

## **Responsibility, Regret, and Protest**

*Oxford Studies in Agency and Responsibility* 4, David Shoemaker, ed., Oxford: Oxford University Press, 2017, pp. 121-40.

Derk Pereboom, Cornell University

Penultimate draft

Abstract: Is morality viable without the notions of desert, moral demand, and moral obligation, notions threatened by possible limitations in human abilities? This article contends that it may well be. Instead of invoking desert, blame can be largely forward-looking, recast as appropriate moral protest, and aiming at protection, moral formation, and reconciliation. Moral demands in relationships can be re-envisioned as commitments deriving from care, and failure to act in accord with one's commitments can be conceived as wrong in the sense that they are appropriately protested. Moral obligations can be reconfigured as axiological recommendations. Revision of morality poses risks, but it is argued that the changes envisioned are practically viable.

Keywords: moral responsibility, free will, obligation, moral demand, moral protest, blame, praise, desert, moral reform, reconciliation, protection, guilt, regret.

### **Introduction**

Morality is shaped in part by our emotions, some of which introduce or enliven features of the practice that threaten to conflict with the limitations of human nature. These features of

the practice include, most prominently, moral desert and moral obligation. The applicability of these notions is put at risk by the causal determination of action, and by the luck that indeterminism arguably introduces. But in addition, the legitimacy of desert within general ethical frameworks such as consequentialism or Kantian universalizability is unclear, and one might question how essential obligation and the related notion of demand are to morality. Here I press forward with the project of exploring the viability of morality without these notions (Pereboom 2013, 2014, 2015), with a specific focus on accounting for blame and regret without desert, and for wrongness without obligation or demand.

### **Different senses of moral responsibility**

Our practice of holding each other morally responsible is complex. It involves a number of different aims, and a range of responses justified by those aims. In recent years a number of theorists have argued that this complexity can be unified, that there is ultimately a single notion of moral responsibility that unifies the practice. Proponents of this view include R. J. Wallace (1994) and George Sher (2006). I believe that a view of this sort misrepresents the practice to at least some degree, and here my potential allies include Gary Watson (1996), Dana Nelkin (2011), and David Shoemaker (2011, 2015). With them, I contend that a certain kind of pluralism about the practice is true.

Contrary to what is at times supposed in the discussion, I believe that it's mistaken to claim that the term 'moral responsibility' has a single sense. Rather, there are multiple senses of 'moral responsibility' corresponding to multiple aims and justifications, all of which are aspects of our practice of holding morally responsible. Purely linguistically, this is credible.

Often linguistic terms that have a long history of use have a number of senses. Even if a referring term with such a history originally had just one specific referent, over time it is apt to be applied to similar but distinct referents, thereby acquiring different senses. The terms 'moral responsibility' and 'blame' plausibly have this profile. Ludwig Wittgenstein (1953) famously makes the general point with the example of 'game.' There is a family resemblance among the referents of this word, but these referents do not exhibit a relatively simple, non-disjunctive necessary and sufficient condition for class membership, and the word may therefore have multiple senses.

The advocate of a single sense has several options. One is to argue that what might seem to be different senses have a common essence; perhaps each sense of blameworthiness features, at its core, that the wrongdoer deserves to be blamed (McKenna 2012), or to feel guilty (Clarke 2013; Carlsson 2016; Duggan ms); or that the wrongdoer is an appropriate target of reactive attitudes (Strawson 1962; Wallace 1994). A second is to weed out all but one sense on the ground that the others are not notions of genuine moral responsibility, or on the theoretical ground that simplicity in theory is preferable. I resist these strategies, partly because I think that they don't withstand scrutiny in their own right. But I also believe that some of the senses are best eliminated from the practice while others remain in place, and this proposal requires distinct senses. The grounds for holding that some are best eliminated are twofold. One is that they can successfully be criticized for ethical reasons, and a second is that they can be challenged on the ground that they presuppose a sort of freedom we might very well not have, and are thus not compatible with skepticism about this sort of free will.

Watson (1996) distinguishes between moral responsibility in the *attributability* sense and moral responsibility in the *accountability* sense. By his characterization, to be morally responsible for an action in the attributability sense is for that action to express what one stands for or who one is. In this way, acts that are courageous or cowardly, magnanimous or petty, license us to attribute to the agent something morally good or morally bad about her, perhaps a character trait. In light of such evaluations, we are often given reasons to adjust our behavior and modify our expectations. But Watson argues that a person's being merely responsible in the attributability sense does not entail that she is a legitimate target of indignation or of a demand that she rectify her bad conduct or apologize to those she wronged. These are all manifestation of moral responsibility in the accountability sense, the core of which is the legitimacy of holding someone to account for her actions. (Shoemaker (2011) provides a different characterization of this distinction.)

It's typically agreed that the holding-accountable aspect of the practice features the notion of desert. In the basic form of desert, someone who has done wrong for bad reasons deserves to be blamed and perhaps punished just because he has done wrong for those reasons, and someone who has performed a morally exemplary action for good reasons deserves credit, praise, and perhaps reward just because she has performed that action for those reasons (Feinberg 1970; Pereboom 2001, 2014; Scanlon 2013). More precisely:

For an agent to be *morally responsible for an action in the basic desert sense* is for the action to be hers in such a way that she would deserve to be blamed if she understood that it was morally wrong, and she would deserve to be praised if she understood that it was morally exemplary. The desert at issue here is basic in the sense that the agent, to

be morally responsible, would deserve to be blamed or praised just because she has performed the action, given sensitivity to its moral status; and not, for example, by virtue of consequentialist or contractualist considerations. (Pereboom 2014, 2001; cf. Feinberg 1970; Scanlon 2013)

There may in addition be senses of moral responsibility that involve a non-basic variety of desert. Essentially forward-looking notions of holding agents deserving of blame and punishment have been defended on consequentialist or contractualist grounds (Dennett 1984, 2003; Lenman 2006; Vargas 2007, 2013; Vilhauer 2009). But there might be no such thing. On one type of revisionary account, our practice of holding agents morally responsible in a desert sense should be retained, not because we are in fact morally responsible in this sense, but because doing so would have the best consequences relative to alternative practices. Daniel Dennett (1984, 2003) advocates a version of this position, as does Manuel Vargas (2007, 2013). One question for Dennett and Vargas is whether, given their positions, what we call deserved responses are really just negative or positive incentives. We might not ordinarily regard them merely as incentives, but on this view their justification is ultimately forward-looking, founded solely on the value of consequences that we expect to result. And if these responses are in fact just incentives, then it would seem that they are not also deserved.

However, not all aspects of our practice of holding morally responsibility invoke desert. Consider, for example, T. M. Scanlon's account of blame. In his analysis, to blame an agent for an action is to judge that it reveals something about the agent's attitude toward oneself and/or others that impairs the relations that he can have with them, and to take one's relationship with him to be modified in a way that this judgment of impaired relations justifies as

appropriate (Scanlon 2009: 128-31; for valuable development of this notion, see Shoemaker 2011: 619-24). Whether blame defined in this way invokes desert depends on how we are to understand the appropriateness to which this characterization refers. Scanlon does specify that he conceives this notion of appropriateness to have an aspect that underwrites basic desert justifications for limited harms. Public acknowledgement of wrongdoers' faults, when it does not involve excessive humiliation, "is made appropriate by these faults themselves (it is deserved on their basis alone), and they therefore have no moral claim against it" (Scanlon 2013).

But there are ways to construe the appropriateness featured in this account that do not involve desert (Pereboom 2013). A first is epistemic and evidential, and an example of Scanlon's illustrates this reading. You trusted Bill, but you then noticed that he behaves in an untrustworthy manner, as a result of which it is now appropriate for you to take your relationship with him to be impaired to the degree that reflects this diminished trust. Here the justification is evidential. You believed Bill was trustworthy to a high degree, but you then acquired good evidence of disloyalty, and hence a good reason to judge that an attitude of his was relationship-impairing. You now make this judgment, and take your relationship with him to be impaired to a degree that it justifies as appropriate. No notion of desert is in play. Second, even if the appropriateness is interpreted as specifically moral, it need not involve desert. You have the right to protect yourself against Bill, and this may require ceasing to confide in him. To justify the appropriateness of this protective measure, one might appeal to the right to self-defense: the justifying aim in ceasing to confide in Bill is to protect yourselves against the danger to you he poses. This self-defense justification does not involve basic desert. The very

same justification will legitimately serve in case Bill's behavior is due to a psychological disorder that precludes moral responsibility in the basic desert sense.

There are reasons to be skeptical of any notion of moral responsibility that involves desert. One worry for the basic desert sense is that for an agent to basically deserve a harmful response she must have a kind of free will that is unavailable to us, and the free will skeptic contends that this concern can't be successfully countered (G. Strawson 1986; Waller 1990, 2011; Pereboom 1995, 2001, 2014). As noted, one might argue that some desert sense of moral responsibility can or should be retained because doing so stands to bring about good results, but on this view what might appear to be deserved responses would seem to be mere incentives. Another worry is that for a number of contending general normative ethical theories the notion of desert seems to have the role of an awkward supplement. Any place for desert in typical consequentialist views is uncomfortable, and despite Kant's well-known (1791/1963) invocation of desert in justifying criminal punishment, that appeal appears not to be justified by any formulation of the Categorical Imperative, which he held to be the supreme and comprehensive moral principle. A further issue is that the conception of deserved harm, at least in its basic form, would seem to involve the idea of harm as an intrinsic good, which is dubious. Motivated by these concerns, I will propose a view that rejects all desert-involving senses of moral responsibility.

### **A desert-free sense of moral responsibility**

The notion of moral responsibility, and blame in particular, that I develop and endorse (Pereboom 2013, 2014, 2015) is largely forward-looking. Blaming is, in its paradigm cases, a kind of calling to account, and is justified by these forward-looking elements:

1. The right of those wronged or threatened by wrongdoing to protect themselves and to be protected from immoral behavior and its consequences.
2. The good of reconciliation with the wrongdoer,
3. The good of the moral formation of the wrongdoer.

Immoral actions are often harmful, and we have a right to protect ourselves and others from those who are disposed to behave harmfully. Immoral actions can also impair relationships, and we have a moral interest in undoing such impairment through reconciliation. And because we value morally good character and resulting action, we have a stake in the formation of moral character when it is beset by dispositions to misconduct.

There is an account of praise that corresponds to this conception of blame. Of our three goals of blaming, moral formation, protection, and reconciliation, the one most clearly amenable to praise is moral formation. We praise an agent for a morally exemplary action to strengthen the disposition that produced it. This can have a protective function, since strengthening such dispositions has the effect of reducing the incidence of dangerous behavior. Corresponding to reconciliation is the notion of celebrating successes in a relationship. Praising an action has this celebratory function as well.

Michael McKenna (2012) has developed a conception of moral responsibility that turns out, with a few key revisions, to be amenable to the view I'm proposing. In his *conversational* theory of moral responsibility, actions of a morally responsible agent are potential bearers of a

type of meaning by indicating the quality of will that resulted in the action (2012, pp. 92-94; see also Arpaly 2006). Blaming an agent who manifests an immoral quality of will in action is an expression of an attitude such as moral resentment or indignation, and its function is to communicate to him a moral response to the indicated quality of will. Morally responsible agents understand that members of the moral community might attribute such a meaning to their actions. When actions are morally charged, they understand themselves to be introducing a meaningful contribution to such a conversational exchange. McKenna labels this initial stage of the conversation *moral contribution*. In the case of a prima facie immoral action, in the second stage the agent is blamed by a respondent; he calls this stage *moral address*. In the third stage, *moral account*, the blamed agent offers an excuse, a justification, or an apology. The respondent might at this point continue the conversation by forgiving or punishing the wrongdoer. In a subsequent stage the blamed agent may be restored to full status in the moral community. McKenna points out that not all blaming conforms to this model; blaming the dead, for instance, does not. Here he invokes a paradigm-similarity model for the meaning and extension of a concept (Rosch 1972, 1973). The blame conversation as he describes it is a paradigm case of blaming, and examples of other sorts, such as blaming the dead, qualify as instances of blaming because they are sufficiently similar to such paradigm cases.

I endorse such a model with a few modifications, in particular that the justifying aims of blame are protection, reconciliation, and moral formation, and that contrary to McKenna's version, desert is excised (Pereboom 2013, 2014, 2015). Here I want to point out that on this model it's the agent's responsiveness to reasons that is engaged in the envisioned process for both blame and praise. In the case of blame, at the stage of moral address we request an

explanation with the intent of having the agent acknowledge a disposition to act badly, and then, if he has in fact so acted without excuse or justification, we aim for him to come to see that the disposition issuing in the action is best eliminated. In the standard sort of case, this change is produced by way of the agent's recognition of moral reasons to make it. More generally, it is an agent's responsiveness to reasons (Fischer (1982, 1994); Fischer and Ravizza (1998)), together with the fact that we have a moral interest in our protection, his moral formation, and our reconciliation with him, that explains why he is an appropriate recipient of blame in this conversational and largely forward-looking sense.

Let me note a difference between a forward-looking and a backward-looking conversational, answerability notion of responsibility. On the backward-looking version, requests for justification of morally questionable attitudes presuppose that those attitudes are grounded in reasons, which the agent can then cite in response (Shoemaker 2011: 609-12). In the forward-looking version, by contrast, agents can also legitimately be called to account for attitudes that up to now float free of their reasons (Shoemaker 2011: 611), so long as moral reasons are apt to be accessed as a result of the conversation, reasons whose acceptance stands to result in protection, reconciliation, or moral formation.

In fact, the key necessary condition for moral responsibility in this largely forward-looking sense is sensitivity or responsiveness to reasons. A condition of this kind is typically advanced as a necessary condition for basic desert responsibility by philosophers who maintain that this sort of responsibility is compatible with the action's causal determination, and that it isn't explained by her ability to do otherwise. The largely forward-looking sort of moral responsibility I advocate is also compatible with agents' being causally determined in their

actions by factors beyond their control, and in an appropriately constructed deterministic manipulation argument the manipulated agent will be morally responsible in this way (Pereboom 2014). In my view, agents will also be morally responsible in the forward-looking sense in Frankfurt examples, and thus, just as in the compatibilist position that Fischer and Ravizza (1998), McKenna (2012) and Carolina Sartorio (2016) develop, the kind of freedom or control required for such moral responsibility for an action will be a matter of its actual causal history.

The compatibilists just mentioned are all committed to sensitivity or responsiveness to reasons as the key condition on basic-desert or reactive-attitudes-involving moral responsibility, while I instead view it as the most significant condition for a notion of responsibility that focuses on protection, reconciliation, and moral formation. It's noteworthy that there is a transparent connection between reasons-sensitivity and the forward-looking notion of responsibility, and that a similar tie is lacking for basic-desert and reactive-attitudes-involving moral responsibility. Sensitivity to reasons is evidently required for an agent to be subject to moral formation on the basis of presentation of reasons, while the link between reasons-sensitivity and an assumption of basic desert or the reactive attitudes is more opaque. We may happen to be so constructed as to assume basic desert or have basic desert-linked reactive attitudes when and only when reasons-sensitive agents act badly, but here the transparent connection that we see with the forward-looking notion of responsibility is absent.

### **Blame as Moral Protest**

What is morally distinctive about blaming given the rejection of desert? One might propose that it essentially involves a judgment of relationship-impairment, or the expression of such a judgment, in accord with Scanlon's (2009) account. Again, on his view, to blame an agent is to judge that an action of his reveals something about his attitude toward oneself and/or others that impairs the relations that one can have with him, and to take one's relationship with him to be modified in a way that this judgment of impaired relations justifies as appropriate. Recall that such a view need not invoke desert, if desert is excised in favor of an epistemic notion of appropriateness, or a moral notion that invokes the right to take protective measures in self-defense.

However, such an account will not accommodate various kinds of blaming that I would want to retain. Blame in human relationships would seem only seldom linked with a judgment of relationship impairment. Blame is fairly frequently directed toward children by parents and teachers in the process of moral education, but in only a small minority of such cases does such blame involve a judgment that a relationship has been impaired. More generally, as Susan Wolf (2011) argues, relationships are resilient to wrongdoing in the sense that when a participant in a relationship does wrong and is appropriately blamed, there is no impairment of the relationship. Your son from time to time plays video games instead of doing his homework. You blame him, and do so appropriately, but the relationship is not impaired at all. Your husband on occasion forgets to do errands he's promised to do. He's appropriately blamed, but without any relationship impairment. These are relatively minor wrongs, but the same can hold for wrongs of a more serious sort. Suppose that your spouse is prone to unjustified and inappropriate expressions of anger and frustration in a way that many people are – not involving violence, but

still burdensome for you. But you knew this prior to marriage; in fact, you made the choice to marry in full knowledge of these traits. The good qualities far outweighed this downside. Although the spouse is appropriately blamed for the inappropriate expressions of anger, the relationship is resilient to and not impaired by this behavior.

Instead, I propose that the notion of *moral protest* is the essence of blame in the sense I'm setting out. For a precursor, Pamela Hieronymi (2001: 546) contends that resentment is best understood as moral protest; "resentment protests a past action that persists as a present threat." Resentment is not a feature of the sense of blame I have in mind. But on Hieronymi's account a key function of blaming someone is to protest a past action of his that persists as a present threat, and I accept, in the account of blame I endorse, that this is one of its core functions.

Angela Smith's recent account of blame also features a notion of moral protest, this time as the essence of blame (2013: 43):

**Smith's Moral Protest Account:** To blame another is to judge that she is blameworthy (i.e. that she has attitudes that impair her relationships with others) and to modify one's attitudes, intentions, and expectations toward that person as a way of protesting (i.e. registering and challenging) the moral claim implicit in her conduct, where such protest implicitly seeks some kind of moral acknowledgement on the part of the blameworthy agent and/or on the part of others in the moral community.

I accept Smith's core idea: in the notion of blame I endorse, appropriate moral protest is its essence. As for the details of her account, I want to register two caveats. First, as I've argued, I don't think that judging blameworthy essentially involves judging that the target has attitudes

that impair a relationship. Second, I don't think that typical cases of blame involve protesting a moral claim implicit in conduct. Smith adduces a case of blame for racial discrimination that arguably does -- perhaps the claim that respect is not due to people of a certain race. But protest against wrongdoing due to lack of self-control and or against minor omissions, such as the husband forgetting to buy the milk (Clarke 2014) doesn't seem to target a moral claim implicit in conduct. But it protests behavior nonetheless.

I propose a leaner view:

**Moral Protest Account of Blame:** For B to blame A is for B to issue a moral protest against A for the immoral conduct that B attributes (however accurately) to A.<sup>1</sup>

The immoral conduct will typically be an immoral action, but there are cases in which the action considered separately from the reasons for which it's performed is not wrong, but the reasons make the overall conduct wrong (e.g., Haji 1998, Hanser 2005). Sometimes blame is misplaced, since no immoral conduct has taken place, but the protest can still count as blame. This can happen when B believes A to have acted badly but the belief is false, due perhaps to misinformation or improper consideration of evidence. This can also happen when B does not believe that A acted badly but nonetheless attributes bad conduct to B, as in cases of politically motivated false accusation. It's often the case that blame functions, as in Hieronymi's proposal, as a moral protest of a past action that persists as a present threat, and I agree that this is one highly important objective for blame. But not all blame has this point, as when we blame the dead, or blame someone who is alive but lacks a persisting disposition to act badly -- someone, for instance, who has already undergone moral reform. In such cases protest can yet have the

---

<sup>1</sup> Thanks to Daniel Telech for discussion of this formulation.

function of explicitly noting immoral conduct as immoral, which might also have the aim of general moral improvement. In the example of the already-reformed wrongdoer, blame might still function as a step in the process of reconciliation. While this moral protest account captures much of blaming behavior, it is consistent with the revisionary view that desert presuppositions, which are widespread, have no justified role in blaming, and that its legitimate aims are instead those I've set out. (Let me note that McKenna's conversational account of blame might also be recast as a moral protest account.)

There are cases of persisting dispositions to act badly where the disposition has yet to be manifested in action. For instance, someone might reveal a disposition to violent behavior, but without ever having acted on it. One might then legitimately protest that disposition, but on the account as stated, this wouldn't all by itself count as blame, since B's blaming A requires, conceptually, that B blame A for what B represents, either truly or falsely, as A's immoral conduct. One might extend the account to allow for the possibility of blame for dispositions not yet manifested, but my own sense is that this would not count as blame but rather as pre-emptive intervention.

### **Regret and Deserved Guilt**

In my account, blame as moral protest is to be justified in the largely forward-looking way I've specified, by the aims of protection, reconciliation, and moral formation. But the immediate target of blame is typically a past action, and in this respect such blaming will have a backward-looking aspect: the badness of the past act is part of what makes the protest

appropriate. Does this backward looking aspect invoke deserved or basically deserved pain or harm?

In accord with Randolph Clarke's (2013) suggestion, it's valuable to consider whether a wrongdoer deserves or basically deserves to feel guilty and the pain that accompanies it (see also Duggan, ms). Clarke proposes, first, that there is value in the recognition by an agent who is blameworthy that he is blameworthy. A further response, the feeling of guilt, would provide a morally valuable and intuitively fitting addition to this acknowledgement. This response would intuitively have value insofar as it expresses moral concern for having done wrong and for those wronged.

I agree that regret, which features pain, is a morally fitting additional reaction to one's own wrongdoing without accepting that wrongdoers ever deserve or basically deserve to feel guilty and to experience the pain of guilt. Two of my allies on this point, Bruce Waller (1990) and Hilary Bok (1998), argue that the fittingness of a pained feeling can be accounted for by a recognition that one has not lived up to one's standards for morality and self-control, without the need to invoke desert. Bok sets out an example in which one has done something wrong, on account of which one suffers a painful response, which she compares to heartbreak (1998: 168-69). She calls this response 'guilt,' but I'd like to substitute 'regret' for Bok's 'guilt,' reserving 'guilt' for a desert-involving feeling:

The relation between the recognition that one has done something wrong and the guilt one suffers as a result... is like the relation between the recognition that one's relationship with someone one truly loves has collapsed and the pain of heartbreak. Heartbreak is not a pain one inflicts on oneself as a punishment for loss of love; it is not

something we undergo because we deserve it... Similarly, the recognition that one has done something wrong causes pain. But this pain is not a form of suffering that we inflict on ourselves as a punishment but an entirely appropriate response to the recognition of what we have done, for two reasons. First, our standards define the kind of life we think we should lead and what we regard as valuable in the world, in our lives, and in the lives of others. They articulate what matters to us, and living by them is therefore by definition of concern to us. If we have indeed violated them, we have slighted what we take to be of value, disregarded principles we sincerely think we should live by, and failed to be the sorts of people we think we should be. The knowledge that we have done these things must be painful to us.

I think Bok is right to contend that feeling pain on account of a recognition that one has not lived up to one's moral standards or standards for self-control need not involve desert.

Here are two additional analogies: (1) one might appropriately feel pained that one failed to meet one's standards for chess playing when one understands that one's substandard performance is due to factors beyond one's control, while this pain is not deserved; (2) it might well be appropriate, and basically so, to feel the pain of grief upon the death of a loved one, while this pain is not deserved. Saliently, these cases feature the appropriateness of feeling pain without feeling pain being deserved.

Accordingly, I contend that it's appropriate that wrongdoers feel regret for what they've done, where regret, unlike guilt, does not involve deserved or basically deserved pain or harm. How can such feelings of pain be appropriate -- and even basically so -- but not deserved or basically deserved? One part of the answer is that the general sort of phenomenon illustrated

by the heartbreak, chess, and grief cases, does in fact occur. In addition, I suggest that feeling guilt involves pain that one regards as *appropriately imposed* – perhaps impersonally -- because one has done wrong. In the case of grief, the pain doesn't have this feature. Thus one possibility is that to count as basically deserved, by contrast with merely basically appropriate or fitting, the pain must in fact be appropriately imposed. In the passage from Hilary Bok, she makes two claims regarding her analogy, the pain of heartbreak. The first is that it is not a pain that one inflicts on oneself as a punishment. The second is that it is not something we undergo because we deserve it. The passage suggests that she conceives of the claims as linked: the pain of heartbreak isn't properly inflicted or imposed and thus not deserved, and the same is true of the pain of regret.

There are other justifications for regret that don't invoke desert. Ben Vilhauer (2004) advocates an account of a pained response upon wrongdoing that grounds it in sympathy with those one has wronged, and according to which such regret is fitting because the sympathy is morally appropriate. It's credible that such sympathy-based regret can serve to motivate repentance and moral self-improvement, for reconciliation with those one has wronged, and subsequent restoration of one's integrity. Vilhauer argues that because such sympathy-based remorse is also other-directed rather than merely self-directed, it is morally preferable to guilt or remorse grounded in basic desert. Guilt on a basic desert conception has no essentially forward-looking moral objective. By contrast, sympathy-based remorse involves taking on the perspective of the agent one has wronged, which has morally beneficial consequences. I can take Vilhauer's notion on board, and I'm happy to do so.

In summary, a painful response to one's own wrongdoing is intuitively appropriate. But

the justification for the appropriateness of such a painful response need not invoke desert. Instead, it might appeal to the appropriateness of backward-looking sentiments that don't involve desert, and to forward-looking considerations such as the moral reform of the wrongdoer and sympathy and reconciliation with those who were wronged.

### **Blame and obligation**

Against the largely forward-looking account of moral responsibility I've proposed one might object that for an agent to be appropriately blamed in the protest-involving sense requires that she ought not to have performed the action, and this in turn requires that she could have avoided the action, which is incompatible with causal determination of action. While free will skeptics need not endorse general causal determination of action, they tend to leave it open as a serious possibility. And according to the Consequence Argument, causal determination is incompatible with the ability to do otherwise. Thus causal determination threatens to undercut blameworthiness even in this largely forward-looking sense.

Still, everyone should agree that there are senses of 'could have done otherwise' or 'could have avoided' that are determinism-friendly. For an example inspired by David Lewis (1976), suppose you have the ability to speak both English and French. You're in café in Paris, but you order your coffee in English. Your friend says to you: "You could have ordered in French instead!" It would seem that your friend was correct to say what she did no matter what the truth about determinism turns out to be. So it appears that any reasonable person should accept that there are determinism-friendly senses of 'could have done otherwise.' Perhaps it's

such a sense that's required for the 'ought' of moral obligation, and this type of view is endorsed, for example, by David Brink and Dana Nelkin (2013).

However, such ability claims may invoke only general sorts of abilities. And as Clarke (2009) has argued, the issue between the compatibilist and the incompatibilist may be whether determinism is compatible with an agent's exercise of such a general ability on a particular occasion. The incompatibilist can, after all, readily allow that there are general abilities, such as the ability to speak French in the above example, which one retains at the time one is not exercising it. But she may not grant that a causally determined agent could in fact have exercised such a general ability at a time when he does not in fact do so.

A conclusion one might draw is that there may be a more specific and stronger notion of an ability to do otherwise that is not determinism-friendly. Carl Ginet (1990, Chapter 5) argues that there is reason to think that there are contexts of inference in which we assume such a more specific and stronger notion of an ability to do otherwise, which on his view cannot be accounted for by any such compatibilist semantic proposal. If he is right, the answer to the question: "Is being able to do otherwise compatible with determinism?" depends on which sense you mean. Yet it may be that a determinism-friendly sense of 'could have done otherwise' is all that's required by the 'ought' of moral obligation. I won't try to resolve that issue here. But suppose that moral obligation requires a sense that's not determinism-friendly. What options do we then have?

A second and related concern is that if moral obligation is undercut by the general causal determination of action, moral wrongness is also undermined. Ishtiyaque Haji (1998, 2002) argues that S has a moral obligation to perform A just in case it is morally wrong for S not to

perform A, and, similarly, that S has a moral obligation not to perform A just in case it is morally wrong for S to perform A. Thus moral wrongness and moral obligation stand and fall together. I've claimed that Angela's Smith (2004) notion of moral judgment, which invokes the normative notion of moral wrongness in the case of immoral action, is not threatened by causal determination. But now it seems as if it is. Let's consider these two objections in turn.

I've developed a response to the first objection that takes advantage the plausible claim that 'ought' has a range of correct uses, and as C. D. Broad (1952) suggests, it may be that not all are linked to an "ought-implies-can" (OIC) requirement, or at least to an OIC requirement that underwrites the incompatibility of causal determinism and 'ought' judgments (Pereboom 2013, 2014, Chapter 6). Let me summarize. Ruth Marcus (1966), Lloyd Humberstone (1971), and Gilbert Harman (1977: 87) distinguish between an 'ought' that applies to action and one that applies to states of affairs. An 'ought to do,' Harman contends, "implies that an agent has a reason to perform an action, while an 'ought to be' evaluates a state of affairs and does not by itself imply that any particular agent has a reason to contribute to bringing about that state of affairs. Mark Schroeder (2011) illustrates this latter sense with 'Larry ought to win the lottery' where Larry has been subject to a series of undeserved misfortunes. The 'ought to do' sense is arguably governed OIC, while the 'ought to be' is not.

Kate Manne (2011) points out that the evaluative 'ought' applies not only to non-agential states of affairs, but also to actions. So even if the abuse-addict is not free to refrain from abusive behavior, it is still the case that he ought not, in the evaluative sense, behave abusively. Thus we now have two notions of 'ought' that apply to action. One, the 'ought to be' we can call the 'ought' of *axiological evaluation*, and it does not entail a 'can' claim. The other,

the 'ought to do,' does entail that the agent can perform the specified action. Nelkin (2011: 111) contends that 'ought' propositions that specify what an agent is morally obligated to do are essentially action-directed, so that if 'S ought not do A' is true, then as a matter of the meaning of such 'ought' judgments, or of the essential nature of moral obligation, S is thereby directed to refrain from A, and this entails that S can refrain from A. I think that this is correct, and I call this sense the 'ought' of *specific action demand*.

Thus it may be that if causal determination precludes alternative possibilities for motivation and action in a sense relevant to the issues under consideration, there is a core notion of 'ought to do' that will be undermined. If I know that in this relevant sense an agent could not have avoided lacking the motivation required for refraining from performing an immoral action, it would be unfair, and I think, mistaken, to claim that she ought not have performed that action at that time. But it might well not be mistaken or unfair for me to recommend to that agent that she not perform an action of that type in the future, supposing that it's reasonable to believe that she has or will have the abilities and opportunity to refrain from performing the action, and that it's epistemically open that she will acquire the requisite motivation, and in particular if it's reasonable to believe that making this recommendation would contribute causally to bringing about the motivation. To recommend refraining from the action to her, I might tell her that she ought not perform the action at the future time, and do so appropriately and without making any kind of mistake. In my view, the sense of 'ought' invoked here would need to be distinct from the 'ought' of specific action demand.

Given determinism and that determinism precludes alternatives, when one tells an agent that he ought to refrain from performing an action of some type in the future, it's not the

'ought' of specific action demand, but rather the 'ought' of axiological evaluation that's legitimately invoked. This use of 'ought' proposes as morally valuable a state of affairs in which the agent refrains from performing the action and recommends that she not perform it. I call this the ought of *axiological recommendation*. Unlike the 'ought' of specific action demand, it is not an 'ought' of obligation. Supposing the general causal determination of action and that such determination rules out the relevant alternative possibilities for action, the use of the 'ought' of specific action demand could be correct and fair only if the agent is in fact causally determined to perform the action and one is reasonably sure that she is, which would typically be untrue. But under these circumstances the 'ought' of axiological recommendation would not be similarly undermined.

Blameworthiness in the largely forward-looking sense I've set out licenses the right sort of respondent to protest the action by telling the agent who has acted badly that he ought not act this way, where 'ought' has the sense of axiological recommendation, which in turn requires that it be epistemically open for the respondent that the agent will comply with this recommendation. Thus for an agent to be liable to my appropriately blaming him for an action in the forward-looking sense it must be epistemically open to me that he refrain from performing actions of this type in the future. But unlike what's required for a legitimate use of the ought of specific agent demand, it need not be that the agent could have refrained from performing the action at issue.

### **Demands, wrongness, and protest**

Dana Nelkin (2014) asks whether relationships such as friendship are possible supposing

that the 'ought' of moral obligation and specific action demand is ruled out. On a plausible proposal, what sets friendship apart from mere acquaintances is a structure of obligations. For example, if A and B are friends, and if A is in dire need, then there is a strong prima facie moral obligation for A to help, with an ought of specific agent demand in play.

But consider replacing the notions of obligation and demand with those of care and commitment. General moral care – care for beings with moral standing – is apt to generate specific commitments in particular contexts. In the case of friendship this sort of care would generate a commitment on the part of each friend to provide help to the other when in dire need. On this account, how do we conceive the interpersonal moral relationship between friends A and B in a situation in which there is a threat that A will not come to B's help in time of dire need, and thus not to honor her own commitment? Here it's natural for A to think of himself as morally obligated to help B, and for B to make a demand of moral obligation that A help her. But instead, we can frame B's expectation in terms of the notion of moral wrongness: A can legitimately think that it would be wrong for him not to help B, and B can communicate to A that it would be wrong of him not to help her.

However, this introduces a further threat, according to which judgments of normative wrongness are undermined by causal determination. Haji (1998, 2002) argues that due to the tight connection between moral obligation and moral wrongness, the threat posed to judgments of moral obligation extends to those of moral wrongness. Crucial to his argument is the following principle:

S has a moral obligation to perform A if and only if it is morally wrong for S not to perform A.

If this principle is true, then if judgments of moral obligation are ruled out by causal determination, judgments of moral wrongness will be excluded as well. But although this biconditional principle may be attractive considered in the abstract, it captures only one aspect of the complex notion of moral wrongness. This can be made plausible by reflecting on the fact that the right-to-left half:

If it is morally wrong for S not to perform A, then S has a moral obligation to perform A

is not clearly secure. There's likely no example in which it's credible that an agent has a moral obligation not to perform an action while it is not morally wrong for her to perform it. But there are cases in which it's intuitive that performing an action would be morally wrong for an agent, while it's at least less clear that she has a moral obligation not to perform it. Imagine that a psychopath could not have avoided an act of deception due to his psychological disorder. The 'ought-implies-can' principle provides an intuitive basis for denying that he was morally obligated to refrain from deceiving, while it remains intuitive that his action was morally wrong. (Pereboom 2001, 2014)

An alternative notion of moral wrongness, one that isn't biconditionally linked to moral obligation, accommodates this intuition. The core of Alastair Norcross's (2006) proposal for a purely axiological ethics involves specifying for each action-relevant situation the pertinent options for acting ranked in order of value of the consequences realized, without obligation to maximize value. This position meshes nicely with the foregrounding of the 'ought' of axiological recommendation that's already in place. We can now specify that actions are wrong when their axiological ranking is sufficiently low. But how low is that? Here we can again appeal to the notion of protest which I've already used to characterize blame:

**An axiological notion of wrongness:** An option for acting is morally wrong when its value is low enough in the axiological ranking for it to be morally appropriate for a relevantly positioned interlocutor to issue a moral protest against actions of this sort.

Given the Norcross view, moral protest functions as an axiological recommendation against performing an action. What makes such protest morally appropriate would itself consist in its having a relatively high ranking among available options. Note also that this proposal does not characterize wrongness independently of when it is appropriate to protest, and so it cannot ground the appropriateness of protest in wrongness. But it does have the consequence that it's appropriate to protest a type of action only when so acting would be morally wrong, and it thus satisfies an important intuition we have about this relationship. Given the previous characterization of blame as appropriate moral protest, this account of wrongness forges an intimate and intuitively correct relationship between the two notions. Since blame is appropriate moral protest and a wrong action is one that's appropriately protested, an action will be blameworthy just in case it's wrong. To pre-empt a counterexample, I'm counting blaming the dead as morally appropriate protest just in virtue of the appropriateness of explicitly noting wrongdoing.

Note that I characterize wrongness in terms of the justifiability of protest, and not in terms of the appropriateness of resentment or indignation. These reactive attitudes are essentially backward looking insofar as they presuppose or are closely associated with a presumption of basic desert. By contrast, the notion of moral protest and the attendant notion of wrongness are largely forward-looking and do not involve basic desert.

This proposal also yields an account of rights without corresponding duties, a further concern Nelkin (2014) raises. Suppose that by virtue of the friendship between A and B, A intuitively has a right to B's help in time of dire need. In Nelkin's view, an essential feature of friendship that becomes salient in this situation is B's moral obligation to help A. On my proposal, what's salient is rather that given the nature of friendship, B has a commitment, grounded in care, to help A, and it's appropriate for A to protest B's not helping her and to claim that it is wrong for B not to help. Thus the role of moral obligation in a relationship has a near functional equivalent whose constituents are care and its resulting commitments, together with the appropriateness of moral protest in cases in which commitments are not honored.

### **Final Words**

Is morality viable without desert, demand, and moral obligation, notions threatened by possible limitations in human abilities? I have argued that it may well be. Instead of invoking desert, blame can be largely forward-looking, recast as appropriate moral protest, and aiming at protection, moral formation, and reconciliation. Feeling bad about a wrong one has done also need not involve desert, but instead might be reconceived on analogy with the appropriate pain of heartbreak or grief. Moral demands in relationships can be re-envisioned as commitments deriving from care, and failure to act in accord with one's commitments can be conceived as wrong in the sense that they are appropriately protested. Obligations can be reconstrued as axiological recommendations. Revision of morality poses risks, but the reduction

of anger and demand holds out the promise of a more compassionate and humane moral practice.<sup>2</sup>

---

<sup>2</sup> Thanks to audiences at Florida State University, the NOWAR conference at Tulane University, and at the University of California, San Diego for commentary and discussion. Thanks specifically to Gunnar Björnsson, David Brink, Randy Clarke, Austin Duggan, Michael McKenna, Dana Nelkin and two anonymous reviewers for valuable comments.

## References

- Arpaly, Nomy. (2006). *Merit, Meaning, and Human Bondage*. (Princeton, NJ, Princeton University Press).
- Brink, David, and Nelkin Dana. (2013). "Fairness and the Architecture of Responsibility," *Oxford Studies in Agency and Responsibility*, pp. 284-313.
- Bok, Hilary. (1998). *Freedom and Responsibility*. (Princeton: Princeton University Press).
- Broad, Charlie. D. (1952). "Determinism, Indeterminism, and Libertarianism," in his *Ethics and the History of Philosophy* (London: Routledge and Kegan Paul), pp. 195-217.
- Carlsson, Andreas. B. (2016). "Blameworthiness as Deserved Guilt," *Journal of Ethics* doi:10.1007/s10892-016-9241-x, pp. 1-26
- Clarke, Raldolph. (2009). "Dispositions, Abilities to Act, and Free Will: The New Dispositionalism," *Mind* 118, pp. 323-51.
- Clarke, Randolph. (2013). "Some Theses on Desert," *Philosophical Explorations* 16, pp. 153-64.
- Clarke, Randolph. (2014). *Omissions*. (Oxford: Oxford University Press).
- Dennett, Daniel. (1984). *Elbow Room*. (Cambridge: MIT Press).
- Dennett, Daniel. (2003). *Freedom Evolves*. (New York: Viking).
- Duggan, Austin. ms. "Behind the Eyes of the Transgressive: Guilt as Definitive of Moral Responsibility"
- Feinberg, Joel. (1970). *Doing and Deserving*. (Princeton: Princeton University Press).

- Fischer, John. M., and Mark Ravizza. (1998). *Responsibility and Control* (Cambridge: Cambridge University Press).
- Haji, Ishtiyaque. (1998). *Moral Appraisability*. (New York: Oxford University Press).
- Haji, Ishtiyae. (2002). *Deontic Morality and Control*. (Cambridge: Cambridge University Press).
- Hanser, Matthew. (2005). "Permissibility and Practical Inference," *Ethics* 115, pp. 443-70.
- Harman, Gilbert. (1977). *The Nature of Morality*. (Princeton: Princeton University Press).
- Hieronymi, Pamela. (2001). "Articulating an Uncompromising Forgiveness," *Philosophy and Phenomenological Research* 62, pp. 529-54.
- Humberstone, Lloyd. (1977). "Two Sorts of 'Oughts'," *Analysis* 32, pp. 8-11.
- Kant, Immanuel. (1797/1963). *The Metaphysical Elements of Justice*, tr. John Ladd (New York: Bobbs-Merrill).
- Lenman, James. (2006). "Compatibilism and Contractualism: The Possibility of Moral Responsibility," *Ethics* 117, pp. 7-31
- Lewis, David. (1976). "The Paradoxes of Time Travel," *American Philosophical Quarterly* 13: 145-152.
- Manne, Kate. (2011). *Not by Reasons Alone*. (MIT PhD Dissertation).
- Marcus, Ruth. B. (1966). "Iterated Deontic Modalities," *Mind* 75, pp. 580-82.
- McKenna, Michael. (2012). *Conversation and Responsibility*. (New York: Oxford University Press).

- Nelkin, Dana. (2011). *Making Sense of Moral Responsibility*. (Oxford: Oxford University Press).
- Nelkin, Dana. (2014). Symposium contribution on *Free Will, Agency, and Meaning in Life*, in *Science, Religion and Culture* 1, pp. 209-17.
- Norcross, Alastair. (2006). "Reasons without Demands: Rethinking Rightness," in J. Dreier, ed., *Blackwell Contemporary Debates in Moral Theory*. (Oxford: Blackwell).
- Pereboom, Derk. (2001). *Living without Free Will*. (Cambridge: Cambridge University Press).
- Pereboom, Derk. (2013) "Free Will Skepticism, Blame, and Obligation," in *Blame: Its Nature and Norms*, N. Tognazzini and D. J. Coates, eds., (New York: Oxford University Press), pp. 189-206.
- Pereboom, Derk. (2014). *Free Will, Agency, and Meaning in Life*. (Oxford: Oxford University Press).
- Pereboom, Derk. (2015). "A Notion of Moral Responsibility Immune to the Threat from Causal Determination," *The Nature of Moral Responsibility* Randolph Clarke, Michael McKenna and Angela Smith, eds. (New York: Oxford University Press), pp. 281-96.
- Rosch, Eleanor. (1972). "Universals in Color Naming and Memory," *Journal of Experimental Psychology* 93, pp. 10-20.
- Rosch, Eleanor. (1973). "Natural Categories," *Cognitive Psychology* 4, pp. 328-50.
- Sartorio, Carolina. (2016). *Causation and Free Will*. (Oxford: Oxford University Press).
- Scanlon, Thomas. M. (2009). *Moral Dimensions*. (Cambridge: Harvard University Press).
- Scanlon, Thomas. M. (2013). "Giving Desert Its Due," *Philosophical Explorations* 16, pp. 101-16.
- Schroeder, Mark. (2011). "Ought, Agents, and Actions," *Philosophical Review* 120, pp. 1-41.

- Sher, George. (2006). *In Praise of Blame*. (New York: Oxford University Press).
- Shoemaker, David. (2011). "Attributability, Answerability, and Accountability: Toward a Wider Theory of Moral Responsibility," *Ethics* 121, pp. 602-32.
- Shoemaker, David. (2015). *Responsibility from the Margins*. (Oxford: Oxford University Press).
- Smith, Angela. (2008). "Control, Responsibility, and Moral Assessment," *Philosophical Studies* 138.
- Smith, Angela. (2013). "Moral Blame and Moral Protest," *Blame: Its Nature and Norms*, N. Tognazzini and D. J. Coates, eds., (New York: Oxford University Press), pp. 27-48.
- Strawson, Galen. (1986). *Freedom and Belief*. (Oxford: Oxford University Press).
- Strawson, Peter. F. (1962). "Freedom and Resentment," *Proceedings of the British Academy* 48, pp. 187-211.
- Vargas, Manuel. (2007). "Revisionism" and "Response to Fischer, Kane, and Pereboom," In John Fischer, Robert Kane, Derk Pereboom, Manuel Vargas, *Four Views on Free Will*. (Oxford: Blackwell Publishers).
- Vargas, Manuel. (2013). *Building Better Beings*. (New York: Oxford University Press).
- Vilhauer, Benjamin. (2009). "Free Will Skepticism and Personhood as a Desert Base," *Canadian Journal of Philosophy* 39, pp. 489-511.
- Vilhauer, Benjamin. (2013). "Persons, Punishment, and Free Will Skepticism," *Philosophical Studies* 162: 143-63.
- Waller, Bruce. (1990). *Freedom without Responsibility*. (Philadelphia: Temple University Press).

Watson, Gary. (1996). "Two Faces of Responsibility," *Philosophical Topics* 24, pp. 205-20.

Wittgenstein, Ludwig. (1953). *Philosophical Investigations*. G.E.M. Anscombe and R. Rhees, eds.,  
G.E.M. Anscombe, trans. (Oxford: Blackwell).

Wolf, Susan. (2011). "Blame, Italian Style," in R. J. Wallace, R. Kumar, and S. Freeman, eds.,  
*Reasons and Recognition. Essays on the Philosophy of T. M. Scanlon*. (New York: Oxford  
University Press), pp. 332-48.