives an important role to our sensitivity to reasons is Vargas's [2013] agency-cultivation model. Vargas, like many philosophers, takes agents to be morally blameworthy for wrongdoing when they had the capacity to avoid wrongdoing, where this capacity is further described in terms of reasons-responsiveness and facts about nearby possible worlds. Somewhat as Sher does, Vargas answers the conceptual question by taking moral blame to primarily be a cognitive judgment that someone is blameworthy, meaning that they are an appropriate target of a certain class of interpersonal reactions (for example, resentment, indignation,

anger), which in turn creates a disposition to engage in blaming reactions (for example, verbal condemnation, calls for censure or shame, avoidance, emotional distance, or retractions of interpersonal warmth), the expression of which communicates to them our dissatisfaction with their violation of moral standards. What distinguishes Vargas's account from other accounts of moral responsibility is the way that he justifies our moral blaming practices. This particular justificatory story not only makes the account uniquely placed to give us what we want from a theory of epistemic blame; it enables us to identify why epistemic blame seems to have some connection to moral blame.

Vargas argues that our practices of blaming and praising are justified by their ability to cultivate a particularly valuable kind of agency—an agency that is responsive to moral reasons and considerations, which is something that ‘stands out against the fabric of the universe’ [ibid.: 140]. Blame is particularly effective at cultivating a reasons-responsive agency because, for creatures like us, blaming-reactions discourage certain kinds of behaviours and habits, while praising-reactions encourage them. These blaming- and praising-reactions start as external motivators but, over time, come to be internalised, and are then experienced as intrinsically motivating. This causes agents to become more adept at governing themselves by those norms, recognising reasons relevant to those norms, and enforcing those norms when interacting with others.

This account initially looks like a ‘moral influence’ account of blame and blameworthiness, which have been extensively criticised for justifying blame towards innocent agents when doing so would influence them (for example, scapegoating), while failing to justify blaming incorrigibly culpable wrongdoers who wouldn't be influenced by such blame. However, Vargas avoids the objections ordinarily levelled at influence theories by locating questions of justification not with particular instances of blame, but at the second-order level of our *practices* of blaming.³ While cultivating agency is a forward-looking goal that justifies practices of blaming as whole, individual instances of blaming are justified by backward-looking considerations—namely, whether the agent committed wrongdoing without an excuse.⁴

To understand this justificatory story, consider an analogy. Basketball has rules against ‘travelling’ with the ball to prevent players from gaining an unfair advantage, which helps the game to stay competitive, serving the goal of providing entertainment to fans. This is a forward-looking consideration, in that we want to prevent certain kinds of behaviour so as to achieve certain ends. However, we all agree that players who travel within a game of basketball are appropriately penalised, even in cases where they travel without gaining an advantage, and even in cases where penalising them won't contribute to the level of entertainment for the fans. Thus a player can be appropriately penalised, with this penalty being appropriate in a backward-

³ This is similar to Rawls's [1955] consideration of a two-tiered utilitarian justification for punishment.

⁴ But shouldn't we want a justificatory story that points to deeper facts about whether agents really are responsible, rather than to some forward-looking teleological justification? Vargas favours finding an independent source of justification for our practices because the facts about responsibility are fiercely contested, and it seems like we have many strongly held but incompatible intuitions. Some points in favour of this particular account over its competitors are that it is naturalistic, compatibilist, treats our acting for reasons as central to our practices, is realistic about the limits and messiness of human psychology, compatible with multiple normative theories and other justifications for our responsibility practices, and preserves our intuitions about many key features of our actual practices (e.g. the role of excuses and justification). Given that we currently lack any minimally satisfying theory of epistemic blame, the fact that this model does all of the above while also answering the conceptual, warrant, and justificatory questions (argued below) counts considerably in its favour.

looking way because of the way in which the player was connected to their act of travelling, even though the justification for the system of penalties as a whole is forward-looking. Likewise, agents can be appropriate targets of blame for wrongdoing, even if blaming in a particular instance won't directly serve the justification that we have offered for the practice as a whole.

Successful internalisation of norms requires that our practices be stable across time, efficient at producing the desired effect, and compatible with our messy psychologies that are the product of our evolutionary history. These features place limits on what sorts of practices and responses could count as justified, by using an appeal to (modestly) teleological considerations. That blame is justified because it cultivates our agency but doesn't entail that other responses (such as scapegoating), which may sometimes cultivate agency, are also justified because those other responses are either not sufficiently stable, effective, or compatible with our psychologies to achieve internalisation. Blaming people who have not done anything wrong, while also failing to blame people who have done something wrong, is unlikely to make groups of people internalise moral norms. Even if it did turn out that practices of scapegoating are effective at cultivating agency, the account is modular in that it is open to us to treat the wrongness of scapegoating as a fixed normative point which our responsibility practices then have to accommodate. These features allow for some degree of flexibility between the contexts in which blame is appropriate, our choice to express blame in those contexts, and the effects of that blame that we are aiming to bring about. What matters is that the expression of blame in response to culpable wrongdoing contributes to the stability and efficacy of our moral responsibility system over time [ibid.: 175–81].⁵

Here's why this justificatory story seems especially apt for giving us what we want from a theory of epistemic blame. Just as moral blame causes us to become more adept at recognising moral reasons by causing us to internalise norms, epistemic blame is ripe for norm internalisation, too. Our desire to avoid being blamed by others provides us with motivation to check our sources and to reflect on our reasoning in order to avoid making mistakes. Indeed, psychologists reliably find that people's reasoning improves when they know that they are going to have to defend their assessments to others [Lerner and Tetlock 2003]. Of course, we don't just want to avoid being blamed; we also want to avoid being *blameworthy*. When we realise that we have made a foolish error, we often experience regret or embarrassment, and we quickly

⁵ An anonymous reviewer asks what we should say if it turned out that engaging in (say) moral teaching and counselling, rather than blaming, would cultivate agency even more effectively. First, our account says nothing about what one ought to do, all things considered, in particular situations. If, in some cases, moral counselling is better, then one ought to do this. It's just that blaming also remains justified, according to the internal logic of the practice, and the practice is justified by its overall efficacy at agency-cultivation. Second, because the account is naturalistic, the norms that are justified need to be efficient at cultivating our agency in a wide range of contexts where local norms are expected to remain stable, and compatible with messy psychologies that have been shaped by our evolutionary history. Although there are contexts where moral counselling would work better than blame, it would take some work to show that giving up our practice of blaming entirely and replacing it with a practice of counselling would work better on average. Third, it is not clear that we require our practices be maximally effective if they are to be justified, as opposed to just being sufficiently effective given the costs that a more effective system might create. Finally, we are only trying to give a plausible explanation of, and justification for, our actual existing responsibility practices, which clearly assign an important role to blame. If it turned out that a practice of counselling without blame was more effective and did not introduce unwelcome costs, then we could be justified in adopting those practices instead and considering this to be a form of moral progress. Moving between practices could carry costs, though, which would count against immediately pushing for wide-scale changes. For further details, see Vargas [2013: 172–95].

feel strong desires to avoid making similar mistakes in the future. We're also likely to ruminate on our blunders and the things that led us to commit them, and are more willing to listen to advice from others who have avoided our errors. And these responses are usually amplified when our errors attract blame from others.

Let's now examine how an agency-cultivation account can easily give us what we want from a theory of epistemic blame. This account can easily answer the conceptual question: epistemic blame consists in a cognitive judgment that someone is blameworthy—namely, that they are an appropriate target of a certain class of interpersonal reactions for violating epistemic standards—which in turn creates a disposition to engage in blaming reactions, the expression of which communicates to the agents our dissatisfaction with their violation of those standards. This is a particularly appealing answer because it explains why epistemic blame manifests in a variety of forms, and it overlaps with our conception of moral blame.

Our account can also easily answer the warrant question: blame is warranted when an agent had the capacity to recognise, and to believe in accord with, epistemic norms.⁶ This answer is also appealing because it entails that our account of epistemic blameworthiness overlaps significantly with our account of moral blameworthiness. As already noted, many philosophers writing on moral blame and moral responsibility take the possession of certain capacities to be necessary for blameworthiness, and failures to exercise those capacities to be blameworthy. It also gives an intuitive explanation for why, say, children are typically excused for forming unjustified beliefs: their diminished belief-forming capacities act as an excusing condition.

Finally, our account answers the justification question: epistemic blame is justified because the practice of blaming agents for failing to abide by epistemic norms helps them to internalise those norms, fostering a very distinctive and valuable kind of agency. Even if any particular instance of engaging in epistemic blame won't directly cultivate the wrongdoer's agency, the practice of blaming as a whole is justified by this goal. This last answer is particularly appealing because it succeeds in explaining not only how moral blame and epistemic blame are similar, but how they share an important foundation. While Vargas focuses on explaining how our capacity to respond to moral reasons is a distinctive form of agency worth cultivating, our ability to respond to reasons when forming beliefs is also especially valuable. Our rational capacities and sensitivity to reasons are commonly held by many philosophers to contribute greatly to the value of our lives, or to be the source of our inherent dignity.

Our answer to the justification question in turn answers a question that arose earlier: why is it that some responses (for example, hostility) to moral blameworthiness seem like inappropriate responses to epistemic blameworthiness, while others (for example, criticism, disappointment, non-retributive anger) remain equally appropriate in either case? The answer is that there are different kinds of considerations to which

⁶ Vargas's account might initially seem to use distinctly moral notions and thus to encounter the difficulties that I raised in section 2, because he claims that culpable agents deserve to be blamed, and that agents need to display a deficient quality of will if they are to be culpable. However, Vargas explicitly doesn't understand desert in any retributive or Pereboomian basic desert sense. His claim that an agent deserves to be blamed means only that blaming them is justified under our responsibility practices [2013: 115, 249–66]. Additionally, his inclusion of a quality-of-will condition is somewhat tentative, since what counts as a poor quality of will is determined by the reasons upon which an agent acts or fails to act. To put this another way, an agent can display deficient quality of will simply by failing to notice a consideration, which is a broader notion than the more everyday sense in which poor quality of will towards someone requires some awareness of that agent [ibid.: 234–9, especially note 2].

we are trying to sensitise agents, and different things that we are trying to communicate in signalling our dissatisfaction with their violations of our standards. When we engage in moral blame, we are not trying only to make wrongdoers recognise that they have violated a norm. We are trying to sensitise them to *what* they have done *to* someone, and what they ought to do now, after having wronged someone, in order to set things right.

When someone commits wrongdoing, they disrespect another agent's dignity, or violate the agent's legitimate demand for good will, or express something objectionable that threatens the agent's status as someone with equal moral standing. Wrongdoing also creates victims, damaging their view of the world, trust in others or shared moral norms, and their sense that their dignity is respected by others [Walker 2006], typically causing perpetrators to incur remedial duties [Piovarchy 2020]. In addition to wanting wrongdoers to recognise the norms that they've flouted, so that they comply in the future, moral blame also aims to make wrongdoers recognise the moral reasons that they have *now* to show repentance, repair relationships, compensate victims, and demonstrate a commitment to the moral norms that they have flouted [Radzik 2009]. If one is the victim, having this blame take a hostile form can also signal a certain kind of vulnerability, in that it reveals that a perpetrator's actions can hurt us, making clearer the nature of their wrongdoing and the reasons that they have to restore the relationship. Hostile blame can also carry with it a threat of action, signalling that a blamer is willing to 'strike back' if the perpetrator does not change their ways [Reis-Dennis 2019].⁷

In contrast, none of these factors is present when considering epistemic blame-worthiness, and so we do not need to sensitise those agents who violate epistemic norms to the reasons that these factors create in the moral case. While we want to blame agents for epistemic norm violations in order to cultivate their agency, violations of epistemic norms *qua* epistemic norms do not harm or disrespect agents, they do not create victims or remedial duties, and there is no need to signal vulnerability or a willingness to fight. No one has their status threatened in some way that warrants protesting, and we do not need to induce guilt *qua* a pained recognition of what the blamed has done to another agent, so as to motivate them to make restorative gestures. Moral blame and epistemic blame have some similar affective and behavioural dispositions because both aim at achieving internalisation of norms to cultivate an agency that is sensitive to reasons. But they take different dispositions to be appropriate responses to norm violations, because they aim to cultivate an agency that is sensitive to very different kinds of reasons.

6. Conclusion

This paper has argued that a number of questions need to be answered if we want to construct a theory of epistemic blame. In particular, we need an account of what epistemic blame is, when agents are epistemically blameworthy, and what justifies one's engaging in epistemic blame. By examining a number of theories of moral blame and blameworthiness, I showed that many existing conceptions of moral blame are unable to answer these questions. I also argued that Brown's [2020] recent account

⁷ Although this 'striking back' need not be aggressive; one can threaten that one will withdraw good will from an intimate partner, for instance.

fails to tell us what makes epistemic blame justified, or why epistemic blame seems similar to moral blame in some ways but different in others. Finally, I argued that Vargas's [2013] agency-cultivation model not only provides a plausible account of what epistemic blame is; it is particularly well-suited to do so. Both epistemic blame and moral blame are attributed in response to failures to act on reasons, given an agent's capacities, and this is justified because an agency that can respond to reasons (and not just moral reasons) is something worth cultivating.⁸

Disclosure Statement

No potential conflict of interest was reported by the author.

Funding

This work was funded by an Anderson Junior Research Fellowship from the University of Sydney.

ORCID

Adam Piovarchy  <http://orcid.org/0000-0002-5169-2030>

References

- Brown, Jessica 2020. What Is Epistemic Blame? *Noûs* 54/2: 389–407.
- Brown, Jessica forthcoming. Epistemically Blameworthy Belief, *Philosophical Studies*.
- Carlsson, Andreas B. 2017. Blameworthiness as Deserved Guilt, *The Journal of Ethics* 21/1: 89–115.
- Darwall, Stephen 2006. *The Second-Person Standpoint: Morality, Respect, and Accountability*, Cambridge, MA: Harvard University Press.
- Franklin, Christopher E. 2013. Valuing Blame, in *Blame: Its Nature and Norms*, ed. D. Justin Coates and Neal A. Tognazzini, New York: Oxford University Press: 207–23.
- Goldberg, Sanford C. 2017. Should Have Known, *Synthese* 194/8: 2863–94.
- Graham, Peter J. 2015. Epistemic Normativity and Social Norms, in *Epistemic Evaluation: Purposeful Epistemology*, ed. David K. Henderson and John Greco, Oxford: Oxford University Press: 247–73.
- Hieronymi, Pamela 2004. The Force and Fairness of Blame, *Philosophical Perspectives* 18: 115–48.
- Hieronymi, Pamela 2008. Sher's Defense of Blame, *Philosophical Studies* 137/1: 19–30.
- Lerner, Jennifer S. and Philip E. Tetlock 2003. Bridging Individual, Interpersonal, and Institutional Approaches to Judgment and Decision Making: The Impact of Accountability on Cognitive Bias, in *Emerging Perspectives on Judgment and Decision Research*, ed. Sandra L. Schneider and James Shanteau, Cambridge: Cambridge University Press: 431–57.
- Levy, Neil 2011. *Hard Luck: How Luck Undermines Free Will and Moral Responsibility*, Oxford: Oxford University Press.
- Littlejohn, Clayton forthcoming. A Plea for Epistemic Excuses, in *The New Evil Demon Problem*, ed. Julien Dutant and Fabian Dorsch, Oxford: Oxford University Press.
- Nelkin, Dana K. 2015. Psychopaths, Incurable Racists, and the Faces of Responsibility, *Ethics* 125/2: 357–90.
- Nelkin, Dana K. 2016. Accountability and Desert, *The Journal of Ethics* 20/1–3: 173–89.
- Pereboom, Derk 2014. *Free Will, Agency, and Meaning in Life*, Oxford: Oxford University Press.
- Piovarchy, Adam 2020. Blame in the Aftermath of Excused Wrongdoing, *Public Affairs Quarterly* 34/2: 142–68.
- Radzik, Linda 2009. *Making Amends: Atonement in Morality, Law, and Politics*, New York: Oxford University Press.

⁸ Thanks to Manuel Vargas, Sandy Boucher, Adrian Walsh, David Coady, and Isabelle Wentworth for helpful comments. Thanks, especially, to two anonymous reviewers for helping greatly to improve this paper.

- Rawls, John 1955. Two Concepts of Rules, *The Philosophical Review* 64/1: 3–32.
- Reis-Dennis, Samuel 2019. Anger: Scary Good, *Australasian Journal of Philosophy* 97/3: 451–64.
- Rettler, Lindsay 2018. In Defense of Doxastic Blame, *Synthese*, 195/5: 2205–26.
- Sher, George 2006. *In Praise of Blame*, New York: Oxford University Press.
- Smith, Angela M. 2013. Moral Blame and Moral Protest, in *Blame: Its Nature and Norms*, ed. D. Justin Coates and Neal A. Tognazzini, New York: Oxford University Press: 27–48.
- Sutton, Jonathan 2007. *Without Justification*, Cambridge, MA: The MIT Press.
- Tierney, Hannah forthcoming. Guilty Confessions, in *Oxford Studies in Agency and Responsibility*, Vol. 7, ed. David Shoemaker, Oxford: Oxford University Press.
- Vanderheiden, Steve 2016. The Obligation to Know: Information and the Burdens of Citizenship, *Ethical Theory and Moral Practice* 19/2: 297–311.
- Vargas, Manuel 2013. *Building Better Beings: A Theory of Moral Responsibility*, Oxford: Oxford University Press.
- Walker, Margaret U. 2006. *Moral Repair: Reconstructing Moral Relations after Wrongdoing*, Cambridge: Cambridge University Press.
- Watson, Gary 2008. Review of George Sher's *In Praise of Blame*, *Mind* 117/466: 515–20.
- Watson, Gary 1996. Two Faces of Responsibility, *Philosophical Topics* 24/2: 227–48.
- Williamson, Timothy 2007. On Being Justified in One's Head, in *Rationality and the Good: Critical Essays on the Ethics and Epistemology of Robert Audi*, ed. Mark Timmons, John Greco, and Alfred R. Mele: New York: Oxford University Press: 106–22.