



Vice epistemology, norm-maintenance and epistemic evasiveness

Adam Piovarchy¹ 

Received: 11 July 2022 / Accepted: 10 February 2023 / Published online: 10 March 2023
© The Author(s) 2023

Abstract

Vice epistemology studies how character traits, attitudes, or thinking styles systematically get in the way of knowledge, while doxastic responsibility is concerned with what kinds of responses are appropriate towards agents who believe badly. This paper identifies a new connection between these two fields, arguing that our propensity to take responsibility for our doxastic failures is directly relevant for vice epistemology, and in particular, understanding the social obstacles to knowledge that epistemic vices can create. This is because responses to norm violations are an important mechanism by which norms are upheld, and maintaining epistemic norms is crucial for our collective epistemic successes. This paper then identifies a new kind of vice, one which is bad precisely because of the way it undermines the epistemic norms that our blaming practices help maintain, and thus the benefits that said norms create. I call this vice epistemic evasiveness, and it concerns the attitude that one takes towards their own performance as an epistemic agent. Evasiveness is bad because it creates uncertainty about which agents are reliable, it prevents holders of this attitude from learning from their mistakes, and it signals to third parties that the norm is not being upheld, making them less likely to follow the norm.

Keywords Epistemic vice · Vice epistemology · Epistemic evasiveness · Blame · Doxastic responsibility

✉ Adam Piovarchy
adam.piovarchy@nd.edu.au

¹ Institute for Ethics and Society, The University of Notre Dame, 2007 Broadway, Sydney, NSW, Australia

1 Introduction

Obstructionist vice epistemology is concerned with understanding how character traits, attitudes, or thinking styles systematically “get in the way of knowledge” (Medina, 2013, p. 30). In contrast to mere defects which may also obstruct knowledge, epistemic vices are blameworthy, criticisable, or reprehensible. Cassam (2018, 2019) proposes that we can identify whether traits, attitudes and thinking styles count as epistemic vices by investigating whether they systematically make it less likely that the beliefs of one’s self or others will be based on evidence, undermine others’ confidence in their true beliefs (since one might reduce their confidence without abandoning the belief entirely), or deprive others of their right to be confident in their beliefs (since one might have a true belief, and not have their confidence undermined, but arrive at that belief in an unjustified manner).

Doxastic responsibility, in contrast, is concerned with whether and when agents are responsible for things like their beliefs, ignorance, traits, inferences, and attitudes (Osborne, 2021; Weatherson, 2008). In particular, it is interested in when agents form their beliefs (etc.) in a way that makes them culpable, what kinds of factors might be excusing or mitigating, and what kinds of responses towards such agents are appropriate, namely blame. It seems to be a common feature of our lives that we blame people not only for what actions or choices they make, but also what beliefs they form, what inferences they make, and what sources they trust or appeal to.

These two fields have significant overlap, in that both are interested in agents’ criticisable epistemic features and conduct. But they typically have distinct temporal features of agents’ conduct as their target of concern. Our interest in epistemic vices is primarily concerned with the traits or attitudes that lead to or manifest criticisable epistemic conduct, as well as what produces those traits or attitudes. Doxastic responsibility, in contrast, is largely concerned with how we ought to respond to agents *after* they exhibit criticisable epistemic conduct.¹ This is not to say these focuses are exclusive: we can account for why a vice is bad in virtue of what happens downstream of said vice manifesting, and factors leading up to a particular performance are relevant for assessing whether blame is apt. Still, there is a clear difference in emphasis here, which has allowed these distinct fields of inquiry to develop and operate mostly independently.

This paper has two goals. The first is to create a new bridge between the literatures on vice epistemology and doxastic responsibility.² By examining the way in which epistemic norms are upheld by holding norm violators responsible and having them take responsibility for their violations, I offer a new angle from which to consider how our epistemic successes are affected by other agents. Though vice epistemologists have illuminated a number of vices that concern an agent’s attitude towards, and ways of thinking about truth, evidence, and other agents, we should also be con-

¹ As well as laudable conduct, in the case of praiseworthiness, though blame has taken the lion’s share of attention given there are higher costs to getting our blaming responses wrong.

² Beyond the existing one provided by those investigating whether agents can be blameworthy for their epistemic vices (e.g. Battaly, 2019; Cassam, 2018, 2019).

cerned with an agent's attitudes and ways of thinking about *their own performance*.³ Because of the social, collaborative way in which knowledge is discovered, produced, transmitted and maintained by norms, and because how well norms are upheld can be affected by how agents respond to their own failures, our propensity to take or shirk responsibility is an important arena in which epistemic virtues and vices can operate. The second goal is to use this understanding to draw attention to a thus far unidentified vice. Though it is one we encounter often enough in our dealings with others, it has not yet received a name or a sustained analysis.

The paper proceeds as follows. First, I identify some important benefits that come from maintaining epistemic norms, and argue that the costs of not maintaining them leads us to sometimes treat said norms as social norms: they are the kind of thing that we prescribe and blame one another for violating. Second, I outline some challenges that communities face in upholding norms (knowing that everyone else will comply with the norm, deterring violations, and producing competence at complying with the norm), and identify a number of ways that holding others responsible for their failures and taking responsibility for our own, can overcome, or at least mitigate, these challenges. I then outline some ways that agents can respond poorly to their own performances, and argue some of these responses constitute the manifestation of a vicious kind of attitude or thinking style I call epistemic evasiveness. Evasiveness has a number of bad effects, which are unified by their tendency to undermine norm-following, both moral and epistemic. It creates uncertainty about which agents are reliable and what the relevant norms are, it prevents norm-violators from learning from their mistakes, and it signals to third parties that the norm is not being upheld, which makes them less likely to follow the norm themselves.

2 The importance of upholding norms

Epistemic norms are typically thought of as 'norms' in the sense that they characterise what counts as good reasoning. One should not form certain beliefs because doing so would e.g. be irrational, or fail to believe according to the balance of epistemic reasons, or violate the law of non-contradiction. But many epistemic norms can also be 'norms' in another sense: they are *socially* prescribed (Graham, 2015).⁴ Revealing

³ Cassam (2019) notes that arrogant agents tend to feel and think of themselves as superior to others, and Tanesini (2021) gives a sustained examination of vices that regard one's measurement of their self (e.g. narcissism, haughtiness, vanity, self-abasement, fatalism, servility), being primarily concerned with an agents' assessments of their own abilities and skill. As we shall see, I mean their performance in quite a specific sense – not one's performance relative to others, performance regarding the identification of vices in themselves, or performance as an accurate prediction of how one will do in the future, but how they have been doing with regard to meeting prescribed standards, the violation of which would entitle others to blame.

⁴ Although most of the literature on holding agents responsible has focused on moral norm violations, we clearly hold agents responsible for poor athletic performance (Shoemaker, 2022), or violating non-moral social norms (e.g. appropriate attire for occasions; Brennan et al., 2013), aesthetic norms (e.g. the blame of many Star Wars fans towards George Lucas's directive choices; Wolf, 2015, 2016), and epistemic norms (Graham, 2015; also evidenced by the recent literatures on doxastic responsibility and epistemic blame, e.g. Brown, 2020b; Chuard & Southwood, 2009; Mchugh, 2014).

that you hold inconsistent beliefs, for instance, can cause your audience to have negative responses towards you, whether that be privately thinking less of you, gossiping about your faults to others, or expressing blame to you.

It's reasonably clear why epistemic norms are often socially prescribed, in the sense that their violation can generate negative responses from others. Our epistemic achievements depend greatly upon co-ordination with others and a division of epistemic labour (Brennan, 2010; Goldberg, 2010). Since there is far too much for any one person to know, and it is too difficult for any given individual to assess what counts as good evidence on every topic, we specialize. We outsource certain roles to others, who we then rely upon. I don't know much about viruses, so I rely on public health agencies to tell me how to avoid them. I haven't been to Berlin, so I rely on someone who has been to tell me what the nightlife is like.

This specialization and outsourcing, however, not only involves many individuals. It requires a high degree of organization, e.g., laws, policies, habits, and networks. And maintaining this organisation requires certain standards. If doctors in the medical system become too unreliable, then important information that needs to be communicated (e.g. from medical researchers) fails to be transmitted effectively as people stop listening. We also need certain language norms with which to communicate this information: if our conceptions of what 'irrational' or 'knows' means are too far apart, it will be hard for me to follow your advice to rely or not rely on someone (Dogramaci, 2012; Henderson and Greco, 2015). We also need certain norms around assertion and testimony: telling someone that P, when you have no idea whether P obtains, is likely to be met negatively once their reliance on your testimony gets them into strife (Graham, 2020b).⁵

Given the benefits of co-operation and the extent to which our success depends upon others, upholding important epistemic norms can clearly have good effects, both morally and epistemically speaking, while violating them can produce bad effects.⁶ There are moral and epistemic costs of getting things wrong, and there are *many* more ways to be wrong than to be right on most topics, which means that ensuring members of our community comply with epistemic norms is very valuable. A lot is needed to produce said compliance with norms (e.g. resources, education, time, social support) and epistemologists are paying more attention to the ways such factors affect our epistemic lives. But there are some very distinctive ways that agents

⁵ Grimm (2009) argues that the potentially limitless ways in which someone might rely on our beliefs is an important source of epistemic normativity.

⁶ It has been noted that—particularly at the collective level, across long time scales—vices can contribute to knowledge in important ways, namely through the evolution of cumulative culture (Bland, 2022; Levy & Alfano, 2020). Though this doesn't show that vices encouraging norm violations are particularly beneficial within a generation, there is an important question here of when epistemic norms should be complied with, and when they can be ignored, which goes on to determine whether a given behaviour means an agent is blameworthy or not. In this paper I am setting aside the question of when norms should be complied with and under what circumstances an agent is blameworthy. All that matters is that there are clear instances in which a norm should be complied with, that we should blame agents for violating, and that our blaming practice will be justified in virtue of the bad effects that violations like these typically have. A two-tier theory of justification can handle questions readers might have here about teleological justifications; see mostly Vargas (2013), with similar endorsements from Piovarchy (2021), Rawls (1955) and Caruso and Dennett (2021).

fail to comply with norms, which communities wanting to maintain said norms need to contend with. Looking at these potential sources of failure can also help us understand the way in which norms are prescribed, and why we respond negatively to norm violations in the ways that we do.

One reason agents might not comply with norms—epistemic or otherwise—is that they simply don't *care* about the norms. They might have more to gain by explicitly violating norms, they might be susceptible to choosing short term gains at the expense of incurring longer term costs, or they might simply feel they have better things to do with their time. Another reason agents might fail to comply with norms, even if they care to some degree about them, is that they might think other agents are unlikely to comply, and thus resign themselves to inaction.⁷ Even if an agent starts out caring about the norm, if they believe others don't care, or won't follow it, or that complying with the norm is imposing an unreasonable cost on them that others are not bearing, then they will be unlikely to comply with the norm (Bicchieri, 2005, 2016; Brennan et al., 2013). It is a well-known feature of collective action problems that agents who want similar things can nevertheless find themselves unable to achieve them because the fact that others want the same thing is not common knowledge, which is often required to co-ordinate (Aumann, 1976; Geanakoplos, 1992). With enough examples of others not following the norm, they might even come to think that there is a norm of *not* doing what they previously took to be prescribed.

Even if agents want to uphold norms, and know that other agents want to too, a third reason they might fail to comply is that they lack the ability to recognise what the relevant norms are or how to comply with them. Norms admit of exceptions, context matters greatly, and epistemic norms can be particularly difficult to follow because we do not always know when we are violating them. The wishful thinker, for instance, often does not know they are engaging in criticisable wishful thinking. As Cassam (2019) identifies, some vices are *stealthy*: they avoid their own detection. To learn they are performing below our shared standards, agents need to have this brought to their attention. They require sustained feedback from others, across a variety of settings, in which they can learn from their mistakes and improve. Even when one desires to comply with a norm, learning how to do so can be difficult. Though we tend to think of norms as relatively obvious, like 'don't hold inconsistent beliefs' or 'don't harm', what *counts* as 'harm' or 'holding inconsistent beliefs' in practice requires a lot of knowledge and training.⁸ Norm-following is a skilled competence requiring intelligent flexibility, especially since norms can change over time and contexts. Producing competence is the kind of thing that requires regular feedback on performance.

It's against this background of norms and the pressures that can cause agents to violate them that we can now see the importance of holding agents responsible for

⁷ This might seem inapt given we are primarily interested in beliefs rather than action, which are not under our volitional control, but there are many things agents can do which indirectly affect the beliefs they end up with, such as choosing to go to college or deciding to ask a question. Whether the belief is brought about by direct or indirect means does not matter for my argument.

⁸ Should the latter seem easy to readers, consider the difficulties of recognising whether a pair of beliefs are in fact inconsistent once vagueness, ambiguity, under determination, and lack of clarity regarding reference classes have been introduced.

their failures to comply with important norms, as well as the importance of taking responsibility for our own failures.

In the moral and doxastic responsibility literatures, the main way we hold agents responsible for their norm violations is considered to be by blaming them. This achieves quite a lot. First, blaming behaviours, being unpleasant for the target, increase the costs of norm violations and thus deter them (Fehr & Fischbacher, 2004). We reduce our opinion of the agent, we feel and express disapproval, we tell others about what has happened, affecting the violator's reputation, we demand that the target answer for what they have done, and we withdraw good will or epistemic trust (Boult, 2021) until we get evidence the agent can be counted on to comply with the norm in the future. Second, blame provides assurance to others that the norm matters, will be complied with, and that any violations will be taken seriously. Blame is a costly (and thus credible) signal that one is committed to the norm and will act in ways to uphold it (Shoemaker & Vargas, 2021). Finally, blame helps develop competence. Blame acts as a form of feedback, letting the target know what kinds of things count as violations in what contexts, and sensitising them to what they have done. When we blame, we draw our target's attention to the norm, its importance, and what impact their violation had. We also draw them into a conversation about the significance of their conduct; blame "includes an RSVP" (Darwall, 2006, p. 40), increasing their understanding of how the other person saw their conduct, producing alignment between the blamer and blamee (Fricker, 2016). Additionally, as agents become aware of others' expectations and continually try to comply with the norm—even if this begins as simply a means of avoiding negative treatment—people eventually begin to internalise the norm, experiencing it as inherently motivating (Brennan et al., 2013; Sripada & Stich, 2006; Ensminger & Henrich, 2014; Gavrillets & Richerson, 2017; Tomasello, 2020).⁹ Internalisation is a particularly useful thing to create, because it means that an agent's motivation to comply with the norm and develop their competences outlasts their incentives to comply, which can be contingent, unstable, and hard to accurately assess. Successful internalisation in turn greatly

⁹ As Piovarchy (2021) puts it, "we don't just want to avoid being blamed; we also want to avoid being blameworthy" (p. 801). A substantial body of evidence shows we possess "a suite of genetically evolved cognitive mechanisms for rapidly perceiving local norms and internalizing them" (Chudek et al., 2013, p. 443); see also Chudek and Henrich (2011) and Gelfand and Jackson (2016). Focusing on three distinct effects may seem untidy for those who find theoretical simplicity a philosophical virtue, but we have a principled basis for doing so because they are united by a common function (upholding norms) and each arises to solve a particular problem that social agents like us face in maintaining norms. Signaling helps provide assurance of co-operation which helps solve co-ordination problems and stabilize a signaling system among senders and receivers (Smith & Harper, 2003; Zahavi & Zahavi, 1999), as well as providing assurance of sanctions which themselves solve the problem of defectors. Sanctions are particularly important as initially small numbers of norm violations, if left unchecked, lead to a raft of violations in public goods games because while most people are conditional co-operators, the amount of co-operation each agent expects from others can be quite different (Fehr & Gächter, 2000). Sanctions can stave off initial, relatively minor violations, and thereby prevent cascades of violations from occurring. Feedback is an obvious way in which norms are learned by beings like us (especially as children), and internalisation further reduces the chances that people will violate the norm by creating motivation that outlasts their incentives to comply. Additionally, this story is naturalistic and consistent with research on mixed motive public goods games and co-ordination problems as understood in economics, psychology, anthropology and sociology (Bicchieri et al., 2018; Kelly & Setman, 2021). For a more thorough explanation of how these features illuminate our folk concept of blame, see Piovarchy (ms).

helps co-ordination, as agents find it easier to know who they can count on. Knowing who cares about the same norms as me is far more reliable means of knowing who I can rely on to comply with them than assessing what contingent incentives each agent has to comply.¹⁰

Having seen that third-person blaming responses play an important role in maintaining norm following by providing deterrence, assurance, and developing competence and internalisation, it should now also be clear that first-person responses to norm violations matter too, and why seeing an agent take responsibility for what they have done often tempers our blaming responses towards them. If I take responsibility for my poor conduct, acknowledging that it was unacceptable, I provide assurance to others that I care about the norm, despite my lapse, and that this caring means I can be counted on to comply with it in the future. I also communicate that blame or some forms of negative treatment are legitimate, given the importance of the norm. Finally, if I am honestly concerned for the norm, and thus for my conduct, I will be more receptive to things which make me less likely to violate it again in the future.

Now that we have outlined the importance of epistemic norms, some particular ways that norms can fail to be complied with, and the ways in which third- and first-party blame can help sustain norm compliance given those common causes of failure, we are finally in a position to consider the vice I have in mind.¹¹

3 Epistemic evasiveness: a field guide

Williams (1981) famously describes a lorry driver who, though taking every care he can while driving, accidentally hits and kills a young child. It is difficult to know what our attitude towards such a driver should be. On the one hand, given he conducted himself impeccably, it seems he should be off the hook. If he were to blame himself, we would console him, try to convince him that blame is unfitting, and remind him that he is not culpable for the child's death. On the other hand, if the driver were to feel nothing, or to only feel the equivalent of what an uninvolved bystander would

¹⁰ Assuming all candidates have some minimal competence with the norm.

¹¹ An interesting question here is whether, when we hold agents responsible for their believing badly, this is in response to their epistemic norm violation *qua epistemic* norm violation, or *qua* failure to follow moral norms to form certain beliefs. I think the answer is both. That our epistemic norms sustain our collective epistemic successes, and that we often have moral reasons to follow said norms, are both good justifications for blaming agents who violate said norms. Our concern with epistemic vices need not be limited to their ability to impede knowledge; Kotsonis (2022) for instance, argues that excessive curiosity, gossip and excessive inquisitiveness are *bona fide* epistemic vices, even though their bad effects primarily consist in epistemic wrongs to others (such as knowing things about an agent that we ought not know; cf. testimonial injustice and believing things about an agent's competence or sincerity that you ought not believe) rather than knowledge obstruction. But for the purposes of this paper we do not have to settle whether norms regarding belief are ever purely epistemic or always grounded in moral considerations. If one insists, we can have distinct systems for moral blame directed at moral norm violations (of which doxastic conduct will count), and epistemic blame directed at epistemic norm violations (Boult, 2021; Brown, 2020a; Piovarchy, 2021). So long as by holding agents responsible we signal our commitment to the norm, sanction the violation, and develop the agent's competence with the norm, we will have a principled connection between our holding responsible practices and the maintenance of our epistemic norms to investigate.

feel, we would not accept this. We would blame him and draw his attention to the fact that *he* has killed a child. This seems to reveal we have a somewhat inconsistent attitude towards the driver's actions.

In *The Moral of Moral Luck*, Susan Wolf (2001) suggests there exists a “nameless virtue” which can help us make sense of our mixed feelings towards cases like the driver.¹² She argues that we sometimes expect people to recognize that they have negative effects on the world that go beyond what they are at fault for, and that this involvement can call for certain responses, such as guilt, agent-regret, and being concerned for others’ ‘psychic health’.

Wolf’s nameless virtue provides a useful contrast to the vice I have in mind. In motivating us to think such a virtue exists, she asks us to consider how the lorry driver would seem to us if he did not display it:

What is problematic is his failure, beyond this, to take the consequences of his faultiness to have consequences for him, to be a significant part of his personal history... He reveals a sense of himself... as one who is, at least in principle, distinct from his effects on the world, whose real quality and value, for better and for worse, is at best impurely indicated but not at all constituted by the goods and the harms, the successes and the failures that comprise his life in the physical world.

(Wolf, 2001, p. 12–13)

It seems this kind of take is often displayed even outside of contentious cases of faultless wrongdoing warranting agent-regret. In more routine instances of fault where someone has clearly violated a norm, it is all-too-common that people display a reluctance to acknowledge the facts about their performance as the rest of the community sees them. There is something particularly criticisable about people who don’t see themselves as implicated in outcomes which they very much brought about, or who don’t see themselves as having performed certain kinds of actions at all.

Agents can refuse to accept that they have broken a norm (or that they have broken it in the relevant sense, e.g., culpably, unreasonably, thoughtlessly) in a few ways. Perhaps they simply do not care all that much about who is responsible because they do not care about the norm itself (even if they know, on some level, that the norm exists). Or perhaps they profess to care, but always find themselves with an excuse. Sure, they acknowledge the norm matters, but *their* situation was extenuating, and how could they have been expected to know better anyway? Or perhaps they focus on shifting the blame—it’s not that they had an excuse, it’s that someone else was at fault, possibly the blamer. Or, they were justified in behaving as they did (though if it turns out they weren’t justified, then they would also have an excuse).

Another way that people can refuse to accept responsibility is to simply deny their actions had any effect at all. Psychologists have shown that it is not uncommon for people

¹² A minor exegetical point: in Wolf’s example, the driver is guilty of a minor degree of negligence, but no more than thousands of other drivers every day who we do not think are especially blameworthy. This is because her target is people who do recognize and take responsibility for their faults, but who draw a sharp line between faults that they are responsible for and those that they are not, which they feel no guilt or agent-regret over. My target is broader; some agents fail to acknowledge their effects on the world even when they are at fault.

to conflate moral responsibility and causal responsibility in accounting for events (Shaver, 2012); rather than rejecting responsibility by admitting our action had an effect, but arguing it was justified, we instead argue *our* action did not produce *that* effect at all, or was not an action under a particular description. Consider, for instance, the slogan ‘guns don’t kill people; people kill people’. The slogan is drawing attention to the fact that guns themselves are causally inert, typically requiring an agent to pull the trigger. But rather than merely implying (correctly) that both guns and agents are necessary for guns to be used to kill (with either being a potential causal lever we can use to decrease gun deaths), the slogan is more often used to imply (falsely) that causal responsibility lies exhaustively with the agent, which in turn absolves anything which causally contributed to more guns being available in the first place (e.g. policies which increase the availability of guns).

The vice I have in mind consists in a certain take on one’s self, in which the agent fails to see themselves as an agent who performs actions and who produces effects in the world. In addition to ordinary norm-breaking among peers, the previous example shows this vice can manifest in people’s professional roles, in politics, and in the creation of laws and policies. While the agent might agree to statements like ‘I am an agent’, they have a habit of interpreting their criticisable epistemic conduct (and perhaps also the norm violations of other members of their in-group) as being almost entirely the product of external forces, constrained options, incapacity, duress, ignorance, or as not being an action, omission, or behaviour at all. Some examples may help make this clearer:

I just work here Angela works for a company and publishes technically true but highly misleading information about the health effects of the company’s products, and it is reasonably foreseeable that this misinformation will cause harm to citizens. When she is blamed by others for the amount of doubt she has caused among the public, she sincerely believes she isn’t at fault. After all, she never professed to be a scientist, she didn’t lie, and people should know to look for multiple sources on contentious topics.

Not my fault Kelly has very good evidence that if she runs in the election, this will have the effect of splitting the vote among her and the other pro-union candidate due to there being a first-past-the-post voting system. Kelly runs, the vote splits, and the pro-business candidate wins and implements anti-union policies. When later asked what caused the policies to come about, Kelly’s causal explanation focuses on the company candidate choosing certain policies, the other pro-union candidate not dropping out of the race, and voters being too ignorant or busy to turn out in higher numbers.

Regional manager Michael appeals to a clearly unreliable source in creating a work policy, which ends up creating significant costs for his employees. Once he learns of his error, he ‘updates’ the policy which has the effect of covering up his mistake, feeling no need to acknowledge the error or costs this imposed. If blamed, he would argue that there was good evidence it was reliable, even though he wouldn’t consider this evidence to be reliable if someone else in his position had made the error.

The politicians we deserve Kevin is an epistemic slacker, and doesn’t care to do any epistemic work that requires effort. He believes that a certain political candidate will deliver great policies, simply because the candidate says they will and it is easy to not

question them. However, Kevin has ample evidence that the candidate is a liar, unreliable, incompetent, and that his promises are infeasible. The candidate is elected, and breaks all of their promises. Kevin feels outraged at the actions of the politician, blaming them for deceiving voters.

Murphy's observed regularity Dwight commits to a project, agreeing to complete it in a short time frame. However, he falls prey to the planning fallacy, imagining everything going right with no set-backs. Inevitably, as almost anyone in his position would have predicted and accounted for, there are set-backs. He completes the project well behind schedule, and is later asked about completing another, similar project. Dwight commits again to completing the project in a time-frame almost identical to the first estimate, because he once again fails to take into account the possibility of unforeseen setbacks. When the error of his previous estimate is pointed out to him, he is dismissive of the possibility that this reflects badly on him or that he could have done otherwise. As a result, he fails to take this to show that he ought to reduce his confidence in his ability to complete this project on time.

In all of these cases, the agent fails to suitably recognise their epistemic failure and role in producing bad states of affairs. However, this failure is importantly distinct from the failure one might exhibit when simply unaware of the norm, or justifiably distracted by other task demands, or perhaps even incapable of complying with the norm due to, e.g., being a child. These agents are competent enough at recognizing the norm (in that they would recognize when other agents broke them, and hold other agents to similar norms), capable of complying with said norm (in that they likely would comply with it if their incentives were different), but exhibit a *resistance* to recognising, or accurately representing, their own performance regarding it. While some 'recognize' their performance in that they are perhaps vaguely aware that they made an error, we want more than this. A morally blind psychopath can 'recognise' moral norms in that they will agree some norms are moral, or bad to break, but many people (e.g. Shoemaker, 2007) think morally blind agents fail to have a proper understanding of what that means, and do not count as recognising the importance of said norms and what they entail.¹³

This resistance is not simply that exhibited by the dishonest agent who knows they are at fault and consciously decides to pretend they are not. All too often, people like the agents above sincerely believe that they are not at fault. Alternatively, they may be somewhat uncertain whether they are at fault, but treat not being at fault as the default, held-until-irrefutably-proven-otherwise stance. Even if they would acknowledge uncertainty upon reflection, when being accused or blamed they immediately opt for finding counterarguments, disputing previously uncontentious assumptions, raising the epistemic standards required for opponents' claims to be treated as common ground, or diverting the conversation to only thinly-related matters, such as the conduct of the blamer. In the moral domain, when we want people to accept responsibility, we don't just want them to say 'sorry' and go through the motions as

¹³ Whether psychopaths are in fact morally blind is disputed (Jurjako & Malatesti, 2018), but what matters is that mere ability to agree a norm is moral is not sufficient.

if they recognized what they did. We want them to *actually* recognize what they did. Likewise, I believe that when it comes to doxastic/epistemic norms we want agents to recognize that they believed badly, not to merely treat their believing badly as common ground.

When we think about relevant counterfactuals and what kinds of circumstances result in a failure to take responsibility, the best explanation of their patterns of responses is that these agents—perhaps unconsciously—recognize there’s a chance that others might think the agent is at fault, are motivated to prevent this from occurring, and this produces a range of epistemic behaviours that will help avert this.¹⁴ They exhibit a certain attitude towards their own agency and conduct, one which resists regarding the role that they play in the lives of others as negative, primarily by doing whatever is necessary to *evade* believing they ought to be held to account by others. Let’s call this vice ‘epistemic evasiveness’.

One of the (morally) worst manifestations of evasiveness comes in the form of victim-blaming from perpetrators: “Look what you made me do”. By putting the blame and conversational focus on the victim, this takes attention away from the actions of the perpetrator, forces a change of topic, introduces doubt (given the perpetrator is usually something like an epistemic peer and there is clear disagreement despite similar evidence and reasoning capacities) and confusion (given it can be hard to keep track of what’s common ground as sub-disputes open up), and can simply wear down the blamer into giving up as the dialectic proceeds without any resolution. In more aggressive forms, it can induce guilt, shame, or fear in its target, who acquiesces to the perpetrator’s new narrative, which then makes it even harder to establish the perpetrator’s fault as common ground, as this would first require having a discussion about the victim’s change of position. It is also not hard to see how evasiveness can manifest in credibility deficits and testimonial injustice, as one refuses to believe they are culpable by way of believing that others are incompetent or insincere.

Again, I think it is worth emphasizing that victim-blaming is not always simple dishonesty or a tactical feigning of outrage (though this can also occur). Perpetrators who victim-blame often experience genuine outrage and hostility, genuinely perceive their victim as having been the one who has committed a transgression, and genuinely feel like they had no other choice, or that their actions were required. Nevertheless, it would be grossly inaccurate to explain these perceptions by saying that they have simply made a mistake, or are merely unaware of their own role. They do not experience themselves as possessed by a wholly alien, external and unfamiliar force, which presumably would result in confusion (rather than anger and rumination), pleading (rather than self-righteousness and justifying), and agent-regret (rather than guilt).¹⁵

Is evasiveness a trait, an attitude, or a thinking style? There are grounds for thinking each or all of these apply. People can be robustly disposed to deflect blame and

¹⁴ A very large body of research demonstrates that we regularly fall prey to motivated reasoning, where the incentives we have to believe a proposition end up causing us to believe that proposition independent of the evidence (Kahan, 2015; Molden & Higgins, 2012).

¹⁵ A more speculative take is that victim-blaming from third parties may also sometimes be caused by evasiveness, since third parties may have duties to create conditions which prevent perpetrators from abusing others. By assigning fault to the victim, this implies that fault does not lie elsewhere, and thus that the third-party does not have any such duty they have been neglecting.

fail to recognise the significance of their own conduct that it makes sense to attribute a trait to them. Concerning attitudes, evasive agents might take their own behaviour to be beyond reproach, or perceive others to be norm violators when engaging in blame-shifting, or take a dismissive attitude to blame from others, perceiving it to be inapt. They might even take said blame to be a threat, or perceive their own conduct to be righteously justified, resulting in anger towards the blamer. We could bolster this answer by pointing out that attitudes of other vices also manifest in a variety of ways, and this is what explains the heterogeneity in evasive agents (e.g., an arrogant attitude can consist in a dismissive attitude towards others, overconfidence towards one's own abilities, or a feeling of benevolence towards one's audience that is actually quite patronizing, despite the differences between these). Finally, we can also consider evasiveness to be a way of thinking. Just as conspiracy thinking is a thinking style which "attempts to tie together seemingly unrelated events and focuses on errant data" (Cassam, 2019, p. 70), evasive thinking attempts to avoid the conclusion that the agent has performed poorly and is now an appropriate target of blame from others. In this respect, evasive thinking is goal-directed: it is directed at avoiding a certain conclusion which affects what kinds of possibilities are searched for. It can be more or less effortful, and concerns what the agent thinks about, affects the weighing of evidence, and affects the drawing of inferences.

Having just drawn our attention to a possible vice, let me now (in the spirit of intellectual honesty) make the unusual move of admitting that there may be grounds for thinking such a vice does not exist. Because evasiveness consists in a *failure* to take responsibility, and there are many different ways one can fail do this (e.g., making an excuse vs. arguing for justification) which can themselves be produced by a range of motivations (e.g., a desire to preserve one's ego vs. a desire to not face social exclusion), one might argue that the factors giving rise to evasive thinking and behaviours are too heterogeneous to consider a single trait, attitude, or thinking style to be at play. One might propose that the phenomena we are interested in can be accounted for using Tanesini's (2021, p. 12) recent taxonomy of epistemic vices regarding mismeasurement of the self.¹⁶

		Attitude Valence			
		Positive		Negative	
		Social comparison present	Social comparison absent	Social comparison absent	Social comparison present
Motivation	Ego Defense	Haughtiness	Arrogance	Servility	Timidity
	Social Acceptance	Vanity	Narcissism	Self-Abasement	Fatalism

It is plausible some of these could underlie the evasive behaviours and beliefs we are interested in. Admitting fault might be a threat to someone's ego in a variety of ways. It might threaten the haughty individual's sense that they are better than others. Since having performed poorly would make it clear they are not better than others, the agent responds by finding ways to believe that they in fact did not perform poorly.

¹⁶ Table modified slightly to remove contrastive virtues.

Alternatively, having everyone be aware of one's poor previous performance is likely to threaten one's desire for social acceptance. If a vain person relishes the thought of everyone admiring them as a paragon of good reasoning, they will be motivated to discredit any arguments to the contrary. In contrast, a servile person might know that they will not be confident enough to mount a sustained defence should their poor performance come to light (due to their very servility) and as a result, hide any evidence implying they have performed poorly.

Even if it is the case that evasive attitudes and thinking are ultimately reducible to the manifestation of other vices of self-measurement, it is quite useful to have a term for vicious thinking and attitudes which concern an agent's own norm-compliance. As I will argue below, because of the way our epistemic successes are socially dependent, and because we maintain our successes through the upholding of norms, how individuals respond to norm violations is a particularly important juncture in which norms are maintained. This means we will sometimes have a principled reason for wanting to identify, analyse, and communicate about epistemically bad thinking, attitudes and traits that interrupt these practices, and the particular kinds of harms that evasiveness produces. In such instances, it is reasonable to be less concerned with whether this failure is explanatorily reducible to the manifestation of some other trait like arrogance or haughtiness, and whether the agent also happens to, e.g., view other agents as beneath them. For the sake of precision though, if it turns out that evasiveness is not a distinct vice, my argument can instead be considered to be identifying an important sphere in which vices of mismeasurement operate, a number of particular harms that they produce, and a new set of grounds for considering said attitudes to be epistemic vices. It is to these harms that I now turn.

4 The harms of evasiveness

It is easy enough to argue that epistemic evasiveness is vicious by thinking of intuitively bad effects that evasive agents can cause. But we can find a firmer theoretical foundation from which to categorize and investigate particular kinds of bad effects by considering the role of norms and what is required to uphold them that I introduced earlier. This will also clarify what makes the vice distinctly epistemic. Though evasiveness can have bad moral effects, if we accept that there are epistemic norms which play a crucial role maintaining our collective epistemic successes, and that one's beliefs about and attitudes towards their own performance qua epistemic agent is an important means by which following of these norms can be fostered or undermined (as we shall see), then we have a principled basis for considering evasiveness to be an epistemic vice.

The central bad-making feature of evasiveness concerns how, if allowed to proliferate, it has corrosive effects on our epistemic environments. In particular, it systematically gets in the way of knowledge precisely by exacerbating threats to our system of norms that we uphold, and thus the systems' ensuing benefits. The first way this happens is, by failing to own up, the agent fails to learn from their mistakes. If the agent genuinely believes they were not at fault, they are thus going to believe that they have no reason to change their habits, or seek more information next time,

or double check their sources, as they did nothing epistemically wrong. This substantially weakens the way that attempts to hold agents responsible can sensitize them to norms, which philosophers have argued helps ‘scaffold’ our agency over time (McGeer, 2013). The agent also won’t be receptive to any assistance from others, since there is nothing calling for improvement.

Failure to acknowledge fault thus means that one is at a higher risk of re-violating. If the evasive agent argues they did nothing wrong, then they also add an additional source of pressure to themselves to act similarly in the future for the sake of consistency. If they were to later instead conform with the norm by ϕ -ing in a relevantly similar situation, this inconsistency between their previous argument that not ϕ -ing was permissible would be noticed and prompt a challenge from others. This potential cost thus creates an incentive to continually ‘double down’ on taking such actions to be faultless.

The second primary effect of evasiveness is that it makes it much harder for third parties to *keep track* of who is a reliable norm follower and who is not. It often is not clear when someone has violated an epistemic norm, because it is not clear what evidence they had at the time they violated it, whether there were any mitigating factors in play, or what reasonable expectations were in place. We can more easily recognize it in politicians because, having so much of their lives in the spotlight, there is a large body of historical evidence we can use in our assessments, but this is clearly not available for most norm violators. If an epistemic norm violator does not own up to their failures, third parties can be unsure whether the agent actually violated a norm, or whether it was the blamer who has made an error. And given how epistemic successes depend upon identifying who is reliable in what domains, being unable to keep track of reliability is bad in a number of ways. Third parties might erroneously maintain their trust in the agent, and continue deferring to them when they ought not to. Alternatively, though they may appropriately reduce their confidence in the unreliable agent, their uncertainty about who was right may mean they opt to also reduce their confidence in the blaming party, or anyone accusing the unreliable agent of violating the norm, meaning that other agents suffer a credibility deficit.

Finally, evasiveness also makes it much more likely that other agents will violate the norm too. In failing to own up, the agent signals to others that they do not think that *that* kind of conduct was unacceptable. And since much of how we learn about norms comes from other agents, this makes third parties in turn think that the norm is not that important, or easily defeated, or that *that* kind of behaviour doesn’t count as violating the norm. A large body of research from social psychology demonstrates that our willingness to follow norms is strongly influenced by whether we think other agents are following the norm too (Gelfand & Harrington, 2015). As Schwitzgebel (2019) puts it, most of us aim at being merely ‘morally mediocre’: we “actively adjust our moral behavior up or down so as to approximately match the morality of others in the groups with which we identify” (p. 350).¹⁷ Given the way we also care about non-moral norms (e.g. sporting, social) and blame violators of said norms, our standards for what behaviours are epistemically acceptable are also affected by how

¹⁷ This is compatible with the common finding that most people report believing they are more moral, competent or intelligent than average (Hoorens, 1993), since individuals won’t identify with most people.

we think our peers are doing. None of us want to be stupid, but what each of us count as stupid or below standard tends to be determined by the groups we identify with (Corcoran et al., 2011; Mussweiler, 2020).¹⁸

Additionally, agents can be less likely to maintain good epistemic standards because they might see the evasive agent as having an unfair advantage in public disagreements. Consider an agent who doesn't need to refer to evidence or reason, and who deploys whatever rhetoric is most convincing to each audience. This forces anyone who disagrees to spend considerable time refuting each of these points, preventing them from outlining their positive proposals on the subject matter.¹⁹ One way to overcome this trap is to establish that the agent is unreliable in general, which acts as higher-order evidence that most of their claims cannot be trusted. But if the evasive agent won't admit fault, and muddies the facts of their performance, this option becomes much harder to pull off. This, in turn, means that agents who started out being intellectually honest have a strong incentive to stoop to the other agents' level, or use other forms of rhetoric, in order to get their point across.²⁰

Can't we just ignore evasive agents? This is much easier said than done. First, identifying evasive agents is harder than it seems. Even viciously evasive agents will find themselves in circumstances in which they did not culpably violate a norm, and there will be plenty of times where they do comply with the relevant norms. (By comparison, viciously dishonest people still regularly tell the truth, particularly on trivial matters, and this makes their behaviour rather noisy, increasing the difficulty of working out when someone is being dishonest). Once we've identified evasive agents, it's also very hard to ignore them entirely. Often they play some important role in our epistemic networks, and it is difficult to find others to replace them entirely. Even if I ignore them, unless everyone else does the same, the evasive agent can still have effects on other members of my network who can then affect me. Additionally, our susceptibility to norm-following is not always consciously represented as doing so because everyone else is, and even when it is, we are very susceptible to having our perceptions of what is appropriate be affected by cues from others, as many studies in the situationist psychology literature famously show (Doris, 2002; Latané and Darley, 1970, pp. 252–253).

To get a grip on the harms of this vice (and to see why it is easier to talk of one vice, rather than several distinct vices manifesting in a similar way), it will help to consider its relation to another vice known as epistemic insouciance. Cassam (2018) argues that insouciance is an attitude which “implies, and is partly constituted by, a marked lack of intellectual seriousness, and flippancy about basing one's views on expert opinion or what the evidence shows. It is a casualness or indifference to the truth and to the need to base one's opinions on the relevant facts” (p. 2). The insouciant agent is indifferent towards knowledge, evidence and inquiry. Since one's own

¹⁸ Note, for example, that IQ is explicitly defined in relative terms.

¹⁹ This is commonly referred to as Brandolini's Law: “The amount of energy needed to refute bullshit is an order of magnitude larger than is needed to produce it” (Williamson, 2016). Many readers will also be familiar with the ‘Gish Gallop’, in which an arguer overwhelms their opponents with misrepresentations and half-truths that take far more time to refute than to say.

²⁰ If this agent starts breaking the norm other onlookers will be more likely to do so too; cf. *supra* fn. 9 on how initial violations by some lead to more violations by others.

rate of success seems to be relevant to inquiry (e.g. inquiry is more successful if one learns from their mistakes), a question here is whether we should think of evasiveness as being similar to insouciance, while simply taking one's one epistemic performance as its object.

I think that trying to understand what is criticisable about evasive agents through the lens of insouciance alone would be too narrow. Our criticism of the agents in the earlier thought experiments isn't simply that these agents don't care about their performance. Indeed, some would be adamant that their performance was *not* subpar, which is evidence they do care. But we might say they don't care about their own performance as an epistemic agent, or care for the right reasons, or they don't care about the norm *de re*. They care about their performance or the norm as seen by others, or as a mark of status, or as means of avoiding negative treatment, but they don't care sufficiently about it as something which reflects on their agency *qua* norm follower. If they were genuinely concerned for that norm, such that they cared sufficiently about the norm being upheld for its own sake, then they would want to recognize their own failure and react appropriately for the sake of the norm.

The distinctness of insouciance and evasiveness (though they are very co-morbid) can be demonstrated by noting that many of the harms of insouciance identified by Cassam are mitigated—and thus not fully accounted for—when not accompanied by evasiveness. We can see this by realizing that, in principle, there could be agents who are insouciant with regard to truth and evidence, but very un-evasive when it comes to assessing their own performance.²¹ Though they don't care about the truth, the bad effects of this are going to be quite limited if they know they don't care and signal this to others. When others have the opportunity to recognize that the insouciant agent is not to be trusted, it is much easier to ignore them and stem the bad effects that Cassam identifies.

Cassam focuses on the insouciance of a number of politicians leading up to the United Kingdom's vote to leave the European Union in 2016. While these politicians were insouciant regarding the truth and evidence about the merits of leaving the EU, it is not so clear that they were insouciant about their own performance. Even if they openly dismissed the relevance of facts to their case, if one of them had said 'it actually doesn't matter whether I've gotten things right previously, or what *my* own track record of reliability is, sometimes I'll stick to what I said and sometimes I'll deny ever saying it. I personally don't care whether I've continually gotten things wrong in the past and I never try to find out', and then went back to being insouciant, it seems their persuasiveness would have been undermined significantly. Note, for instance, the tendency for even insouciant politicians to argue, upon having failed to deliver on a promise, *not* that they failed, or that they promised it but the facts about what they

²¹ We can also imagine separation between insouciance and evasiveness in the other direction. Someone might care very much about truth and basing their beliefs on evidence, while paying little concern for whether they were blameworthy when they previously got things wrong. They might care indirectly about their past performance insofar as getting things wrong means one is at risk of getting things wrong again, but such a concern is primarily forward-looking: the fault only matters insofar as it affects what they believe now and in the future. In contrast, the object of most people's responses to their norm violations are backward-looking: rumination, regret, and a focus on fixing what one *did* (as opposed to, say, fixing things to increase future utility or outcomes).

promised don't matter, but that, e.g., they didn't promise *that*, or that circumstances changed, or that someone else was at fault. When it comes to the topic of their own performance, they are quite resolute that there is a fact to be had, namely, that their performance was not below standard. Famously, the Leave campaign had a giant bus that read "We send the EU £350 million a week – let's fund our NHS instead". When this amount was not delivered to the NHS after the vote, Leave campaigner Iain Duncan Smith said, "I never said that during the course of the election... What we actually said was a significant amount of it would go to the NHS... There was talk about it going to the NHS, but there are other bits and pieces like agriculture, which is part of the process. That is the divide up. It was never the total" (Perraudin, 2016). Though head of UKIP Nigel Farage said, "[the Leave campaign] made a mistake in doing that", he was adamant "I would never have made that claim... It wasn't one of my adverts – I can assure you!" (Stone, 2016).

For their arguments to have an effect on voters, politicians need to maintain some veneer of trustworthiness, if not based in competence regarding truth and evidence then perhaps based in their own optimism, sharing voter's values, or having a sincere intention to work through any problems that came up in the future. Admitting that nothing they said was reliable, or displaying a sustained indifference to their own performance would have been interpreted as higher-order evidence by voters that their insouciant claims could not be relied upon, perhaps more successfully than that provided by experts' conflicting claims.

Agents who are insouciant and evasive are much more damaging than agents who are only one or the other. They don't care about evidence, truth or inquiry, *and* they have nothing to apologise for. They have no reason to change their ways, or to be concerned with other people's need to assess their past performance, and thus little motive to stop polluting the epistemic commons.

5 Conclusion

Epistemic evasiveness is bad in a number of ways. It causes agents to fail to learn from their mistakes. It makes it harder for other agents to keep track of who is reliable. And it makes it much more likely that other agents will violate norms too. I've argued that if allowed to proliferate, it can undermine the benefits that our system of epistemic norms provides. At least, if I later turn out to be wrong, I ought to admit it.

Funding Open Access funding enabled and organized by CAUL and its Member Institutions
Open Access funding enabled and organized by CAUL and its Member Institutions

Declarations

Conflict of interest none (paper has not been shared).

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this

article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Aumann, R. J. (1976). Agreeing to disagree. *The Annals of Statistics*, 4(6), 1236–1239.
- Battaly, H. (2019). Vice epistemology has a responsibility problem. *Philosophical Issues*, 29(1), 24–36. <https://doi.org/10.1111/phils.12138>.
- Bicchieri, C., Muldoon, R., Sontuoso, A. S., & Norms, *The Stanford Encyclopedia of Philosophy* (Winter 2018 Edition), Edward N. Zalta (ed.), <https://plato.stanford.edu/archives/win2018/entries/social-norms/>
- Bicchieri, C. (2005). *The Grammar of Society: the Nature and Dynamics of Social norms*. Cambridge University Press.
- Bicchieri, C. (2014). Norms, Conventions, and the power of expectations. In N. Cartwright, & E. Montuschi (Eds.), *The philosophy of Social Science: a new introduction* (pp. 208–229). Oxford University Press.
- Bicchieri, C. (2016). *Norms in the Wild: how to diagnose, measure, and change social norms*. Oxford University Press USA.
- Bland, S. (2022). In defence of epistemic vices. *Synthese*, 200(1), 1–22.
- Boult, C. (2021). There is a distinctively epistemic kind of blame. *Philosophy and Phenomenological Research*, 103(3), 518–534.
- Brennan, G. (2010). The Division of Epistemic Labour. *Analyse & Kritik*, 32(2), 231–246.
- Brennan, G., Eriksson, L., Goodin, R. E., & Southwood, N. (2013). *Explaining Norms (paperback)*: Oxford University Press UK.
- Brown, J. (2020a). Epistemically blameworthy belief. *Philosophical Studies*, 177(12), 3595–3614.
- Brown, J. (2020b). What is Epistemic Blame? *Noûs*, 54(2), 389–407.
- Caruso, G. D., & Dennett, D. C. (2021). *Just deserts: debating Free Will*. Polity.
- Cassam, Q. (2018). Epistemic insouciance. *Journal of Philosophical Research*, 43, 1–20.
- Cassam, Q. (2019). *Vices of the mind: from the intellectual to the political*. Oxford University Press.
- Chuard, P., & Southwood, N. (2009). Epistemic norms without voluntary control. *Noûs*, 43(4), 599–632.
- Chudek, M., & Henrich, J. (2011). Culture–gene coevolution, norm-psychology and the emergence of human prosociality. *Trends in cognitive sciences*, 15(5), 218–226.
- Chudek, M., Zhao, W., & Henrich, J. (2013). Culture-gene coevolution, large-scale cooperation, and the shaping of human social psychology. In K. Sterelny, R. Joyce, B. Calcott, & B. Fraser (Eds.), *Cooperation and its evolution*. MIT Press.
- Corcoran, K., Crusius, J., & Mussweiler, T. (2011). Social comparison: motives, standards, and mechanisms. In D. Chadee (Ed.), *Theories in social psychology* (pp. 119–139). Wiley-Blackwell.
- Darwall, S. L. (2006). *The second-person standpoint: morality, respect, and accountability*. Harvard University Press.
- Dogramaci, S. (2012). Reverse Engineering Epistemic evaluations. *Philosophy and Phenomenological Research*, 84(3), 513–530.
- Doris, J. M. (2002). *Lack of Character: personality and Moral Behavior*. Cambridge University Press.
- Ensminger, J., & Henrich, J. (Eds.). (2014). *Experimenting with social norms: Fairness and punishment in cross-cultural perspective*. Russell Sage Foundation.
- Fehr, E., & Gächter, S. (2000). Cooperation and punishment in public goods experiments. *American Economic Review*, 90(4), 980–994.
- Fehr, E., & Fischbacher, U. (2004). Third-party punishment and social norms. *Evolution and human behavior*, 25(2), 63–87.
- Fricker, M. (2016). What's the point of blame? A paradigm based explanation. *Noûs*, 50(1), 165–183.
- Gavrilets, S., & Richerson, P. J. (2017). Collective action and the evolution of social norm internalization. *Proceedings of the National Academy of Sciences*, 114(23), 6068–6073.
- Geanakoplos, J. (1992). Common knowledge. *Journal of Economic Perspectives*, 6(4), 53–82.

- Gelfand, M. J., & Harrington, J. R. (2015). The motivational force of descriptive norms: for whom and when are descriptive norms most predictive of behavior? *Journal of Cross-Cultural Psychology*, *46*(10), 1273–1278.
- Gelfand, M. J., & Jackson, J. C. (2016). From one mind to many: the emerging science of cultural norms. *Current Opinion in Psychology*, *8*, 175–181.
- Goldberg, S. C. (2010). *Relying on others: an essay in Epistemology*. Oxford University Press.
- Graham, P. J. (2015). Epistemic normativity and social norms. In David, Henderson, & J. Greco (Eds.), *Epistemic evaluation: Purposeful Epistemology* (pp. 247–273). Oxford University Press.
- Graham, P. (2020a). Assertions, Handicaps, and social norms. *Episteme*, *17*(3), 349–363.
- Graham, P. (2020b). The function of assertion and social norms. In S. Goldberg (Ed.), *The Oxford Handbook of assertion* (pp. 727–748). Oxford University Press.
- Grimm, S. R. (2009). Epistemic Normativity. In Adrian Haddock, Alan Millar & Duncan Pritchard (eds.), *Epistemic Value*. Oxford: Oxford University Press. pp. 243–264.
- Henderson, David & Greco, John (2015). Introduction: The Point and Purpose of Epistemic Evaluation. In David K. Henderson & John Greco (eds.), *Epistemic Evaluation: Purposeful Epistemology*. Oxford University Press. pp. 1–28.
- Horens, V. (1993). Self-enhancement and superiority biases in social comparison. *European review of social psychology*, *4*(1), 113–139.
- Jurjako, M., & Malatesti, L. (2018). Neuropsychology and the criminal responsibility of psychopaths: reconsidering the evidence. *Erkenntnis*, *83*(5), 1003–1025.
- Kahan, D. M. (2015). The politically motivated reasoning paradigm. *Emerging Trends in Social & Behavioral Sciences*, *Forthcoming*.
- Kelly, D., & Setman, S. The Psychology of Normative Cognition, *The Stanford Encyclopedia of Philosophy* (Spring 2021 Edition), Edward N. Zalta (ed.), <https://plato.stanford.edu/archives/spr2021/entries/psychology-normative-cognition/>.
- Kotsonis, A. (2022). A novel understanding of the nature of epistemic vice. *Synthese*, *200*(1), 1–16.
- Latané, B., & Darley, J. M. (1970). *The unresponsive bystander: why doesn't he help?* Appleton-Century Crofts.
- Levy, N., & Alfano, M. (2020). Knowledge from vice: deeply social epistemology. *Mind*, *129*(515), 887–915.
- McGeer, V. (2013). Civilizing blame. In D. J. Coates, & N. A. Tognazzini (Eds.), *Blame: its nature and norms* (pp. 162–188). Oxford University Press.
- McHugh, C. (2014). Exercising Doxastic Freedom. *Philosophy and Phenomenological Research*, *88*(1), 1–37.
- Medina, J. (2013). *The epistemology of resistance: gender and racial oppression, epistemic injustice, and resistant imaginations*. Oxford University.
- Molden, D. C., & Higgins, E. T. (Eds.). (2012). *Motivated thinking*: Oxford University Press.
- Mussweiler, T. (2020). How social comparison affects the self: the selective accessibility mechanism. In J. Suls, R. L. Collins, & L. Wheeler (Eds.), *Social comparison, judgment, and behavior* (pp. 32–51). Oxford University Press.
- Osborne, R. C. (2021). What do we epistemically owe to each other? A reply to Basu. *Philosophical Studies*, *178*(3), 1005–1022.
- Perraudin, F. (2016, 27 June). Iain Duncan Smith backtracks on leave side's £350m NHS claim. *The Guardian*. Retrieved from <https://www.theguardian.com/politics/2016/jun/26/eu-referendum-brexit-vote-leave-iain-duncan-smith-nhs>
- Piovarchy, A. (2021). What do we want from a theory of Epistemic blame? *Australasian Journal of Philosophy*, *99*(4), 791–805.
- Piovarchy, A. (manuscript) (Ed.). Signalling, Sanctioning and Sensitising: How to Uphold Norms with Blame.
- Rawls, J. (1955). Two concepts of rules. *Philosophical Review*, *64*(1), 3–32.
- Schwitzgebel, E. (2019). Aiming for Moral Mediocrity. *Res Philosophica*, *96*(3), 347–368.
- Shaver, K. G. (2012). *The attribution of blame: causality, responsibility, and blameworthiness*. Springer Science & Business Media.
- Shoemaker, D. (2007). Moral address, moral responsibility, and the boundaries of the moral community. *Ethics*, *118*(1), 70–108.
- Shoemaker, D. (2022). The trials and tribulations of Tom Brady. In A. Carlsson (Ed.), *Self-blame and Moral responsibility* (pp. 28–47). Cambridge University Press.

- Shoemaker, D., & Vargas, M. (2021). Moral torch fishing: a signaling theory of blame. *Nous*, 55(3), 581–602. <https://doi.org/10.1111/nous.12316>.
- Smith, J. M., & Harper, D. (2003). *Animal signals*. Oxford University Press.
- Sripada, C., & Stich, S. (2006). A Framework for the psychology of norms. In P. Carruthers, S. Laurence, & S. P. Stich (Eds.), *The innate mind, volume 2: culture and cognition*. Oxford University Press.
- Stone, J. (2016, 25 June). Nigel Farage backtracks on Leave campaign's '£350m for the NHS' pledge hours after result. *The Independent*. Retrieved from <https://www.independent.co.uk/news/uk/politics/eu-referendum-result-nigel-farage-nhs-pledge-disowns-350-million-pounds-a7099906.html>
- Tanesini, Alessandra (2021). *The mismeasure of the self: a study in vice epistemology*. Oxford University Press.
- Tomasello, M. (2020). The moral psychology of obligation. *Behavioral and Brain Sciences*, 43, 1–33.
- Vargas, M. R. (2013). *Building Better Beings: a theory of Moral responsibility*. Oxford University Press.
- Weatherston, B. (2008). Deontology and Descartes's demon. *Journal of Philosophy*, 105(9), 540–569.
- Williams, B. A. O. (1981). *Moral Luck: philosophical Papers, 1973–1980*. Cambridge University Press.
- Williamson, P. (2016). Take the time and effort to correct misinformation. *Nature*, 540(7632), 171–171.
- Wolf, S. (2001). The Moral of Moral Luck. *Philosophic Exchange*, 31(1), 1–16.
- Wolf, S. (2015). Responsibility, Moral and otherwise. *Inquiry : A Journal Of Medical Care Organization, Provision And Financing*, 58(2), 127–142.
- Wolf, S. (2016). Aesthetic responsibility. *The Amherst Lecture in Philosophy*, 11, 1–25.
- Zahavi, A., & Zahavi, A. (1999). *The handicap principle: a missing piece of Darwin's puzzle*. Oxford University Press.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.