Here is a puzzle about *exculpation.* Imagine Anne, who tortures innocent people because she believes that she ought to, and who has met all of her epistemic obligations when forming that belief.[1] Now consider:

(a) Many of us distinguish between objective wrongdoing and culpable wrongdoing; that is, we think that one can perform an impermissible action and yet not be *culpable* for doing so. Someone who isn't culpable for objective wrongdoing is *exculpated.*

(b) Many of us think that if one's objectively impermissible action is traceable to a "responsibly formed" belief, then one is exculpated for one's objective wrongdoing; so, Anne is exculpated.

(c) And yet many of us think that there are certain *especially horrible* impermissible actions for which one cannot be exculpated, even when those actions can be traced back to "responsibly formed" beliefs. So, Anne remains morally responsible.

Given natural interpretations of these three commitments, they cannot all be true. The puzzle, then, is which of these commitments we ought to reject.

With respect to this puzzle, current work on exculpation—and specifically on the role played by an agent's beliefs in exculpation—is split into two camps. One camp says that only false *non-moral* beliefs can exculpate, and that false *moral* beliefs do not exculpate; this view is usually motivated by a *quality of the will* account of moral responsibility.[2] This view, as it's often developed, entails the counterintuitive conclusion that almost anyone who acts on a false moral belief (no matter the moral content or history of that belief) is culpable for doing so.[3] The second camp says that any belief can, in principle, exculpate; this camp is motivated by the idea that an agent is only culpable for objective wrongdoing when the agent's action is traceable to an instance of akrasia.[4] This view, as it's usually developed, leads to the counterintuitive conclusion that most people are not culpable for most of what they do.[5] The two camps disagree about whether one can be exculpated for objective moral wrongdoing when (i) the wrongdoing is based on

---

[1] If the reader thinks this scenario is impossible, then they can supply any example in which someone does something horrible on the basis of a responsibly formed—but false— moral belief.

[2] Alvarez and Littlejohn 2017; Arpaly 2002a; Arpaly 2002b; Harman 2011; Harman 2017; Mason and Wilson 2017; A.M. Smith 2005; Talbert 2017. This camp rejects (b).

[3] See, e.g., Harman 2011.

[4] Levy 2009; Rosen 2004; Zimmerman 1997. This camp rejects (c).

[5] Rosen 2004; Zimmerman 1997. Some philosophers have tried to develop this view

1

a false moral belief and (ii) one has met all of one's epistemic obligations in forming that belief.[6]

In my paper, I develop a new way of conceptualizing exculpation that solves the puzzle mentioned earlier while avoiding the counterintuitive conclusions entailed by other current views of exculpation. Moreover, my view isn't ad hoc; it's motivated by plausible claims about the nature of subjective obligation (and its relationship to exculpation), the nature of evidence, and the nature of moral theories. My view of exculpation also rests on an *attributability* interpretation of moral responsibility.[7] However, even those who are more interested in answerability or accountability will still be interested in the view I develop (insofar as attributability is a necessary condition for answerability and accountability[8]).

My view comprises four core claims:

1. One is exculpated for objective moral wrongdoing when one's objectively impermissible action satisfies a *subjective moral obligation/permission*.

2. In order to have a subjective moral obligation/permission to ϕ, one must acquire adequate *moral evidence* for the proposition that one is morally obligated/permitted to ϕ.

3. In order to acquire adequate moral evidence for a moral proposition, one must employ a *moral background theory*.

4. There are *constraints* on the moral propositions that a moral background theory can include.

Together, these four core claims entail that there can be false moral beliefs that yield exculpating subjective moral obligations/permissions; this is

---

further so that it no longer entails this counterintuitive conclusion (Clark 2017; FitzPatrick 2008; FitzPatrick 2017; Montmarquet 1999; Peels 2011; Sher 2017). However, I find these extensions of the view ad hoc, and believe that my view—which can be positively motivated—manages to accommodate the intuitions that participants in this debate think need to be accommodated.

[6]It might also turn out that the source of the two camps' disagreement is that they focus on different notions of culpability, namely, attributability, answerability, and accountability. However, participants in the debate disagree about whether they're using different notions of moral responsibility (see Smith 2012). In this paper I'll focus on (and defend my focus on) attributability.

[7]See Scanlon 2008; Shoemaker 2011.

[8]This is a contentious claim; nevertheless, I find it plausible, and provide further motivation for it in my paper.

possible under the assumption that a moral background theory can contain *some* false moral propositions. However, they also entail that there are false moral beliefs that cannot exculpate, because it is impossible to acquire moral evidence for those beliefs using a moral background theory. Thus, my view entails that the solution to the puzzle mentioned earlier is to reject (b), but not in exactly the same way (or for the same reasons) that others reject (b).

My view is similar to quality of the will accounts of moral responsibility, because it exculpates an agent for objective moral wrongdoing when the agent "handles" moral evidence well and thereby exhibits a good will. Moreover, handling moral evidence well requires more than bare rationality; the agent must actually engage with evidence in a way that's recognizably *moral*.

However, my view diverges from standard quality of the will accounts, because it allows for some variation across moral background theories. That is, it allows for the possibility that one can engage with evidence in a recognizably moral way even when one doesn't hold only true moral beliefs. Moreover, my view does not require perfect or near-perfect "de re" moral motivation;[9] instead, it evaluates an agent's quality of will by looking at whether they engage with the world around them using a recognizably moral framework (even if that framework doesn't align perfectly with whichever moral theory is objectively true).

I begin my paper by defending my focus on attributability, and by arguing for my four core claims, with special attention paid to the fourth. From there, I use my four core claims to develop two arguments. The first argument has a weak conclusion: that for any agent, there is at least one possible false moral belief that the agent cannot be exculpated for acting on. The second argument has a stronger conclusion: that there are false moral beliefs for which no one can be exculpated for acting on. (This stronger conclusion requires the introduction of a further assumption, which I also take to be plausible.) I end my paper by showing that when we place reasonable constraints on moral theories, my view avoids the counterintuitive results faced by most views of exculpation; given those reasonable constraints, my view entails that some, but not all, false moral beliefs can exculpate. However, the main upshot of my paper is that *how permissive* we are when it comes to exculpation ultimately depends on *how many constraints* we place on what counts as a moral theory.

---

[9]Cf. Arpaly 2002a; Arpaly 2002b; Harman 2011.

References

1. Alvarez, M. and Clayton Littlejohn, "When Ignorance is No Excuse," in *Responsibility: The Epistemic Condition*, ed. Robichaud and Wieland (Oxford: Oxford University Press, 2017), pp. 64-81.

2. Arpaly, N. *Unprincipled Virtue: An Inquiry into Moral Agency* (Oxford University Press, 2002a).

3. Arpaly N. "Moral Worth," *Journal of Philosophy*, Vol. 99, No. 5 (2002b), pp. 223-245.

4. Clark, R. "Ignorance, Revision, and Common Sense," in *Responsibility: The Epistemic Condition*, ed. Robichaud and Wieland (Oxford: Oxford University Press, 2017), pp. 233-251.

5. FitzPatrick, W. "Moral Responsibility and Normative Ignorance: Answering a New Skeptical Challenge," *Ethics*, Vol. 115, No. 4 (2008), pp. 651-691.

6. FitzPatrick, W. "Unwitting Wrongdoing, Reasonable Expectations, and Blameworthiness," in *Responsibility: The Epistemic Condition*, ed. Robichaud and Wieland (Oxford: Oxford University Press, 2017), pp. 29-46.

7. Harman, E. "Does Moral Ignorance Exculpate?" *Ratio*, Vol. 24, No. 4 (2011), pp. 443-468.

8. Harman, E. "When is Failure to Realize Something Exculpatory?" in *Responsibility: The Epistemic Condition*, ed. Robichaud and Wieland (Oxford: Oxford University Press, 2017), pp. 117-126.

9. Levy, N. "Culpable Ignorance and Moral Responsibility: A Reply to FitzPatrick," *Ethics*, Vol. 119, No. 4 (2009), pp. 729-741.

10. Mason, E. and A.T. Wilson, "Vice, Blameworthiness, and Cultural Ignorance," in *Responsibility: The Epistemic Condition*, ed. Robichaud and Wieland (Oxford: Oxford University Press, 2017), pp. 82-100.

11. Montmarquet, "Zimmerman on Culpable Ignorance," *Ethics*, Vol. 109, No. 4 (1999), pp. 842-845.

12. Peels, R. "Tracing Culpable Ignorance," *Logos and Episteme*, Vol. 2, No. 4 (2011), pp. 575-582.

13. Rosen, G. "Skepticism About Moral Responsibility," *Philosophical Perspectives*, Vol. 18, No. 1 (2004), pp. 295-313.

14. Scanlon, T. *Moral Dimensions: Permissibility, Meaning and Blame* (Cambridge, MA: Harvard University Press, 2008).

15. Sher, G. "Blame and Ignorance," in *Responsibility: The Epistemic Condition*, ed. Robichaud and Wieland (Oxford: Oxford University Press, 2017), pp. 101-116.

16. Shoemaker, D. "Attributability, Answerability, and Accountability: Toward a Wider Theory of Moral Responsibility," *Ethics*, Vol. 121 (2011), pp. 602-632.

17. Smith, A. "Responsibility for Attitudes: Activity and Passivity in Mental Life," *Ethics* Vol. 115, No. 2 (2005), pp. 236-271.

18. Smith, A. "Attributability, Answerability, and Accountability: In Defense of a Unified Account," *Ethics*, Vol. 122, No. 3 (2012), pp. 575-589.

19. Talbert, M. "Akrasia, Awareness, and Blameworthiness," in *Responsibility: The Epistemic Condition*, ed. Robichaud and Wieland (Oxford: Oxford University Press, 2017), pp. 47-63.

20. Zimmerman, M. "Moral Responsibility and Ignorance," *Ethics*, Vol. 107, No. 3 (1997), pp. 410-426.