



Epistemic Responsibility and Criminal Negligence

Alexander Greenberg¹

© The Author(s) 2019

Abstract

We seem to be responsible for our beliefs in a distinctively epistemic way. We often hold each other to account for the beliefs that we hold. We do this by criticising other believers as ‘gullible’ or ‘biased’, and by trying to persuade others to revise their beliefs. But responsibility for belief looks hard to understand because we seem to lack control over our beliefs. In this paper, I argue that we can make progress in our understanding of responsibility for belief by thinking about it in parallel with another kind of responsibility: legal responsibility for criminal negligence. Specifically, I argue that that a popular account of responsibility for belief, which grounds it in belief’s reasons-responsiveness, faces a problem analogous to one faced by H.L.A. Hart’s influential capacity-based account of culpability. This points towards a more promising account of responsibility of belief, though, if we draw on accounts of negligence that improve on Hart’s. Broadly speaking, the account of negligence that improves on Hart’s account grounds culpability in a (lack of) concern for others’ interests, whereas my account of epistemic responsibility grounds responsibility for belief in a (lack of) concern for the truth.

Keywords Belief · Responsibility for Belief · Epistemology · Negligence · H.L.A. Hart

We seem to be responsible for our beliefs in a distinctively epistemic way. We often hold each other to account for the beliefs that we hold. We do this by criticising other believers as ‘gullible’ or ‘biased’, and by trying to persuade others to revise their beliefs. But responsibility for belief looks hard to understand because we seem to lack control over our beliefs. In this paper, I argue that we can make progress in our understanding of responsibility for belief by thinking about it in parallel with another kind of responsibility: legal responsibility for criminal negligence.

✉ Alexander Greenberg
alexanderdavidgreenberg@gmail.com

¹ Department of Philosophy, University College London, 19 Gordon Square,
London WC1H 0AG, UK

I start by outlining the puzzle about responsibility for belief in more detail (Sect. 1). I then give some options for responding to it (Sect. 2). I identify the most plausible of these options as a ‘reasons-responsiveness’ model, but I argue that the reasons-responsiveness model nevertheless fails to explain *which* unjustified or irrational beliefs we are responsible for. I then draw a parallel between this problem for the reasons-responsiveness model and a problem faced by H.L.A. Hart’s important and influential account of criminal negligence (Sect. 3), and describe how Hart’s account has been adjusted to overcome this problem. Finally (Sects. 4, 5), I apply a parallel fix to the reasons-responsiveness model of epistemic responsibility. In brief, the account of negligence that improves on Hart’s account grounds culpability in a (lack of) concern for others’ interests, whereas my account of epistemic responsibility grounds responsibility for belief in a (lack of) concern for the truth.

1 The Puzzle of Epistemic Responsibility

We can understand the puzzle created by responsibility for belief as an inconsistent triad, three plausible but jointly incompatible claims:

- (1) We are responsible for our beliefs.
- (2) If one lacks control over something, one cannot be responsible for it.
- (3) We lack control over what we believe.

Each of these three claims is initially plausible.

(1) finds support in the various ways in which we epistemically evaluate one another’s beliefs. First, we make epistemic *criticisms* of one another, describing others as ‘biased’, ‘gullible’, or ‘short-sighted’ in virtue of the fact that they hold certain beliefs. Second, we offer epistemic *advice*, telling people, for example, not to believe everything they read on the internet. Third, we engage in epistemic *persuasion*, trying to get people to share our beliefs or come to believe what we think they should. Rather generally, it doesn’t make sense to criticise, advise, or persuade people in relation to things for which they are clearly *not* responsible. For instance, it wouldn’t typically make sense to criticise someone for her height, or to advise or persuade her to be shorter or taller. These phenomena of epistemic evaluation, therefore, suggest that we are responsible for our beliefs.

(2) also seems plausible, given that showing that one lacks control in relation to some behaviour typically excuses responsibility for that behaviour. People with Tourette syndrome, for example, are excused from responsibility for swearing when this is a result of their illness, and thus something over which they do not have control.

(3) seems initially plausible too. We don’t have the same kind of control over what we believe as over what we do. As William Alston points out, while one can do things if offered a reward, no level of reward can induce one to form a belief. A £100,000 cheque might get you to, say, send me a stamped addressed envelope; but it could not get you to believe, for example, that the U.S. is still a colony of the U.K.

(Alston 1988, 263). This suggests that, whereas we have control over what we do, we lack control over what we believe.

So each of (1), (2), and (3) look initially plausible. But since they are jointly inconsistent, one of them has to go. Figuring out which claim to reject is the puzzle of epistemic responsibility.¹

2 An Opinionated Map of the Options

We can understand the different options for understanding epistemic responsibility as different ways of denying either (1), (2), or (3). The most common ways of denying each of these claims are given in the table below:

Deny (1)	(a) Scepticism about responsibility for belief
Deny (2)	(b) Recast responsibility for belief, for example, as role-responsibility
Deny (3)	(c) Indirect control over belief, via gathering evidence, cultivating intellectual virtues, etc. (d) Direct , but non-voluntary , control over belief, in virtue of belief's reasons-responsiveness

I'll argue that each of these options is unsatisfactory. Of these options, grounding responsibility for belief in reasons-responsiveness—option (d)—is the most promising. I'll argue that this account still faces a problem, but the problem it faces points towards a more plausible account of responsibility for belief.

The main account that denies (1)—option (a)—is explicit scepticism about epistemic responsibility (Alston 1988, 259, 264; cf. Levy 2007).

The problem with this strategy is that it brings with it the demand to either provide an *error theory* or an *alternative explanation* of the epistemic evaluation phenomena we canvassed above. The former would explain why it seems appropriate to hold each other to account for what we believe when, in fact, it isn't appropriate, whilst the latter would explain the epistemic evaluation phenomena without appealing to epistemic responsibility. But sceptics about epistemic responsibility (e.g., Alston 1985, 1988; Levy 2007) have not yet provided either an error theory or an alternative explanation of our epistemic evaluations.

¹ The puzzle I've outlined differs from a related argument given by Alston, who claims that our lack of control over belief means that we should reject epistemic 'oughts'—i.e., claims about what we ought to believe (1985, 64ff.; 1988). Alston claims that 'ought implies can' means epistemic 'oughts' require direct voluntary control over our beliefs (1985, 64; 1988, 259). However, Alston's rejection of epistemic 'oughts' is unconvincing as no plausible version of 'ought implies can' creates a problem here. 'Ought implies can' does not state that one can only be obliged to ϕ if one can ϕ *voluntarily*, just that one can only be obliged to ϕ if one *can* ϕ (for this point, see Chuard and Southwood 2009, §5; McHugh 2012, §3.2). And it's plausible that we *can* believe what epistemic 'oughts' prescribe. Without a further principle, 'ought implies can voluntarily', there is no problem. But such a principle looks false. We say, for example, 'you ought to feel ashamed of yourself' even though we can't *voluntarily* feel ashamed of ourselves (cf. Chuard and Southwood 2009, 620). For this reason, we should instead think that lack of control over belief primarily creates a clash with *responsibility* for belief, because lack of control does intuitively excuse responsibility.

The most obvious way to deny **(2)**—option **(b)**—is to *recast* epistemic responsibility as a kind of responsibility that is consistent with lack of control. An example is what Hart calls ‘role-responsibility’—the responsibilities a person bears in virtue of having a certain role, such as the responsibility a ship’s captain has *qua* captain for the condition of his ship (1968, 212–214). The parallel idea in the epistemic case would be that we are, *qua* believers, responsible for the ‘epistemic condition’ of our beliefs, i.e., for whether they are justified or rational.²

The problem with this approach is that it simply pushes the question back a step. Consider the captain’s role-responsibility for the condition of his ship. Surely this depends on his capacity to be responsible for performing certain actions required to carry out the requirements of that role, such as repairing the ship. Likewise in the epistemic case, being role-responsible for the epistemic condition of our beliefs requires that we are capable of being responsible for whether our beliefs are justified or rational. But this is the very responsibility we were trying to explain in the first place.

Let’s move on to the strategy of denying **(3)**—arguing that we do, in fact, have sufficient control over what we believe to ground epistemic responsibility. Philosophers attempt to deny **(3)** in two main ways.

Proponents of option **(c)** point out that we can control what we believe *indirectly*, by acting in ways aimed at influencing our beliefs. For example, we gather evidence and we cultivate our intellectual virtues, such as our skills of evidence assessment.³

An indirect control model fails because the kinds of indirect control appealed to cannot explain responsibility for how our beliefs are formed in *response* to the evidence, a central way we seem responsible for our beliefs. Being good or bad at *gathering* evidence isn’t directly relevant to whether one’s beliefs are good or bad *responses* to the evidence. If someone responds to the evidence in a biased or selective manner, gathering more evidence isn’t going to help. And cultivating skills of evidence assessment looks too *distantly* related to whether a belief is a good response to the evidence to ground responsibility. Indirect control over belief thus can’t ground responsibility for how our beliefs are formed in response to the evidence, which is a central form of epistemic responsibility.

Proponents of option **(d)** ground responsibility in *direct* control over belief. These philosophers do not accept ‘doxastic voluntarism’, the view that we have the same kind of voluntary control over our beliefs as we have over our actions.⁴ They instead claim that we have a direct, though non-voluntary, control over our beliefs. This

² The closest to a defence of this view is given by Feldman, though he primarily defends the claim that epistemic ‘oughts’ are ‘role oughts’, i.e., claims that people ought to do things because they hold certain roles. In response to Alston’s argument discussed in the previous footnote, Feldman claims such ‘oughts’ are consistent with lack of control over belief (2000, 676; 2008, 350–351; for a similar view, see Chrisman 2008, 358ff.).

³ For this approach, see Heil (1992), Nottelmann (2007, Chaps. 10–11), Peels (2017, Chap. 3).

⁴ I won’t consider doxastic voluntarism, as it doesn’t do justice to the intuitions behind **(3)**, such as the fact that one can’t believe something for a reward. In any case, while some defend doxastic voluntarism, the views they defend often end up being very close to the reasons-responsiveness view discussed below (Steup 2000, 2012; Ryan 2003); or they only claim that we have voluntary control over belief in some circumstances—such as when the evidence isn’t decisive (Ginet 2001; Nickel 2010)—and such voluntary control could only account for responsibility for an implausibly restrictive set of beliefs.

control is understood as embodied in the fact that belief is an attitude that is responsive to reasons.

Many have defended this kind of account, and the reasons-responsiveness involved has been unpacked in a variety of ways. Some have understood it in terms of *sensitivity* to reasons, i.e., a belief that *p* is reasons-responsive just in case one would revise it if one no longer had reason to believe that *p* (or if one no longer took oneself to have reason to believe that *p*).⁵ Others have understood the relevant reasons-responsiveness in terms of beliefs resulting from reasons-responsive *processes*. Reasons-responsive processes are those that tend to output beliefs that are sensitive to epistemic reasons, examples of such processes being perception, memory, and conscious reasoning. On this understanding, a belief is reasons-responsive just in case it's formed or revised by a reasons-responsive process.⁶ I will focus on the latter kind of view, though the objection I'll make applies to the former as well.

Reasons-responsiveness in either of these two forms is understood by these philosophers to be a form of control that we have over our beliefs. This is *not* understood to amount to *voluntary* control—i.e., the claim is not that our beliefs are the products of choice or decision—but a distinctive form of non-voluntary control. Why it is supposed to amount to control is not always clear. Some claim that it shows that there is a doxastic parallel to the ability to do otherwise, i.e., the ability to believe otherwise if we have reason to (Pettit and Smith 1996, 442, 444–446). Others point to what we say about beliefs that are not responsive to reasons, such as psychotic or delusional beliefs. Given that we say of psychotic and delusional believers that they 'can't help it' or that their rational faculties are 'taken over', some suggest that this indicates reasons-responsiveness is a form of control (McCormick 2011, 175–176; McHugh 2013, 144).⁷

In my view, the reasons-responsiveness account is clearly the most plausible option I've canvassed so far. Unlike scepticism, the 'role-responsibility' idea, and the indirect control model, it genuinely attempts to account for the phenomenon to be explained, our seeming responsibility for whether our beliefs are well formed. Unlike doxastic voluntarism, it appeals to a capacity we uncontroversially have, the

⁵ For views of this kind, see Pettit and Smith (1996), Scanlon (1998, 18–22), Smith (2005, esp. 264–270), Hieronymi (2006, 2008). Some of those of who defend doxastic voluntarism (in particular, Steup and Ryan—see previous footnote) also defend views involving very similar claims, but add the claim that this suffices for voluntary control over our beliefs.

⁶ This view has been defended by McCormick (2011) and McHugh (2013, 2014), both of whom draw on Fischer and Ravizza's reasons-responsiveness account of *moral* responsibility for *actions* (1998). McCormick and McHugh claim that responsibility for belief also requires *ownership* over the reasons-responsive process, which results in the belief. Ownership of a process, as McHugh describes, involves "see[ing] oneself as responsible—as appropriately held responsible—for its outputs" (McHugh 2013, 142; cf. McCormick 2011, 173), and is a notion developed by Fischer and Ravizza *viz.* moral responsibility (1998, Chap. 8).

⁷ Most of those who ground epistemic responsibility in reasons-responsiveness claim that it is a form of control. An exception is Owens, who claims that reasons-responsiveness suffices for responsibility but *not* control (2000, 123–126). If Owens is correct, then the reasons-responsiveness view will amount to another way of denying (2)—i.e., denying that responsibility requires control. For my own part, I'm somewhat ambivalent about whether reasons-responsiveness is best thought of as a form of control, but my objection applies whether or not it is.

capacity of responding to reasons. But the reasons-responsiveness model nevertheless fails to fully explain responsibility for belief. In particular, it can't account which *non*-reasons-responsive beliefs we are responsible for.

The issue is that we are responsible for some of our beliefs that are influenced by non-reasons-responsive processes, but not others. Non-reasons-responsive beliefs for which we *are* responsible include beliefs resulting from wishful thinking (i.e., belief-formation motivated by what the believer wants to be the case), and from confirmation bias (i.e., reasoning biased towards confirming the believer's initial hunch). Non-reasons-responsive beliefs for which we are plausibly *not* responsible include certain beliefs symptomatic of mental illness. People who are plausibly not responsible for such beliefs include, for example, a schizophrenic who believes in a conspiracy theory, when this is a symptom of schizophrenia, or a depressive who believes that she is worthless or unloved, when this is a symptom of depression.

Another kind of example of non-reasons-responsive beliefs for which we are plausibly not responsible is provided by beliefs resulting from abusive social situations. One such example is what is known as 'gaslighting'. This is a form of psychological manipulation paradigmatically carried out by an abusive partner, who makes the abused partner doubt the reliability of her own perceptions and memories. This is typically done by the abusive partner challenging her recollections, saying things like 'you're very tired', 'you're not at your best', 'you've been very forgetful recently'. A victim of gaslighting might believe that her memory is faulty, but plausibly she is not responsible for this belief, and shouldn't be criticised for holding it. Other examples are beliefs formed under the influence of propaganda.⁸ For example, consider someone who, having grown up in the Soviet Union under Stalin, believes that Western criticism of the Moscow show trials is merely part of a capitalist plot to subvert communism. Again it's plausible that such a person is not responsible for her belief, given the malign influence of Soviet propaganda on such a person's sense of who to trust and ability to assess the evidence. It's plausible that, in these cases, the belief in question is influenced by a non-reasons-responsive process and that the believer is not responsible for it. And, while I admit that intuitions may differ, it's plausible that there are at least some abusive social situations that excuse responsibility for belief, but in which the belief is influenced by a non-reasons-responsive process.

This gives us the following set of cases:

Beliefs influenced by Non-Reasons-Responsive Processes

Beliefs for which we are responsible:	Wishful thinking; confirmation bias
Beliefs for which we are not responsible:	Mental illness; gaslighting; propaganda

In short, reasons-responsiveness and epistemic responsibility come apart. This undermines the claim that epistemic responsibility is grounded in belief's reasons-responsiveness.

⁸ I am grateful to an anonymous referee for suggesting this kind of example to me.

The defender of the reasons-responsiveness view might claim here that there is a sense in which, or a degree to which, in the first set of cases, the beliefs in question originate in a reasons-responsive process, such as reasoning—even if this process is influenced by non-reasons-responsive processes (wishful thinking; confirmation bias). If mere origination in a reason-responsive process suffices for responsibility, this would mean—correctly—that we are responsible for the first set of beliefs.

But this doesn't provide a good answer to my objection, because nothing prevents us from saying exactly the same about the second set of cases. Reasoning isn't completely absent for a schizophrenic, a depressive, someone subjected to gaslighting, or someone growing up subjected to propaganda. Thus, if we understand the reasons-responsiveness model as holding that mere origination in a reasons-responsive process is sufficient for responsibility, it gives us the wrong result in the second set of cases, implausibly classifying the mentally ill and victims of gaslighting and propaganda as responsible for their beliefs. So, our intuitions about responsibility don't track reasons-responsiveness; reasons-responsiveness can't ground responsibility for belief.⁹

3 The Model of Negligence Liability

This section introduces Hart's account of criminal negligence and the key problem it faces. People act negligently, according to the standard account, when they take a risk that meets two conditions, (i) the risk is unjustified, and (ii) they are unaware of the risk, but should have been. These two features are clearly reflected in the American Model Penal Code's definition of negligence (American Law Institute 1962):

A person acts negligently with respect to a material element of an offense when he should be aware of a substantial and unjustifiable risk that the material element exists or will result from his conduct. The risk must be of such a nature and degree that the actor's failure to perceive it, considering the nature and purpose of his conduct and the circumstances known to him, involves a gross deviation from the standard of care that a reasonable person would observe in the actor's situation. (§2.02(2)(d))

In this definition, what determines that one *should* be aware of the relevant risk is unpacked in terms of a *standard of care*, specifically the level of care a reasonable person would observe in the circumstances.

In practice, cases of criminal negligence tend to involve defendants who fail to take precautions against risk because they fail to be aware of that risk. For example, in *R v. Adomako* (1995), a landmark case in English and Welsh criminal law,

⁹ One might think an appeal to an additional ownership condition (see footnote 6) might deal with my cases, i.e., one might claim that the believers see themselves as responsible for, and thereby owning, the processes in the first set of cases but not the latter. But this is implausible. Wishful thinking is not plausibly owned. Those who engage in it do not see themselves as responsible for wishful thinking identified as such.

Adomako was an anaesthetist convicted of gross negligence manslaughter for causing his patient's death because he failed to notice the obvious evidence that the oxygen tube had become disconnected. Another example is the much discussed case of *State v. Williams* (1971), in which a Native American couple were convicted of manslaughter for failing to seek medical attention for their obviously seriously ill son, who went on to die, because they thought he only had a toothache (and, at least in part, because they feared the authorities would take their son away).¹⁰ In both of these two cases, the defendants failed to take precautions against risks because they were unaware—but should have been aware—of those risks.

There is broad, if not complete, convergence across Anglo-American jurisdictions on negligence necessarily involving lack of awareness—as the Model Penal Code definition requires—and distinct from *recklessness* in this respect, with the latter necessarily involving awareness of risk (Stark 2016, Chap. 2).¹¹ I will assume this standard account of what distinguishes negligence from recklessness, as it naturally leads on to the objection to criminal negligence that will be my focus. This objection, which—following Michael Moore and Heidi Hurd (2011, 150)—we can call 'the standard objection', is that criminal culpability—and the prospect of state punishment—is inappropriate if the defendant lacks awareness.

My focus is on this objection because Hart's account of criminal culpability for negligence was developed in response to it (specifically, the early version in Turner 1936, 37–48). I'll argue that, while Hart's account provides a good answer to some versions of the standard objection, it faces an objection that parallels the one I raised for reasons-responsiveness accounts of responsibility for belief. I will also, along the way, argue that the standard objection to criminal negligence fails. Lack of awareness does not, I'll argue, make criminal culpability—and the prospect of criminal punishment—inappropriate. This does not, of course, suffice for a defence of criminal culpability for negligence, just an answer to a key objection to it.

I will focus specifically on criminal negligence—i.e., negligence that makes one liable for punishment—rather than ordinary or tort negligence—i.e., negligence that makes one liable to pay compensation. A key difference from tort negligence which merits mention is that criminal negligence sometimes requires that the deviation from the reasonable standard of care is 'gross'—i.e., very serious. This is included in the Model Penal Code's definition of negligence, but is not universally required across Anglo-American jurisdictions. In English law, for example, gross negligence manslaughter requires a gross deviation from the standard of care,¹² whereas the offence of careless, or inconsiderate, driving does not.¹³ My focus is on criminal negligence specifically because it is more controversial among legal philosophers,

¹⁰ This case is controversial, and we shouldn't presume that a justification of criminal culpability for negligence should justify it in this case (see footnote 22 below).

¹¹ Some dispute that the recklessness/negligence distinction *should* be drawn in this way, and claim some inadvertent wrongdoing can amount to the more serious category of recklessness (e.g., Duff 1990, 157, 165–167; for a reply in defence of the standard account, see Stark 2016, 219–225).

¹² See *R v. Bateman* (1927), at 11–12.

¹³ Road Traffic Act 1988, §3ZA. For more examples, see Simester et al. (2016, 160–161); Stark (2016, 29).

evident in the fact that criminal negligence sceptics sometimes accept that tort negligence liability is justified.¹⁴

3.1 The Standard Objection to Criminal Negligence

The standard objection to criminal negligence, in brief, is that that criminal culpability requires awareness of wrongdoing, but that this is—by definition—absent in cases of negligence (Turner 1936, 37–48; Williams 1961, 122–123; Hall 1963). A recent version of the standard objection, given by Larry Alexander and Kimberly Ferzan, frames it in terms of *control*. Alexander and Ferzan claim that negligent actors cannot be culpable because culpability requires having control over what one does. And having control over whether one takes a risk, they claim, requires awareness that one is taking that risk, which rules out culpability for negligence, as negligence precludes awareness. They illustrate this point with the fictional example of Sam and Ruth, a self-absorbed couple who are bathing their young child, when some guests turn up unexpectedly early for the party they are throwing. Sam and Ruth go to greet their guests, initially planning on returning to the bathtub in a few moments. But they get so absorbed in greeting the guests that they forget about the child, who drowns in the bathtub (Alexander and Ferzan 2009, 77).

Alexander and Ferzan claim that, even despite Sam’s and Ruth’s genuine moral flaws, they cannot be held culpable for their child’s death because it was caused by their forgetting, and their forgetting was not something over which they had the relevant kind of control. Once the thought of their child “slipped out of their minds”, Alexander and Ferzan write, “they had no power to retrieve it. They were at the mercy of its popping back into their minds, which it did not” (2009, 78). On this version of the standard objection, the fundamental reason why unawareness excludes culpability is that unawareness excludes control.

Alexander and Ferzan don’t explicitly define control. But it is clear that they specifically have in mind an ability to ‘guide’ one’s conduct in accordance with the law, as outlined in the following passage:

Because the purpose of the criminal law is to prevent harm by giving us reasons to act and to refrain from acting, the criminal law does not reach the negligent actor *at the time he undertakes the negligent act*. At that time, the negligent actor is not aware that her action unjustifiably risks causing harm, and

¹⁴ J.W.C. Turner, an early criminal negligence sceptic, says this explicitly: “although ... negligence may be blameworthy and may ground civil liability, it is at the present day not sufficient to amount to mens rea in crimes at common law” (1936, 39). In the contemporary literature, the situation is complicated by underlying debates about tort law, in particular, by the question of whether tort law’s purpose is primarily economic, i.e., to efficiently redistribute economic burdens (see, e.g., Posner 1972), or whether the purpose of tort law is to enact corrective justice, i.e., for wrongdoers to correct wrongs (see, e.g., Weinrib 1995; Beever 2007, Chap. 2). Contemporary criminal negligence sceptics sometimes argue that if the latter understanding of tort law is correct, tort negligence liability is unjustified (see Hurd 2014; cf. Alexander and Ferzan 2014, 415).

thus cannot be guided to avoid creating that risk by the injunction to avoid creating unjustifiable risks. (Alexander and Ferzan 2009, 80)

This passage illustrates the kind of control that, according to Alexander and Ferzan, the negligent defendant lacks. The relevant kind of control is an ability for one's conduct to be guided by facts concerning potential harm. Alexander and Ferzan's claim is that negligent defendants like Sam and Ruth lack this ability because they are unaware that they are taking a risk; because they have forgotten about their child, they cannot, according to Alexander and Ferzan, be guided by a demand to not put their child at risk.

Alexander and Ferzan's understanding of control, given that one lacks it merely if one lacks awareness that one is taking a risk, is very restrictive. We can find in Hart, I'll suggest, an argument to the effect that it is too restrictive. Lack of awareness that one is taking a risk does not plausibly result, on its own, in the lack of control that Alexander and Ferzan allege it does.

3.2 Hart's Account of Negligence

Hart answers the standard objection to criminal negligence by claiming that it relies on an overly restrictive understanding of *mens rea*—the aspect of the defendant's psychology that makes him culpable. Instead of understanding *mens rea* purely in terms of the defendant's awareness of the consequences of his actions, Hart argues that we should understand it in terms of the defendant's *capacities* to do what the law requires. This gives us Hart's two-part 'capacity test' for criminal culpability (1961, 154):

- (i) "Did the accused fail to take those precautions which any reasonable man with normal capacities would in the circumstances have taken?"
- (ii) "Could the accused, given his mental and physical capacities, have taken those precautions?"

Condition (i) is objective, because it doesn't take into account the defendant's capacities; it only considers whether a hypothetical reasonable person would have taken the precautions the defendant failed to. It thus only specifies whether the defendant has violated the standard of care, i.e., whether the risk-taking was unjustified.

Condition (ii) refers to the defendant's capacities, and it is our focus. Focusing on what it takes for (ii) to be false of someone, Hart argues that the standard objection assumes an overly restrictive understanding of capacities. Hart asks us to consider someone who "does not know the consequences of his action, even though he failed to examine or to think about the situation before acting" (1961, 150). Would we say that such a defendant 'could not have helped it', thus viewing him as non-culpable? Hart points out that the standard objection assumes that we must indeed say that such a person 'could not have helped it'. But this assumption is false, Hart argues: we don't say 'he couldn't have helped it' of everyone who is unaware of the consequences of his actions, but only of someone who has, for example, a defective

memory, who cannot distinguish dangerous situations from harmless ones, etc. (1961, 150). Hart's idea is that the capacities that determine whether we could have helped it, in the sense that the criminal law cares about, include *the capacity to recognise risk*.

Understanding Hart's thought here requires noting a couple of things about the capacity to recognise risk. First, Hart's official statement of condition (ii)—in terms of capacities to 'take precautions'—must be read as centrally referring to *cognitive* capacities, given that his key examples are capacities to think before acting, to pay enough attention, and to recognise dangerous situations (1961, 148, 150–151).¹⁵ Second, and relatedly, the capacity to recognise risk is not one that we necessarily exercise or fail to exercise deliberately or voluntarily.¹⁶ In this second respect, Hart's account resembles the reasons-responsiveness account of responsibility for belief: our control over whether we recognise risk—like our control over our beliefs—is not *voluntary* control, i.e., it's not control that we can exercise at will or if we choose to.

The plausibility of Hart's response can be brought out by comparing Alexander and Ferzan's case of Sam and Ruth—who intuitively do seem culpable—to the case of *Elliott v. C (A Minor)* (1983).¹⁷ C, the defendant, was a 14-year-old girl of low intelligence, who set fire to a shed by igniting white spirit, and was convicted of criminal damage. This is despite the fact that the court found that she "had given no thought" to the risk she was causing and that the risk "would not have been obvious to her or appreciated by her if she had given thought to the matter" (*Elliott v. C*, at 945).

Where Sam and Ruth intuitively seemed culpable for their baby's death, C intuitively does not seem culpable for setting fire to the shed. The difference between the cases is plausibly precisely to do with the differing capacities to recognise risk. Given her low intelligence, C plausibly did lack the relevant capacity to recognise risk; she lacked the capacity to tell dangerous situations from harmless ones. And, for this reason, the conviction in *Elliott v. C* looks clearly unjust. It looks appropriate to say of C that she couldn't have helped it.

Alexander and Ferzan claim the same is true of Sam and Ruth, in their claim that "once the thought [of their child] was out of their minds, they had no power to retrieve it" (2009, 78). But once we have in mind Hart's broader sense of capacity, this is simply false. Sam and Ruth weren't drugged, hypnotised, or brainwashed into forgetting about their child. They did have the capacity or the power to remember

¹⁵ This is noted by Stark (2016, 184), who understands negligence as a failure of cognitive capacities, specifically as a defendant's failure to have a *belief* about a particular risk they are taking (2016, Chap. 8).

¹⁶ Hart does not say this in so many words, but it is implied. He suggests that a grossly negligent agent "has failed, *though not deliberately*, to take the most elementary precautions that the law requires him to take in order to avoid harm to others" (Hart 1961, 147, my italics). This implies that if an agent has the capacity to take these precautions in a case of negligence, they do not *deliberately* fail to exercise this capacity.

¹⁷ This case is also used to illustrate where Hart's account gets it right by Simester (2000, 104) and Stark (2016, 184).

about their child, and so—on Hart’s account—they *could* have prevented putting their child at risk.

Alexander and Ferzan give a response to Hart, but it’s one that doesn’t really engage with Hart’s fundamental point. They claim that an agent only ‘could have done otherwise’ in the sense relevant to the criminal law when the agent has an “internal reason” to do otherwise. Alexander and Ferzan don’t define an “internal reason” explicitly, but they seem to mean a reason to do otherwise that is accessible to the agent “at the time of the ‘negligent’ choice, [given] what he is conscious of and adverting to, his background beliefs, and so forth” (2009, 83). “To have such a reason,” Alexander and Ferzan claim, “an actor will have to advert to that to which he is not adverting. But one has no control at such moments over what one is adverting to or is conscious of: try thinking of what you are not thinking of, but should be!” (2009, 83).

This is not a good response to Hart because it doesn’t provide any reason to think that someone who is merely unaware that he is taking a risk genuinely lacks the capacity to recognise that risk. Someone may fail to exercise that capacity on a particular occasion, but this obviously does not mean he lacks the capacity itself. And we can agree it’s true that it is, in some sense, impossible to comply with an instruction to try thinking of what you are not thinking of. But, insofar as this is true, it’s merely the product of the logical fact that it’s impossible to exercise a capacity while failing to exercise it.¹⁸ It does not show that someone unaware of a risk thereby lacks the capacity to recognise that risk. People often do have this capacity, and Hart’s fundamental point is that we don’t say that agents ‘couldn’t have helped it’ but take a risk unless we have a specific reason for thinking they lack this capacity (a faulty memory, low intelligence, etc.). Alexander and Ferzan have not, therefore, given a convincing reason that a negligent defendant’s lack of awareness that he is taking a risk means that he lacks control over whether he takes that risk.

In giving Hart’s response to the standard objection to criminal negligence, I’ve so far focused on defendants who forget about a risk, like Alexander and Ferzan’s Sam and Ruth. One might worry that such defendants are not clear cases of negligent defendants, but are instead more plausibly thought of as reckless. Douglas Husak raises this kind of challenge, claiming that “many (but not all) cases of forgetting involve recklessness rather than negligence” (2011, 201). Husak claims that awareness involves belief, and claims that defendants who forget—like Sam and Ruth—may pass some key tests for having a tacit or dispositional belief about the relevant risk. Specifically, Husak suggests that forgetful defendants like Sam and Ruth will plausibly respond to a non-leading question, like ‘Where is your baby?’, with an immediate recognition that their baby is at risk, and if so they plausibly count as believing that their baby is at risk (2011, 212–213). If Husak is correct, that would undermine my discussion of the standard objection to criminal negligence, as we will have been focusing on a case that is not a genuine case of negligence.

While Husak raises an important question about how we should understand forgetting, I’m not sure he gives good reason to think that defendants who forget

¹⁸ I am grateful to Lucy Campbell for this point.

about risks count as believing they are taking a risk. This is because the key test on which Husak relies—if someone responds to a non-leading question about whether *p* with an immediate recognition that *p*, then they believe that *p*—doesn't distinguish between those who have a *dispositional belief* from those who merely have a *disposition to believe* (for this distinction, see Audi 1994). Having a *dispositional belief* that *p* involves having certain dispositions—exactly which dispositions differ, depending on the account; commonly cited ones include a disposition to act in ways that would satisfy one's desires if *p* were true, a disposition to assert that *p*, and a disposition to use *p* as a premise in one's reasoning. On the other hand, having a *disposition to believe* that *p* involves having a disposition to form those dispositions. To use Robert Audi's example, most of us plausibly only count as having a disposition to believe that 98.124 is larger than 98, rather than a dispositional belief with that content (1994, 419). With this distinction in mind, does the test on which Husak relies test for dispositional belief, or would it also be passed by someone who merely has a disposition to believe? It seems to me that Husak's test would be passed by both, and, therefore, can't settle the question of whether someone genuinely counts as believing something. And, given that the test doesn't settle this question, it seems more plausible to understand Sam and Ruth as merely having a disposition to believe—and not having a dispositional belief—that their baby is at risk, given that they seem to lack the relevant dispositions, i.e., given that they don't act in ways that would satisfy their desires if it were true that their baby is at risk, they're not disposed to assert that their baby is at risk, or to use this as a premise in their reasoning, etc. Therefore, Alexander and Ferzan's Sam and Ruth plausibly is a genuine case of negligence, and so we can draw more general conclusions about negligence from our discussion of it.

3.3 Improving on Hart's Account

My key aims in outlining Hart's account of negligence are to (a) highlight the parallel between it and the reasons-responsiveness account of responsibility for belief, and (b) show how the way in which it has been improved on points towards a more plausible account of responsibility for belief. I have done the first of these tasks: the key similarity is that they both appeal to specifically cognitive capacities as grounding the relevant kind of responsibility.

Carrying out the second of these tasks involves turning to a different objection to Hart, one developed by A.P. Simester. Simester argues that a failure to exercise the capacity to recognise risk cannot, by itself, ground culpability. Simester's criticism of Hart is an instance of what he calls "the basic challenge of culpability" (2013, 179). This challenge is raised by the fact that culpability requires more than just a bad act. We need, as Simester puts it, "to trace our negative evaluation of *q*ing back to *D*, the person who does it" (2013, 179), and justify the "evaluative link between act and defendant—that link which allows us to transmit judgments of the deed across to the person" (2013, 180).

Applying this to the case of unjustified risk-taking, we can note that risk-taking that is genuinely unjustified, such as leaving one's infant child unattended in a

running bath, is clearly something bad and to be avoided. But, in order to hold a defendant culpable for taking such a risk, we need more than just the fact that his risk was unjustified. We must show that the badness of his risk-taking speaks badly of the defendant.

Simester argues that Hart's capacity test fails to meet this challenge. The problem is *not* with those who fail to meet condition (ii), i.e., those who lack a capacity to recognise risk, such as the defendant in *Elliot v. C*. The fact that someone like C lacks an ability to recognise risk plausibly does explain why her risk-taking does *not* speak badly of her.

The problem is instead with those who have a capacity to recognise the relevant risk, but fail to exercise that capacity, such as Sam and Ruth, who do have the capacity to remember about their child. The fact that they *had* the capacity to recognise the risk does not explain *by itself* why *their* risk-taking speaks badly of *them*, because it only tells us that they are *not* exculpated (Simester 2000, 104; 2013, 184). Sam and Ruth do, of course, still meet condition (i), i.e., they fail to recognise a risk that a reasonable person would have. But Simester points out that, since condition (i) just tells us what a reasonable person would have done, it can only explain why their risk-taking is bad (*ibid.*); it cannot explain why their risk-taking speaks badly of *them*, which it must if it's going to explain why it's justified to hold *them* culpable.

Simester's objection to Hart highlights what seems to me to be the grain of truth in the standard objection to negligence. *If* the defendant were aware that he was taking a risk, as in a case of recklessness, then we would have an easy answer as to why his risk-taking speaks badly of him—he was aware of the risk he would cause to others, but he chose to go ahead regardless. But, because negligence excludes awareness, we don't have this easy answer. With a negligent defendant, we need an alternative explanation of why his risk-taking speaks badly of him.¹⁹

Again there is a parallel here to my earlier discussion of responsibility for belief. Specifically, Simester's key objection—that Hart's account fails to explain why a negligent defendant is at fault, or why his risk-taking speaks badly of them—is parallel to the problem I raised for the reasons-responsiveness account of responsibility for belief. Reasons-responsiveness, I argued, fails to explain how we are responsible for some unjustified beliefs—wishful thinking and confirmation bias—but not others—those resulting from mental illness, gaslighting, and propaganda. With the former cases, we want to say that the believer is *at fault*, and that his unjustified belief speaks badly of him epistemically. But reasons-responsiveness cannot *by itself* explain this; or, more accurately, it cannot explain it while also explaining why the believer in the latter cases is plausibly *not* at fault.

But as well as highlighting a shared problem, the parallel also points towards a more plausible account of responsibility for belief. In order to see this, we need to look at how Hart's account can be improved to overcome Simester's objection.

¹⁹ Some negligence critics seem to agree here. Moore and Hurd, for example, give the standard objection only as a preliminary to criticism of different attempts to justify culpability without awareness (2011, 165–191), and so presumably only view it as putting the burden of proof on defenders of negligence liability.

Specifically, the kind of account on which we should focus claims that criminal culpability for negligence requires not only that:

- (a) The negligent defendant has the capacity to recognise the risk he is taking, and fails to exercise that capacity,

but also that:

- (b) The negligent defendant's failure to exercise his capacity to recognise risk manifests *insufficient concern for others' interests*.

An account of more or less this kind has been given by Simester (2013, 192–194), but also by Victor Tadros (2002, 2005, Chap. 9) and Findlay Stark (2016, Chap. 8). These accounts have a number of subtle differences,²⁰ but these need not concern us here. The fundamental point they all make is that failures to recognise risks only ground criminal culpability if they manifest insufficient concern for others' interests.

This kind of account crucially relies on the claim that insufficient concern can be manifested by a failure to recognise risk, which may be initially difficult to understand. It is clear, as I admitted earlier, how someone's behaviour can manifest insufficient concern when he acts recklessly ('You knew you could have hurt her, but you went ahead regardless'). It may not be initially clear how it's possible for someone to manifest the same kind of insufficient concern when he is unaware that he is putting someone at risk.²¹ However, I think closer attention shows that failures to recognise risk can indeed manifest insufficient concern even though they are inadvertent and not chosen. This is because it's a quite general fact that how much we care about things can be manifested in the exercise of our cognitive capacities—or in the failure to exercise those capacities—as well as in what we do deliberately. How much or how little we care about things can influence what occurs to us, how much attention we pay to things, and how likely things are to slip our minds.

This point is most clearly illustrated in the context of personal relationships. Consider R.A. Duff's example of a man who's in the pub with his friends and forgets to turn up to his own wedding, thus manifesting "an utter lack of concern for his bride and their marriage" (1990, 163); Duff continues: "Had he cared at all for her, he *could* not have forgotten their wedding" (ibid.). Another example is one that Angela Smith finds in George Eliot's *Scenes of Clerical Life*, in which Miss Asher, when offered jelly by her fiancé, Captain Wybrow, expresses surprise that he has not yet learned that she does not like it, concluding "You don't seem to take much interest

²⁰ For example, Simester speaks of failures to recognise risk as manifesting "moral shortcomings" or "moral character flaws" (2013, 192, 194), rather than as manifesting "insufficient sufficient concern for others' interests" as Tadros and Stark do (Tadros 2005, 250–254; Stark 2016, 243–247). Also, Stark holds that negligence essentially involves holding people culpable for the beliefs they form (or fail to form) (2016, 228), whereas Simester instead characterises it as a "failure to act for the right reasons" (but one which stems from one's failure to recognise the right reasons) (2013, 191).

²¹ Thanks to an anonymous reviewer for pressing this point.

in my likes and dislikes” (quoted in Smith 2005, 242). For more examples, I invite readers to consider all of the times they have forgotten engagements with people because they didn’t really want to see them, or when they have failed to notice that they are offending people because they were too engrossed in being right. Cases like these—in which we fail to remember what we have to do or fail to notice things because we don’t care enough about them, or care more about other things—are a common fact of life in personal relationships. They indicate that our cognitive capacities, and in particular our failure to exercise them, can manifest our concern or lack thereof.

Now cases like these are not, of course, criminal cases, but they do plausibly suffice to show that concern can be manifested by our failures to recognise things, and not only by what we are aware of doing. And, once this is accepted, it becomes plausible that the kind of insufficient concern relevant to criminal culpability can also be manifested by failures to recognise risk. In the criminal case too, forgetting clearly looks like it can manifest the lack of concern relevant to culpability, at least in some cases. Indeed, Alexander and Ferzan’s case of Sam and Ruth looks exactly like such a case: they forget about the risk *because* they were narcissistically engrossed in their party, and didn’t care enough about their child.²²

Let’s return to how this improves on Hart’s account. The key requirement for culpability which is added is that the defendant’s failure to exercise his capacity to recognise risk must manifest insufficient concern for others’ interests. This feature provides the answer to the problem Hart’s account faced. That problem was that a mere failure to exercise a capacity to recognise risk is not the kind of thing that can, by itself, explain why someone’s unjustified risk-taking speaks badly of him. But *if* that failure to recognise risk manifests insufficient concern for others’ interests, that plausibly does look like it *can* explain why his risk-taking speaks badly of him.²³ This added requirement, therefore, provides a plausible fix to Hart’s account. And, returning to the case of epistemic responsibility, I’ll argue in the following sections that a parallel requirement provides a plausible fix to the reasons-responsiveness account of responsibility for belief.

4 Grounding Responsibility for Belief in Concern for the Truth

On the account I’ll defend, I’ll assume that reasons-responsiveness plays *some* role in grounding our responsibility for our beliefs. I think that it’s necessary for one to be responsible for a belief that it results from or can be influenced by a reasons-responsive process. Those whose beliefs fail to meet this minimal

²² Relatedly, this kind of account also means culpability for negligence may not be justified in a case like *Williams* (see footnote 10 above), because, unlike the case of Sam and Ruth, the *Williams*’s failure to recognise that their son was at risk plausibly did *not* manifest insufficient concern for his interests (Stark 2016, 267–269).

²³ Importantly, Alexander and Ferzan claim that manifesting insufficient concern for others’ interests is what grounds culpability—and justifies punishment—in the case of *advertent* wrongdoing (2009, Chap. 2).

condition—either generally or in specific cases—would plausibly not be responsible for those beliefs. But we saw above that reasons-responsiveness is not *sufficient* for responsibility. As with Hart’s account of negligence, we must supplement the reasons-responsiveness idea with a reference to concern.

The difference is that, whereas criminal culpability is a matter of whether one’s behaviour manifests insufficient concern for others’ interests, epistemic responsibility, I’ll suggest, is a matter of whether one’s beliefs manifest insufficient concern for the truth. More specifically, on the account I’ll defend, one is epistemically responsible for those of one’s beliefs that fail to respond to reasons *if* that failure manifests one’s insufficient concern for the truth.

It’s easiest to state this account not in terms of responsibility for belief simpliciter, but in terms of responsibility for a belief that has a certain epistemic *status*, such as being justified or unjustified, rational or irrational. For ease of illustration, I will focus on justification, but to the extent that other epistemic statuses, such as rationality or reasonableness, are distinct from justification, we will be able to make analogous claims about responsibility for beliefs that have these statuses.

On the account I propose, one is epistemically responsible for the epistemic status of one’s beliefs only insofar as they manifest one’s concern for the truth, or lack thereof. More specifically:

Justified S is epistemically responsible for a justified belief only if the fact that the belief is justified manifests S’s sufficient concern for the truth.

Unjustified S is epistemically responsible for an unjustified belief only if the fact that the belief is unjustified manifests S’s insufficient concern for the truth.

More informally, if your belief is unjustified because you don’t care enough about the truth of the matter, then you’re on the hook, you’re responsible, and it’s appropriate to hold you to account and criticise you for holding that belief. This feature of my account enables it to give a plausible account of *which* of our beliefs we are responsible for, and so represents an improvement on the reasons-responsiveness account. But, before I show that, I want to make a few clarifications about the notion of concern for the truth.

First, the relevant concern is not concern for the truth as such (as if ‘the truth’ were a friend one is worried about), but concern for *believing* the truth.

Second, there is the question of when one’s concern for the truth is *sufficient*. There is much to say about this question, but for the purposes of this paper it can be left undefined when someone counts as caring sufficiently about the truth, as it’s intuitively clear in particular cases whether someone cares enough about the truth. This is enough to illustrate the role played by concern for the truth in my account of epistemic responsibility. The complete account of when concern for the truth is sufficient will depend on what epistemic norms there are on belief and the strength of those norms, which is a question for normative epistemology. This is not to say that in order to be concerned for the truth believers themselves must be concerned about epistemic norms, just that the strength of what epistemic

norms there are on belief will determine what counts as sufficient concern for the truth.

Third, we should not think of concern for the truth as only operating consciously. Concern for the truth can be manifested in conscious reasoning, but also more implicitly, for example, showing up in one's patterns of belief formation and revision. The claim that our beliefs can manifest concern for the truth should not, therefore, be understood as the claim that belief formation is motivated by a conscious desire to believe the truth, a claim that looks false and likens belief formation too closely to intentional action. Relatedly, concern for the truth needn't be transparent: one can be *mistaken* about whether one is concerned with the truth. Consider a (non-clinical) conspiracy theorist. Such a person would typically describe himself as extremely concerned about the truth, but in a given case his patterns of reasoning and belief formation might make it clear that he is instead motivated by a desire to discover secrets and plots and to be the person 'who found it all out', or by a desire to uphold a very general world-view or ideology. In these respects, concern for the truth is like concern for a *person*. As we saw above, how much one cares about someone can be manifested by inadvertent behaviour, and not just by what one is aware of doing. And one can also be mistaken about one's concern for another person. In Jane Austen's *Emma*, for example, Emma only realises that she loves Mr. Knightley when she notices how annoyed she feels when Harriet, who she is trying to set up with someone else, tells her that she is interested in Mr. Knightley. In this example, which I take it illustrates a familiar kind of situation, Emma realises that she has been mistaken about what Mr. Knightley means to her.

5 Responsibility for Unjustified Beliefs Revisited

We can now return to the cases that posed a problem for reasons-responsiveness accounts. The problem was that we seem to be responsible for some reasons-non-responsive beliefs (those stemming from wishful thinking or confirmation bias) but not others (those stemming from mental illness, gaslighting, or propaganda). I think the account I've put forward can give a much more natural explanation of these cases.

All of these look like unjustified beliefs, in virtue of the fact that they are influenced by non-reasons-responsive processes. Therefore, on the account I've put forward, the believers in question will be responsible for them only if the fact that the beliefs are unjustified manifests insufficient concern for the truth. And this, I suggest, has the result that the believers are responsible in the first set of cases, and that they are not responsible in the second set of cases.

Let's consider the first set of cases: beliefs resulting from wishful thinking and confirmation bias. Intuitively, these are beliefs for which we are responsible, which on my account means that their lack of justification is a manifestation of insufficient concern for the truth on behalf of the believers. And this looks very plausible. Indeed, it's plausible that part of what it is to be engaged in either of these two forms of fallacious reasoning is that one's concern for the truth is defeated or overridden by one's other concerns. With wishful thinking, the overriding concern is what one

wants to be the case; with confirmation bias, the overriding concern is one's interest in one's initial judgement being correct. So my account gives the right result in the first set of cases.

Let's now consider the second set of cases: beliefs resulting from mental illness and abusive social situations. Although these beliefs are unjustified, the believers in question do not seem responsible for this. On my account, this is because their status as unjustified *does not* manifest insufficient concern for the truth on behalf of the believers.

Again this looks plausible. A schizophrenic whose propensity to believe conspiracy theories is a symptom of her illness doesn't manifest insufficient concern for the truth, but a kind of psychological breakdown that impairs her ability to think and reason clearly and assess the evidence as a healthy person can. Someone suffering from depression who believes that no one cares about her again has an unjustified belief, but not one that manifests insufficient concern for the truth on her part. This will be especially clear for a depressive who has insight, recognising her depressive beliefs as symptomatic of her illness, but unable to shed them nonetheless. In both cases, my account again gives the right result.

Now, think of the victim of gaslighting, who comes to doubt that her memory is reliable. In such a case—and in many cases of gaslighting—it's plausible that the victim's belief is unjustified.²⁴ But she doesn't have an unjustified belief because she cares insufficiently about believing the truth; she has an unjustified belief because her partner is placing coercive pressure on her to interpret the evidence in the way that he suggests. Indeed, this looks like a case where the abusive partner exploits the victim's concern for the truth, using it against her. Now, consider the case of someone who, having grown up under the influence of Stalinist propaganda, believes that criticism of the Moscow show trials are a capitalist plot. In this case too, it doesn't look like such a belief is unjustified because the believer cares insufficiently about believing the truth; instead the belief is unjustified because of the oppressive influence of the state on how such a person assesses the evidence. In both of these two cases, therefore, my account again gives the right result.

6 Conclusion

I started by outlining a puzzle about responsibility for belief. This puzzle arises because it looks like we hold each other responsible for what we believe, and that we lack control over what we believe, but it also looks like responsibility requires control. I argued that the most plausible way of answering this puzzle—which claimed that we have a distinctive kind of non-voluntary control over our beliefs—failed. It failed because it could not explain which of our unjustified beliefs we are

²⁴ This may not always be true. It's conceivable that the controlling partner arranges the evidence so that the victim is perfectly justified in doubting her memory. I'm not considering cases like that because they wouldn't be non-reasons-responsive beliefs. I'm considering cases in which the manipulation results in the victim having unjustified beliefs, which will include many, if not most, real-life cases of gaslighting.

responsible for. But the parallel failure of Hart's account of culpability for negligence pointed towards a more promising account of responsibility of belief, specifically, by drawing on accounts of culpability for negligence that improve on Hart's. An account of responsibility for belief based on this model—one which is grounded in the believer's concern for the truth—gives a plausible account of which beliefs we're responsible for and which we're not. In this way, we can further our understanding of epistemic responsibility by drawing on the philosophy of criminal law.

Acknowledgements For helpful comments on previous drafts, I'd like to thank Lucy Campbell, Findlay Stark, Hallvard Lillehammer, Sophie Archer, Daniel Whiting, Conor McHugh, Antony Duff, Sandra Marshall, Christopher Cowley, and an anonymous reviewer from *Criminal Law and Philosophy*, as well as to Jill Flohil for helpful and diligent editing. I'm also grateful to audiences at Reading, Southampton, Birkbeck, Edinburgh, and Warwick. This work was funded by a Leverhulme Trust Early Career Fellowship.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Alexander, Larry, and Kimberly Kessler Ferzan. 2009. *Crime and Culpability: A Theory of Criminal Law*. Cambridge: Cambridge University Press.
- Alexander, Larry, and Kimberly Kessler Ferzan. 2014. "Confused Culpability, Contrived Causation, and the Collapse of Tort Theory." In *Philosophical Foundations of the Law of Torts*, edited by John Oberdiek, 406–426. Oxford: Oxford University Press.
- Alston, William P. 1985. "Concepts of Epistemic Justification." *The Monist* 68 (1): 57–89.
- Alston, William P. 1988. "The Deontological Conception of Epistemic Justification." *Philosophical Perspectives* 2: 257–299.
- American Law Institute. 1962. *Model Penal Code: Official Draft and Explanatory Notes: Complete Text of Model Penal Code*. 1985. Philadelphia: American Law Institute.
- Audi, Robert. 1994. "Dispositional Beliefs and Dispositions to Believe." *Noûs* 28 (4): 419–434.
- Beever, Allan. 2007. *Rediscovering the Law of Negligence*. Oxford: Hart.
- Chrisman, Matthew. 2008. "Ought to Believe." *The Journal of Philosophy* 105 (7): 346–370.
- Chuard, Philippe, and Nicholas Southwood. 2009. "Epistemic Norms without Voluntary Control." *Noûs* 43 (4): 599–632.
- Duff, R.A. 1990. *Intention, Agency and Criminal Liability*. Oxford: Basil Blackwell.
- Elliott v. C (A Minor)* [1983] 1 WLR 93.
- Feldman, Richard. 2000. "The Ethics of Belief." *Philosophy and Phenomenological Research* 60 (3): 667–695.
- Feldman, Richard. 2008. "Modest Deontology in Epistemology." *Synthese* 161 (3): 339–355.
- Fischer, John Martin, and Mark Ravizza. 1998. *Responsibility and Control: A Theory of Moral Responsibility*. Cambridge Studies in Philosophy and Law. Cambridge: Cambridge University Press.
- Ginet, Carl. 2001. "Deciding to Believe." In *Knowledge, Truth, and Duty: Essays on Epistemic Justification, Responsibility, and Virtue*, edited by Matthias Steup, 63–76. Oxford: Oxford University Press.
- Hall, Jerome. 1963. "Negligent Behavior Should Be Excluded from Penal Liability." *Columbia Law Review* 63 (4): 632–644.
- Hart, H.L.A. 1961. "Negligence, *Mens Rea*, and Criminal Responsibility." In *Punishment and Responsibility: Essays in the Philosophy of Law*, 2nd ed., 136–157. 2008. Oxford: Oxford University Press.
- Hart, H.L.A. 1968. *Punishment and Responsibility: Essays in the Philosophy of Law*. 2nd ed. 2008. Oxford: Oxford University Press.
- Heil, John. 1992. "Believing Reasonably." *Noûs* 26 (1): 47–61.
- Hieronymi, Pamela. 2006. "Controlling Attitudes." *Pacific Philosophical Quarterly* 87 (1): 45–74.

- Hieronymi, Pamela. 2008. "Responsibility for Believing." *Synthese* 161 (3): 357–373.
- Hurd, Heidi M. 2014. "Finding No Fault with Negligence." In *Philosophical Foundations of the Law of Torts*, edited by John Oberdiek, 387–405. Oxford: Oxford University Press.
- Husak, Douglas. 2011. "Negligence, Belief, Blame and Criminal Liability: The Special Case of Forgetting." *Criminal Law and Philosophy* 5 (2): 199–218.
- Levy, Neil. 2007. "Doxastic Responsibility." *Synthese* 155 (1): 127–155.
- McCormick, Miriam. 2011. "Taking Control of Belief." *Philosophical Explorations* 14 (2): 169–183.
- McHugh, Conor. 2012. "Epistemic Deontology and Voluntariness." *Erkenntnis* 77 (1): 65–94.
- McHugh, Conor. 2013. "Epistemic Responsibility and Doxastic Agency." *Philosophical Issues* 23 (1): 132–157.
- McHugh, Conor. 2014. "Exercising Doxastic Freedom." *Philosophy and Phenomenological Research* 88 (1): 1–37.
- Moore, Michael S., and Heidi M. Hurd. 2011. "Punishing the Awkward, the Stupid, the Weak, and the Selfish: The Culpability of Negligence." *Criminal Law and Philosophy* 5 (2): 147–198.
- Nickel, Philip J. 2010. "Voluntary Belief on a Reasonable Basis." *Philosophy and Phenomenological Research* 81 (2): 312–334.
- Nottelmann, Nikolaj. 2007. *Blameworthy Belief: A Study in Epistemic Deontology*. Dordrecht: Springer Science.
- Owens, D.J. 2000. *Reason Without Freedom: The Problem of Epistemic Normativity*. Abingdon: Routledge.
- Peels, Rik. 2017. *Responsible Belief: A Theory in Ethics and Epistemology*. Oxford: Oxford University Press.
- Pettit, Philip, and Michael Smith. 1996. "Freedom in Belief and Desire." *The Journal of Philosophy* 93 (9): 429–449.
- Posner, Richard A. 1972. "A Theory of Negligence." *The Journal of Legal Studies* 1 (1): 29–96.
- R v. Adomako [1995] 1 AC 171.
- R v. Bateman [1927] 19 Cr App R 8.
- Road Traffic Act. 1988 (U.K.) London: Her Majesty's Stationary Office. Available at: <https://www.legislation.gov.uk/ukpga/1988/52/contents> (accessed 9th September 2019).
- Ryan, Sharon. 2003. "Doxastic Compatibilism and the Ethics of Belief." *Philosophical Studies* 114 (1): 47–79.
- Scanlon, T.M. 1998. *What We Owe to Each Other*. Cambridge, MA: Belknap Press.
- Simester, A.P. 2000. "Can Negligence Be Culpable?" In *Oxford Essays in Jurisprudence, Fourth Series*, edited by Jeremy Horder, 85–106. Oxford: Oxford University Press.
- Simester, A.P. 2013. "A Disintegrated Theory of Culpability." In *The Sanctity of Life: The Legacy of Glanville Williams*, edited by Dennis J. Baker and Jeremy Horder, 178–203. Cambridge: Cambridge University Press.
- Simester, A.P., J.R. Spencer, Findlay Stark, G.R. Sullivan, and G.J. Virgo. 2016. *Simester and Sullivan's Criminal Law: Theory and Doctrine*. 6th edition. Oxford: Hart Publishing.
- Smith, A.M. 2005. "Responsibility for Attitudes: Activity and Passivity in Mental Life." *Ethics* 115 (2): 236–271.
- Stark, Findlay. 2016. *Culpable Carelessness: Recklessness and Negligence in the Criminal Law*. Cambridge: Cambridge University Press.
- State v. Williams 484 P 2d 1167 (1971, WA).
- Steup, Matthias. 2000. "Doxastic Voluntarism and Epistemic Deontology." *Acta Analytica* 15 (24): 25–56.
- Steup, Matthias. 2012. "Belief Control and Intentionality." *Synthese* 188 (2): 145–163.
- Tadros, Victor. 2002. "Recklessness and the Duty to Take Care." In *Criminal Law Theory: Doctrines of the General Part*, edited by Stephen Shute and A.P. Simester, 227–258. Oxford: Oxford University Press.
- Tadros, Victor. 2005. *Criminal Responsibility*. Oxford: Oxford University Press.
- Turner, J.W.C. 1936. "The Mental Element in Crimes at Common Law." *The Cambridge Law Journal* 6 (1): 31–66.
- Weinrib, Ernest J. 1995. *The Idea of Private Law*. Oxford: Oxford University Press.
- Williams, Glanville. 1961. *Criminal Law: The General Part*. 2nd ed. London: Stevens.