

DISCUSSION NOTE

## Responsibility Skeptics Should Be More Skeptical

Aarthy Vaidyanathan

Department of Philosophy, Princeton University, Princeton, New Jersey, USA

Email: aarthyv@princeton.edu

### Abstract

Menges (2022) seeks to identify the kind of blame that should be at issue in debates between skeptics and anti-skeptics about responsibility. Menges argues that such blame is constituted by responses that the target has a claim against, and by the blamer's thought that they have forfeited this claim due to their bad action and state while engaged in that action. I identify a class of blame responses that Menges mistakenly excludes and offer an alternative, more general, account in which the distinctive feature of controversial blame isn't claim forfeiture, but the defeat of reasons grounded in the target's interests.

**Keywords:** moral responsibility; claim forfeiture; defeat of interests; blame; blameworthiness; skepticism about moral responsibility

Skeptics about responsibility typically argue that metaphysical possibilities like determinism or luck can undermine our practice of holding people responsible.<sup>1</sup> Anti-skeptics reject this claim.<sup>2</sup> In his article, "The Kind of Blame Skeptics Should Be Skeptical About," Leonhard Menges (2022) identifies an important problem for skeptics and anti-skeptics—the ambiguity regarding the forms of backward-looking blame or praise that they should disagree about. I will call these forms of blame and praise the 'controversial' forms of blame and praise. Blame constituted wholly by a belief that someone acted badly is uncontroversial since such beliefs are plausibly warranted even in a deterministic universe, while blame constituted by retributive punishment is paradigmatically controversial. However, the status of many familiar everyday forms of blame like rebukes or resentment is murkier.

Menges argues that a blame response is controversial if the target has an antecedent claim against it that the blamer takes them to have forfeited due to their action and state while engaged in that action. On this view, rebukes are controversial, while resentment isn't. I will argue that while Menges's account captures many controversial forms of blame, there is a large, important, and commonplace class of cases of controversial blame that Menges should capture but cannot. These cases involve responses that can be justified without claim forfeiture.

I will sketch an alternative account in which controversial forms of blame are constituted by responses that can be bad for the target and that rely on reasons provided by the target's bad actions and state while engaged in it to defeat reasons against the response grounded in the target's interests. This account is more general than Menges's; it captures the cases that Menges's account does, as well as cases that it doesn't.

<sup>1</sup>Examples include Pereboom (2014), Strawson (1994), Rosen (2002), and Levy (2011). Caruso (2018) provides an overview.

<sup>2</sup>Examples include Strawson (1962), Wallace (1994), and Mckenna (2008).

## 1

Menges articulates the Claim Forfeiture View on Blame (CFB) to “characterize the responses that should be at issue between those philosophers who argue that, in our world, humans are never appropriate targets of blame and those who deny this” (2022, 1). According to the CFB: “Skeptics about blameworthiness should argue that, in our world, humans are never worthy of responses that (i) would violate a claim of theirs if it is not waived, overridden, or forfeited and (ii) involve the thought that they forfeited the claim against this response because they did something morally objectionable and because of how they were when they did it” (10).

Menges argues that the Claim Forfeiture View captures controversial forms of blame because it meets three constraints: the neutrality constraint, the revisionary constraint, and the justice constraint. The view meets these three constraints because it is not trivially true that someone’s bad action can make them forfeit a moral claim regardless of the metaphysical possibilities that concern the skeptic, it would be revisionary to eschew all blame responses that can violate claims, and it would be clearly unjust to blame the blameless in a way that violates their claims.

I will argue that although meeting the conditions of the Claim Forfeiture View may well be sufficient for a blame response to be controversial, it is not necessary. There is a class of cases of controversial blame that the view should capture but cannot. Consider two such cases:

**The Racist Joke:** You hear an acquaintance make a racist joke at a party. A few days later, he asks to borrow your kayak. You refuse, although you are generally inclined to say yes despite the inconvenience, because he made that joke.

**The Profligate Neighbor:** A neighbor gets behind on her mortgage payments after buying a luxury car. She requests help; concurrently, another neighbor also requests help due to a sudden rise in fuel prices. You can’t help them both and decide to help the neighbor affected by the fuel prices because the other neighbor shouldn’t have bought such an expensive car.

I wish to make two claims about your responses in these cases. First, they amount to blame. Second, they amount to controversial blame.

Your responses in these cases amount to blame because in each case you withhold a benefit that you would otherwise have good reason to grant, and you do so on the ground that they have acted badly.<sup>3</sup> This is a way of holding their action against them—a way in which the person from whom the benefit is withheld is made worse off on account of their bad action. If you withheld benefits from them for reasons that do not include their misconduct, say because you didn’t feel like it, your response would not amount to blame. And if you offered to buy them a drink because they acted badly, again, your response would not amount to blame.

Note that in these cases the target has no claim to the benefit you are withholding. So, if withholding a benefit is to amount to controversial blame, it cannot be because it is a response that would violate a claim against it unless the claim is forfeited. Your acquaintance does not have a claim to borrow your kayak and your neighbor does not have a claim to your financial assistance.

Your responses in these cases do amount to controversial blame because they meet the three constraints that Menges has identified: the neutrality constraint, the revisionary constraint, and the justice constraint. First, it is not trivially true that someone’s bad action can ground rational support for withholding discretionary benefits regardless of metaphysical possibilities like determinism. If the profligate neighbor or the acquaintance who made the racist joke suffer from a brain tumor that causes their actions, it seems less plausible that their actions give you reason to refuse their requests. Skeptics would argue that determinism is like the brain

---

<sup>3</sup>I would like to thank Reviewer #1 for suggesting that I elaborate on how your responses amount to blame. I also thank Gideon Rosen for suggesting this way of putting the point.

tumor.<sup>4</sup> Anti-skeptics would reject this strategy. So, your blame responses meet the neutrality constraint.

Second, I suspect that we often withhold morally discretionary benefits in the manner illustrated in these cases (as a way of holding the agent's bad actions against them). Interactions involving discretionary benefits are commonplace. So, giving up such blame responses would be revisionary.

Third, although the target has no claim to the benefit as such in these cases, they do have a claim against the benefit being withheld on the basis of reasons that are illegitimate. Your blame responses are impermissible and unjust in the bad versions of the cases (in which the target is in fact blameless) because they violate this claim. Note, however, that in the good versions of the cases—where your responses are permissible—they aren't permissible in virtue of this claim being forfeited. The claim remains fully in place. Rather, your responses are permissible because the benefit is not withheld for an illegitimate reason and so the claim isn't violated. Thus, your blame responses in these cases meet the justice constraint. However, they do so in a manner that is independent of claim forfeiture. Menges's Claim Forfeiture View of controversial blame cannot, therefore, capture your responses.

Note that although I focus on cases that involve withholding benefits, actions that cause permissible harm can be made similarly intelligible as forms of controversial blame: consider a variation of the racist joke case in which you decide to build a wall on your property that would block your acquaintance's view of the lake.<sup>5</sup> The Claim Forfeiture View is inadequate because it cannot capture such cases of controversial blame constituted by withholding discretionary benefits or causing permissible harm.

## 2

I will now consider objections to my argument. First, we might worry that the blame responses I'm pointing to are not really controversial and so are not a problem for the Claim Forfeiture View because the justification provided by the target's bad action is in fact independent of determinism and the like in virtue of being grounded in the norms of relationships or the requirements of respect. Menges suggests, for example, that refraining from inviting a friend to your birthday party because of her promise-breaking constitutes a form of blame that is appropriate regardless of luck or determinism (2022, 4);<sup>6</sup> perhaps because the requirements of self-respect or the norms of friendship make it appropriate to refrain in virtue of her promise breaking. You are no doormat, and you value healthy friendships. Such reasons for refraining from inviting the friend (or refusing to lend the kayak) seem to be independent of determinism and of luck.

<sup>4</sup>Mele (1995, 168) notes that the oft-used manipulation cases need not rely on the element of *intentional* manipulation to generate skeptical intuitions. A variant of the case involving a suitable brain disorder would generate similar intuitions.

<sup>5</sup>I thank Gideon Rosen for this example.

<sup>6</sup>Menges uses the case to argue that skeptics should not be skeptical about the kind of blame that T. M. Scanlon is talking about in *Moral Dimensions*, blame that is constituted by relationship revision made appropriate by actions that indicate something about the target that impairs that relationship. Menges claims that revising your relationship with the promise-breaking friend by "giving up the intention to invite them to your birthday party" can be appropriate regardless of luck or determinism (4). I argue, contra Menges, that skeptics should be skeptical about at least some types of blame constituted by relationship revision, exemplified by the case of the promise-breaking friend. See also Pereboom (2014, 131). Pereboom argues that the acceptability of Scanlonian blame to responsibility skeptics depends on how we understand the notion of appropriateness in Scanlon's account. He notes that the notion has a basic desert aspect as well as an epistemic aspect, and that while the latter is unproblematic for the skeptic, the former is not. The blame responses in the cases I've presented as a challenge to Menges's account are similar to cases of Scanlonian blame in that "they deprive a person of things that he or she has reason to want. But they are not things to which anyone has an unconditional claim" (Scanlon 2008, 187). However, in my cases the target has *no* claim (not even a conditional claim) to the benefit that is withheld from them, and Scanlonian blame is not limited to such cases. The blame responses in my cases are also different from Scanlonian blame in that they need not constitute the modification of a relationship. I would like to thank Reviewer #2 for pressing me to clarify the relation between my cases and Scanlonian blame.

In response to this objection, I will first note that a justification grounded in relationship norms or the requirements of respect is unavailable in the profligate neighbor case. It is implausible that the improper management of our finances wrongs our neighbors as a general matter. Next, I will consider my case of the racist joke. This case seems analogous to the case of the promise-breaking friend. It is arguable that requirements of respect or the norms of social/moral relationships make it appropriate to withhold certain discretionary benefits from the person who made the racist joke.<sup>7</sup> The objection suggests that justification of this kind is not threatened by facts about determinism or luck, and that the responses in the case of the racist joke or that of the promise-breaking friend do not, therefore, constitute a kind of blame that moral responsibility skeptics need find problematic. I disagree.

Cases like that of the promise-breaking friend or the racist joke *are* cases of controversial blame.<sup>8</sup> Suppose that you find out that the promise-breaking friend has a brain tumor that causes her promise-breaking. Inviting them to your party no longer seems to constitute behaving like a doormat. You might, of course, have reason to refrain from inviting the friend simply because of the risk of their future promise breaking and its impact on you. Such self-protection, however, would not constitute blame. Now, as I noted earlier, skeptics and anti-skeptics disagree about the use of the brain tumor thought experiment, and so such cases are cases of controversial blame.

A second worry is regarding whether the forms of blame I point to can really raise a question of justice. The targets have no claim to your kayak or to your financial assistance, and that seems to entail that withholding the benefit can't be unjust although it can be unkind or petty. It would sound rather odd were the target to say, "I have no claim to your kayak, but your refusal to let me borrow it is unjust."<sup>9</sup> This is an important worry because, as Menges notes, prominent responsibility skeptics are concerned about blame responses that are, all else equal, *unjust* when directed towards the blameless (2022, 4). Their concern isn't about blame responses being petty or unkind.

In response to this worry, consider this: it would not sound odd were someone to say, "I have no claim to your kayak, but your refusal to let me borrow it because I'm black is unjust." It is unjust to withhold even discretionary benefits from someone if your reasons for doing so include their membership in certain racial groups. Analogously, it is unjust to withhold benefits from someone who is blameless if your reasons for doing so include their bad action.

Your reasons for withholding benefits shape the meaning of this act.<sup>10</sup> When performed for certain reasons, the act of withholding benefits can express that the person from whom the benefits are withheld is not worthy or less worthy of these benefits. This can make the action morally impermissible and unjust if what is expressed is false.<sup>11</sup> The target may not have a claim to the benefit, but they do have a claim against being denied the benefit in ways that misrepresent their worthiness. If you refuse to let someone borrow your kayak because you dislike them, your action does not express that they are not worthy or less worthy of this benefit. However, if you refuse because of their race or because of their bad action, then your refusal does express that they are not worthy or less worthy of this benefit.

<sup>7</sup>Note, however, that although such a ground is available, it need not be the ground of my withholding benefits from the person who made the racist joke. I might take the racist joke to be a moral mistake that warrants such a response regardless of its implications for the perpetrator's relationships.

<sup>8</sup>I do not focus on this promise-breaking friend case in my main argument against the Claim Forfeiture View because I find it plausible that the friend has a claim to being invited to the party that she might forfeit due to her promise-breaking.

<sup>9</sup>I would like to thank Reviewer #2 for articulating and pressing the objection in this form.

<sup>10</sup>I'm understanding the meaning of an action here in the way that T. M. Scanlon does in *Moral Dimensions* (2008, 52–56).

<sup>11</sup>See also Scanlon's discussion in *Moral Dimensions* of how the meaning of a discriminatory action affects its permissibility (2008, 69–74).

## 3

The cases I've presented suggest that in blaming someone, we take their bad actions to contribute to defeating reasons grounded in their interests that favor benefitting them (or that count against burdening them). I will now sketch an account of controversial blame based on this observation.<sup>12</sup>

**The Defeat of Interests View:** Controversial forms of blame can be constituted by responses that (i) are bad for the target in some way and that (ii) rely on the reasons provided by their actions falling short of relevant standards and their state while performing these actions (iii) to defeat the reasons against such a response that are grounded in the target's interests.

Note that the target's bad action need not in fact provide reasons to withhold benefits in order for blame to be controversial.<sup>13</sup> However, the subject of the blame response must *rely* on reasons provided by the target's action in the sense that the response (withholding benefits) is rationally supported from the subject's point of view by these reasons. These reasons are special: they presuppose the existence of reasons against the response that are grounded in the target's interests. They can *defeat* the reasons against the response in the sense that the rational force of reasons against the response that are grounded in the target's interests can be extinguished or exceeded by reasons in favor of the response grounded in the target's action.

The Defeat of Interests View meets Menges's neutrality constraint, revisionary constraint, and justice constraint. It is not trivially true that acting badly can ground a reason that can defeat the reasons provided by the agent's interests regardless of metaphysical possibilities like determinism; it would be revisionary to eschew blame responses that rely on such reasons, and it would be unjust to the blameless to take their interests to be defeated by the reasons grounded in their actions.

The Defeat of Interests View is more general than Menges's Claim Forfeiture View. It can accommodate every form of controversial blame captured by the latter. Consider, for instance, Menges's case in which you take yourself to be justified in aggressively reproaching a friend for her promise-breaking (2022, 11). It is a case of controversial blame on the Claim Forfeiture View because you take the friend to have forfeited her claim against such reproachment. It is also a case of controversial blame on the Defeat of Interests View. Your friend's claim against such treatment constitutes a reason against reproaching her that is grounded in her interests. To defeat this reason, you rely on a reason you think is provided by her promise-breaking—your friend's forfeiting her claim.

## 4

Skeptics about responsibility should be skeptical about cases in which blaming someone does not involve taking any of their moral claims to have been forfeited. This is because blaming the blameless is unjust, even if it involves no thought about claim forfeiture, to the extent that it incorrectly accounts for their interests.

**Acknowledgments.** I am grateful to Gideon Rosen, Grace Helton, and Mercedes Corredor for their feedback on earlier versions of this paper and for discussion. I would also like to thank the two anonymous reviewers of this journal for their comments and suggestions; they were incredibly helpful.

**Aarthy Vaidyanathan** is a graduate student in the Department of Philosophy at Princeton University.

<sup>12</sup>I would like to thank Gideon Rosen for helpful conversation and suggestions regarding the formulation of the account.

<sup>13</sup>I would like to thank Reviewer #2 for pressing me to further explain elements of my account and for suggesting ways of doing so.

## References

- Caruso, Gregg D. 2018. "Skepticism about Moral Responsibility." In *The Stanford Encyclopedia of Philosophy* (Spring), edited by Edward N. Zalta.
- Levy, Neil. 2011. *Hard Luck: How Luck Undermines Free Will and Moral Responsibility*. New York: Oxford University Press.
- McKenna, Michael. 2008. "A Hard-Line Reply to Pereboom's Four-Case Argument." *Philosophy and Phenomenological Research* 77 (1): 142–59.
- Mele, Alfred. 1995. *Autonomous Agents*. New York: Oxford University Press.
- Menges, Leonhard. 2022. "The Kind of Blame Skeptics Should Be Skeptical About." *Canadian Journal of Philosophy*: 1–15. <https://doi.org/10.1017/can.2021.38>.
- Pereboom, Derk. 2014. *Free Will, Agency, and Meaning in Life*. New York: Oxford University Press.
- Rosen, Gideon. 2002. "The Case for Incompatibilism." *Philosophy and Phenomenological Research* 64 (3): 699–706.
- Scanlon, T. M. 2008. *Moral Dimensions: Permissibility, Meaning, Blame*. Cambridge, MA: Harvard University Press.
- Strawson, Galen. 1994. "The Impossibility of Moral Responsibility." In *Free Will*, edited by Gary Watson, 212–28. New York: Oxford University Press, 2003.
- Strawson, Peter F. 1962. "Freedom and Resentment." In *Free Will*, edited by Gary Watson, 72–93. New York: Oxford University Press, 2003.
- Wallace, R. Jay. 1994. *Responsibility and the Moral Sentiments*. Cambridge, MA: Harvard University Press.