

A Forward-Looking Account of Self-Blame

Derk Pereboom, Cornell University

in *Self Blame*, Andreas Carlsson, ed., Cambridge: Cambridge University Press,
forthcoming.

Penultimate version

Introduction

Morality is shaped in part by our emotional attitudes, some of which introduce features of the practice of holding morally responsible that threaten to make demands that exceed our capacities, given how we are situated in the world of natural causes. These emotions include resentment and indignation, which, I argue, presuppose moral desert (Pereboom 2001, 2014; McKenna 2012). However, such desert is put at risk by the prospect of the causal determination of action and by the absence of control that indeterminism would introduce. In addition, the place of moral desert within general ethical frameworks such as consequentialism or Kantian universalizability theory is fraught. Here I continue the project of exploring the viability holding morally responsible absent desert, with a specific focus on self-blame and the related emotional attitude of regret (Pereboom 2014, 2017).

The conception of holding morally responsible and of blame in particular that I endorse is largely forward-looking. While on this view blame has a backward-looking element, since it is a matter of conceptual fact that appropriately blaming someone for an action requires that the action has already been performed and was in fact wrong,

blame's objectives are forward-looking, with objectives such as the moral formation of the wrongdoer and reconciliation with those who have been wronged. To blame is to adopt a stance of moral protest (Hieronymi 2011, Talbert 2012, Smith 2013), whose aim is in part to be communicated (McKenna 2012), while this aim may not in be realized in every case (Macnamara 2015, Chislenko 2019). Moral protest itself need not presuppose desert (Pereboom 2017), and emotions such as disappointment and sorrow, which also do not presuppose desert, may accompany blame as moral protest. Here I advocate a parallel account of self-blame. To blame oneself is to take on a stance of moral protest toward oneself in virtue of an action one regards as morally wrong. The reasons one has for doing so are forward-looking, and include one's moral formation and one's reconciliation in a relationship that has been impaired as a result of one's wrongdoing. Regret, distinguished from guilt, may accompany self-blame, and this attitude does not presuppose desert, or so I will argue.

Moral Responsibility and its Various Senses

Our practice of holding morally responsible is complex. It involves a number of different aims, and a range of responses justified by those aims. A number of theorists have argued that this complexity can be regimented, and that there is ultimately a single notion of moral responsibility that unifies the practice. Proponents of such a unitary view include R. J. Wallace (1994) and George Sher (2006). I believe that a view of this sort misrepresents the practice, 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.

Consider, specifically, the blaming part of the practice. One might argue that blaming essentially involves a supposition of morally deserved feeling of guilt, which it's blame's function to produce. Often blame does have the aim of inducing a feeling of guilt conceived as deserved due to the wrong done (e.g., Clarke 2013, Carlsson 2017, Duggan 2018). But a mother may blame her child just for the reason that it is her duty to see to his moral formation, specifically to moderate or extinguish the disposition to wrongdoing manifested in his action by presenting him with moral reasons to alter his behavior. What she does is to blame him for what he has done, but a supposition of a deserved feeling of guilt need not have a role in her calling him to account.

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. For instance, one might contend, perhaps despite apparent counterindications, that each sense of blameworthiness features, at its core, the supposition that the wrongdoer deserves to be blamed (McKenna 2012), or to feel guilty (Clarke 2013; Carlsson 2017; Duggan 2018), 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 are legitimate and remain in

place. This proposal requires distinct senses. Reasons for holding that some are best eliminated are twofold. The first is that they can successfully be criticized on moral grounds, and the second is that they can be challenged because they presuppose a sort of freedom we do not have.

It's typically agreed that one aspect of the practice of holding morally responsible indeed features the notion of desert. In the basic form of desert, agents deserve to be blamed punished just because they have knowingly acted wrongly, and agents deserves credit or praise just because they have knowingly acted rightly (Feinberg 1970; Pereboom 2001, 2014; Scanlon 2013). Here is a more formal characterization of basic desert:

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)

This characterization can be extended to include basically deserved punishment and reward.

There may in addition be senses of moral responsibility that involve a non-basic variety of desert. Essentially forward-looking notions of holding agents to be deserving

of blame and punishment have been defended on consequentialist grounds (Dennett 1984, 2003; Dennett and Caruso 2021; Vargas 2013) or for contractualist reasons (Lenman 2006; Vilhauer 2009). On one such account, our practice of holding agents morally responsible in a desert-involving sense should be retained because doing so would have the best overall consequences relative to alternative practices. Daniel Dennett (1984, 2003; Dennett and Caruso 2021) advocates a version of this position, as does Manuel Vargas (2007, 2013). Such options must be considered seriously, but the proposal I envision involves a more resolutely forward-looking approach on the part of practitioners.

There are reasons to be skeptical of any notion of moral responsibility that involves basic desert. One concern 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; Caruso 2021). As just noted, one might argue that some desert sense of moral responsibility can or should be retained because doing so stands to bring about good consequences, but such desert is not basic. Another concern is that for a number of contending general normative ethical theories the notion of desert seems to have the role of an awkward supplement (Pereboom 2014). A place for desert in typical consequentialist views is arguably uncomfortable, and despite Kant's well-known invocation of desert in justifying criminal punishment (Kant 1797/1963), that appeal is not plausibly justified by any formulation of the Categorical Imperative, which he held to be the supreme and

comprehensive moral principle. An additional issue is that the conception of deserved pain or harm that is imposed in blaming or punishing, would at least in its basic form seem to involve the idea of harm as a non-instrumental good (McKenna 2019), an idea that might well be contested. Motivated by these considerations, I've proposed a view that rejects desert-involving senses of moral responsibility altogether.

Moral Responsibility without Desert

The notion of moral responsibility, and blame in particular, that I develop and endorse (Pereboom 2014, 2017) is largely forward-looking. Blaming is, in its paradigm cases, a kind of calling to account, and is justified by forward-looking elements, including the following:

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.
4. The retention of integrity of victims of wrongdoing.

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. Because we value morally good character and action that results from it, we have a stake in the moral formation of character when it is beset by dispositions to

misconduct. For those whose sense of integrity has been undermined by having been victims of wrongdoing, blaming can be instrumental to restoring that integrity.

There is an account of praise that is parallel to this conception of blame. Of the forward-looking aims just cited, the one most clearly amenable to praise is moral formation. We may praise an agent for a morally exemplary action to strengthen the disposition that produced it. Praise can also have a protective function, since strengthening dispositions to act rightly stands to have the effect of reducing the incidence of harmful behavior. Corresponding to reconciliation is the notion of celebrating success in a relationship, and praising may have this objective as well.

Michael McKenna has proposed a conversational account of moral responsibility that, with a few revisions, is amenable to the forward-looking view I advocate (McKenna 2012, 2019; Fricker 2016). The actions of a morally responsible agent are potential bearers of a type of meaning by virtue of indicating the quality of will that resulted in the action (2012, 92-94; see also Arpaly 2006). Blaming an agent who manifests an immoral quality of will in acting expresses an attitude such as resentment or indignation, and its aim 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 they address actions that are morally charged, they understand themselves to be initiating a meaningful interaction in such a conversational exchange. McKenna labels this initial stage of the conversation *moral contribution*. In the case of an ostensibly immoral action, in the second stage the agent is blamed by an interlocutor; McKenna calls this stage *moral*

address. In the third stage, *moral account*, the blamed agent offers an excuse, a justification, or an apology. The interlocutor might at this point carry on the conversation by forgiving or punishing the wrongdoer. In a subsequent stage of the interaction the blamed agent may be restored by other participants 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. At this point he invokes a paradigm-similarity model for the meaning and extension of a concept (Rosch 1972, 1973). A conversation with a living and present participant as is a paradigm case of blaming, while blaming the dead, for instance, qualifies as a case of blaming in virtue of its similarity to paradigm cases.

I can and do endorse such an account, on the supposition that the objectives of blaming are forward-looking goals such as protection, reconciliation, moral formation, and retention of integrity, and that deserved blame excised (Pereboom 2014, 2017). As in McKenna's account, on the forward-looking view blame may be painful for the wrongdoer who is blamed. It may be painful to be called out for having done wrong, and the pain of regret (to be discussed), may result. But by contrast with blame conceived as involving basically deserved pain, the pain on the forward-looking conception is not conceived as a non-instrumental good imposed by blaming, but rather as an instrumental good that serves the forward-looking goals to be achieved by blaming. This rejection of this pain as a non-instrumental good is significant, and lies at the core of what motivates the forward-looking account.

On the forward-looking conversational model, as in McKenna's, it's the agent's responsiveness to reasons that is engaged in the envisioned process for both blame and praise. For blame, at the stage of moral address, which in the case of wrongdoing is the blaming stage, one may request an explanation with the intent of having the agent acknowledge a disposition to act wrongly, and then, if he has in fact acted wrongly and is without excuse, one may intend for him to come to see that the disposition issuing in the action is best modified or extinguished. In the paradigm case, such dispositional change is effected by way of the agent's recognition of moral reasons to make it. More generally, it is an agent's responsiveness to reasons, together with forward-looking objectives, that explains why he is an appropriate recipient of blame in this forward-looking conversational sense.

Reasons-responsiveness is often advanced as the key necessary condition for basic desert responsibility by philosophers who maintain that this sort of responsibility is compatible with the action's causal determination by factors beyond the agent's control, and that it isn't explained by her ability to do otherwise (Fischer 2007, 82; McKenna 2012, Sartorio 2016). The largely forward-looking conception of moral responsibility I advocate is also compatible with an agent's being causally determined to act by factors beyond her control, and in an appropriately constructed deterministic manipulation argument the manipulated agent will be morally responsible in this way (Pereboom 2014; Pereboom and McKenna 2019). In my view, agents can also be morally responsible in the forward-looking sense in Frankfurt examples (Frankfurt 1969, Pereboom 2001, 2014), and thus, just as in the compatibilist position that John Fischer

(1994, 2007), Michael McKenna (2012) and Carolina Sartorio (2016) develop, the kind of freedom or control required for basic desert moral responsibility for an action will be a matter of its actual causal history. The compatibilists just mentioned are all committed to responsiveness to reasons as the central condition on basic-desert or reactive-attitudes-involving moral responsibility, while I view it as the most significant condition for a notion of responsibility that instead focuses on goals such as protection, reconciliation, moral formation, and retention of integrity.

Against anger as the core blaming attitude

Pamela Hieronymi (2001), Matt Talbert (2012), Angela Smith (2013) have proposed that blame should be understood as moral protest, and I endorse a position of this kind (Pereboom 2017). Mine differs in that for these other proponents the negative reactive attitudes of resentment and indignation are central to blame, while in my view they are not. As I see it, moral protest is a psychological stance, a posture of mind that has certain aims that are manifest as dispositions to act (Schwitzgebel 2013). It is a stance of opposition to specific wrongful actions and their general type, whose aims include moral engagement with the wrongdoer by communication of this opposition to such wrongdoing together with reasons to refrain from it. Hieronymi (2001) connects moral protest to the negative reactive attitudes; on her account moral protest in fact is a reactive attitude such as resentment, which is commonly understood as a second-person attitude I have toward someone who has wronged me (Darwall 2002; Shabo 2012). By extension, one might include indignation, understood as a third-person

attitude one has toward someone who has wronged someone else. I deny this connection.

David Brink and Dana Nelkin (2013) view blame as having a “core and syndrome structure,” and they argue that there is a core to blame present in all cases of blame. So far I agree. On their view, that core is an aversive attitude toward the target that is predicated on the belief or judgment that the target is blameworthy. The syndrome features the disposition to manifest this aversive attitude in various ways. I contend, by contrast that the core is the stance of moral protest, which does not essentially involve resentment or indignation, both forms of moral anger (Pereboom 2001, 208-13; D’Arms and Jacobson 2003).

David Shoemaker (2017) maintains that moral anger is essential to blame in the accountability sense (more on this later), and he develops and defends a Strawsonian response-dependence account of such blame in which the designated response is anger:

Fitting Response-Dependence about the Blameworthy: The blameworthy (in the realm of accountability) just is whatever merits anger (the angerworthy); that is, someone is blameworthy (and so accountable) for X if and only if, and in virtue of the fact that, she merits anger for X.

Shoemaker argues that what unifies all of the properties that make anger an appropriate response to wrongdoing is just that it merits anger, and this is what makes the account truly a response-dependent one. Here I don’t wish to focus on the response-dependence feature of the account, but rather on the choice of anger as the response. As Shoemaker points out, there is a response-*in*dependent account that also

features anger, but as a response which is independent of the property in which blameworthiness consists, while anger is made appropriate in virtue of that property:

Response-Independence about the Blameworthy: The blameworthy consists in a property (or properties) of agents that makes anger at them appropriate, a property (or properties) whose value-making is ultimately independent of our angry responses. Anger at someone for X is appropriate if and only if, and in virtue of the fact that, she is antecedently blameworthy (and so accountable) for X. What makes her blameworthy is thus ultimately response-independent.

On this account, blameworthiness is not essentially dependent on the response of anger. But moral anger, in the form of reactive attitudes such as resentment and indignation, is the property that fixes the referent of the term 'blameworthy.' The blameworthy consists in properties that in fact actually merit such anger, even if blameworthiness doesn't consist in just whatever merits anger.

I have raised two concerns about the choice of anger in each of these accounts (Pereboom 2020; cf., Nussbaum 2016). First, there are cases of blameworthiness that are plausibly not cases of angerworthiness. Melanie is a high school teacher, whose students respect her but misbehave in ways typical for this age group in that context. They come unprepared not having done the assigned reading, or talk about non-class-related matters in distracting ways, or surf the internet instead of participating and paying attention. However, Melanie responds with firm protest but not with anger, understanding the level of moral and social development for late adolescents. Here the angry response stands to be counterproductive and to undermine her effectiveness and

the respect students have for her. Or take Nicole, a parent, whose teenage children misbehave in typical ways; they squabble, text their friends when they should be sleeping, and fail to expedite household chores. Nicole responds from the sense of a duty to correct and educate, combined with care, but not with anger. Here an angry response would tend to engender resistance, and is generally less effective for moral education than the approach Nicole adopts. For such cases of teachers and parents, evidence for an angry response actually being inappropriate is that those who become angry are routinely criticized and are not respected to the degree enjoyed by those who forgo anger.

The second concern is that anger has a strong tendency to distort judgments of blameworthiness, and that it's questionable whether being blameworthy is being worthy of a reactive attitude that systematically distorts judgments of blameworthiness. Surveys conducted by Mark Alicke and his associates indicate that subjects who spontaneously evaluate agents' behavior unfavorably are apt to exaggerate their causal control and any evidence that might favor it while deemphasizing counterevidence (Alicke, Davis, and Pezzo 1994; Alicke 2000; Alicke, Rose, and Bloom 2012). Alicke calls this tendency *blame validation*. In addition, experimental evidence that blaming behavior is subject to problems of these kinds has been mounting (e.g., Nadelhoffer 2006).

There is reason to believe that the anger that accompanies blame is what leads to these problems (Duggan, ms). Psychological research indicates that anger, once activated, degrades subsequent reasoning processes in various ways (e.g., Lerner, et.al.

1998, Goldberg et.al. 1999, Litvak et.al. 2010). Anger increases tendencies to overlook mitigating features of the circumstances before blaming, to perceive ambiguous behavior as hostile, to rely on stereotypes, concerning, for example, ethnicity in blaming, and to discount the role of uncontrollable factors when attributing causality. Anger makes subjects slower to associate positive traits than negative traits with an out-group. Julie Goldberg and her associates find in one of their studies that when the retributive desire to harm is not satisfied, anger "activate[s] an indiscriminate tendency to punish others in unrelated situations without regard for whether their actions were intentional" (Goldberg *et. al.*, 1999)

This second concern may be less decisive than the first, since even if anger has the distorting propensities cited, it may yet be the best candidate for a general emotional attitude to accompany blame. However, I contend that moral protest, as a psychological stance, is a better fit. This is especially true in virtue of the first problem, since, as the examples indicate, in relationships between teachers and students, and parents and children, for common sort of wrongdoing anger is actually inappropriate.

One might respond by arguing that these are not paradigm cases of blame-involving relationships, since they are not relationships of mutual regard, but rather relationships in which there is a relevant discrepancy in maturity. But consider relationships between faculty members at a university, or relationships between administrators and faculty, which are relationships of mutual regard. Non-major but yet significant wrongdoing in such relationships is not infrequent and to be expected. For example, university faculty are partial to their close colleagues and political allies when

it comes to appointments and honors, and often the resulting advocacy is wrong. Suppose Olivia is a university administrator who often faces these sorts of issues with faculty in her purview. Imagine that she responds not with anger, but calmly with arguments that invoke the rules that best govern the situations at issue. In such cases, angry responses typically reduce an administrator's effectiveness, and tend to cause false judgments which in turn motivate defective solutions. Here too, anger is arguably an inappropriate response.

A moral protest account of blame

These observations call for a general characterization of blame and blameworthiness in which anger is not the core attitude. I've proposed a conception of blame as moral protest that does not essentially involve moral anger as the designated attitude. When Melanie, Nicole, and Olivia morally protest the wrongful behavior at issue, they are morally concerned, but not resentful or indignant. Again, I conceive moral protest as a psychological stance, one that features a disposition to engage in overt protest against an agent for having performed an action that the protester perceives to be morally wrong. I've endorsed the following simple version of a moral protest view of blame, in which moral anger does not appear (Pereboom 2017):

Moral Protest Account of Blame: For B to blame A is for B to adopt a stance of moral protest against A for immoral conduct that B attributes (however accurately) to A.

The immoral conduct will typically be a wrongful 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, Markovits 2010), and worthy of protest. Sometimes blame is misplaced, since no immoral actions have been performed, but the protest can still count as blame. This may happen when B believes A to have acted wrongly but the belief is false, perhaps due to misinformation or improper consideration of evidence. This can also happen when B does not believe that A acted wrongly but nonetheless represents A as having acted wrongly, as in cases of false accusation motivated by anger, envy, or fear. It's often the case that blame has the goal, as in Hieronymi's (2001) proposal, of moral protest against an agent for a past action that persists as a present threat, and this is one highly important objective for blame. But not all blame has this point, as when we blame the dead, or blame an agent who is alive but lacks a persisting disposition to act badly -- someone, for instance, who has already undergone moral reform. In such cases protest may yet have the aim of explicitly noting immoral behavior in order to encourage moral improvement on the part of an audience. In the example of the already-reformed wrongdoer, blame might still be intended as a step in the process of reconciliation. Or the aim may be that the victim of wrongdoing reassert and retain her integrity.

An objection to the protest account of blame is that while unexpressed blame is possible, the idea of unexpressed protest is incoherent, and thus blame cannot be identified with moral protest. The objection is that protest is essentially communicative,

and unexpressed protest is not communicative. Eugene Chislenko (2019) has recently provided a reply to this concern, citing the distinction Coleen Macnamara draws between the activity of communicating—of which mental states kept private are not instances—and the idea of a communicative entity (Macnamara 2015b: 217). An unsent e-mail, even though it does not actually perform the function of communicating, nevertheless has the function of evoking uptake of representational content in a recipient (Macnamara 2015a: 548). An unsent e-mail is thus communicative in nature; and similarly, unexpressed protest is communicative in nature. For a salient case in point, an unsent email might be an unexpressed message of protest. Chislenko writes: “We can even say, as [Angela] Smith does of blame, that the email “expresses protest, and... seeks some kind of moral reply” (2013: 39), even when the email is unsent.” (Chislenko 2019). We can add that the email can express moral protest even if its author never intends to send it. Similarly, someone who privately blames may never intend to communicate it, even if that blame has the function of moral protest.

Thus on my proposal, moral protest is fundamentally a psychological stance one takes, one that is apt for being communicated, and has that aim, but that aim may not be implemented in specific instances. An entity having an aim or function that it does not actually implement in certain instances is familiar from biology – for example, a heart that fails to pump blood – and should not be regarded as unusual.

In accord with the protest account of blame, I would propose the following amended version of Shoemaker’s response-dependent view about the blameworthy:

Fitting Response-Dependence about the Blameworthy: The blameworthy (in the realm of accountability) just is whatever merits moral protest (the protestworthy); that is, someone is blameworthy (and so accountable) for X if and only if, and in virtue of the fact that, she merits moral protest for X.

This account has a response-independent correlate, which also invokes moral protestworthiness but claims that there are properties that make wrongdoing protestworthy that are independent of our protest responses, while appropriate moral protest can serve to fix properties which 'blameworthiness' picks out:

Fitting Protest-Response-Independence about the Blameworthy: The blameworthy (in the realm of accountability) consists in a property (or properties) of agents that makes morally protesting their wrongdoing appropriate, a property (or properties) whose value-making is ultimately independent of our responses of moral protest.

Following Shoemaker, in these formulations I've retained the idea that the notion to be characterized is blameworthiness in the realm of *accountability*. He provides the following characterization of accountability he provides:

To be accountable for something is to be liable to being appropriately *held* to account for it, which is to be eligible for a range of fitting responsibility responses with a built-in confrontational element. (2015, 87)

Moral protest, as I conceive it, is essentially confrontational, at least to some degree. But being confrontational is compatible with not involving anger. One's attitude toward the wrongdoer might be exclusively compassionate, while believing that in this case

compassion requires confrontation. There are other characterizations of accountability on which blaming someone in this sense essentially involves the supposition that the wrongdoer deserves or basically deserves to be the target of blame. Given that notion of accountability, I reject it, and would want to frame the discussion in terms of a different notion of blame.

On the proposed account of blame, blame does not essentially involve anger, while it does essentially involve taking on the stance of moral protest. But despite moral anger involving a presupposition of desert, and indeed basic desert, which I claim is a false presupposition, angry blame may still be *practically* rational for the desert-denier. It may be, for instance, that in certain cases opposition to violence and abuse is most effective if it is motivated partly by moral anger. Then the putative epistemic irrationality of having the false supposition may be overridden. In this sort of case, moral progress needn't be held hostage to epistemic rationality.

Self-Blame and Regret

Just as one might take on a stance of moral protest against another for his having done wrong, one might also adopt such a stance in response to one's own immoral behavior. One might, by virtue of one's general moral commitment, view an action one has performed as wrong and the disposition that issues in it as morally defective, and as a result take on a stance of opposition against one's action and the disposition. In adopting this stance, one may, for instance, aim at one's own moral formation or with reconciliation with someone one has wronged. If one has mocked and

embarrassed someone, and the relationship with him has been impaired as a result, one might assume a stance of protest against that action and one's disposition to act badly in this way. One might do so in order to extinguish this disposition and to reconcile with the person one has wronged.

Which emotional attitudes aptly accompany self-directed moral protest? Just as in the case of one's child having done wrong, one might feel disappointment or sorrow, without desert being invoked. But 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 it features (cf., Carlsson 2017, Duggan 2018). Clarke proposes, first, that there is value in the recognition by an agent who is blameworthy that he is blameworthy, and a further response, the feeling of guilt, provides an 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.

Let's adopt the convention that 'guilt' refers to an attitude that presupposes basic desert, that is, one's basically deserving to feel pain accompanying the recognition that one has done wrong, and that 'regret' refers to a similar attitude, which also involves feeling pain that accompanies the recognition that one has done wrong, but which does not presuppose that this pain is basically deserved. Given this conception, I can agree that regret is a morally fitting additional reaction to one's own wrongdoing without committing to desert.

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 several additional analogies. One might appropriately feel pained that one failed to meet one's own standards for playing chess when one understands that one's substandard performance is due to factors beyond one's control, while this pain is not deserved. A similar example is due to Shoemaker (forthcoming), one in which a baseball player feels pain upon making a mistake, yet he doesn't deserve to feel this pain. A different kind of example is due to McKenna: it would be appropriate, and basically so, to feel the pain of grief upon the death of a loved one, while this pain is not deserved (McKenna, 2012, 2019). Saliiently, these cases feature the appropriateness of feeling pain without its being deserved.

In accord with these analogies, 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. How can the feeling of pain upon recognition of wrongdoing be appropriate – and even basically so – but not be basically deserved? What's required is an indicator for distinguishing basically deserved pain from merely basically appropriate pain. Both Andreas Carlsson (2017) and I (Pereboom 2017) have suggested that the pain of guilt, given its desert presupposition, would be (prima facie) appropriately *imposed* while the pain of grief is not. Let me try to make the suggestion more precise. If pain is basically deserved on account of wrongdoing, this gives rise to a prima facie moral permission for appropriately situated agents to intentionally impose it

on the wrongdoer for a non-instrumental reason and thus for its own sake. If a wrongdoer basically deserves to be punished, there is then a prima facie presumption that the right sorts of authorities, perhaps parents or state officials, are permitted to intentionally impose it on him for a non-instrumental reason and thus for its own sake. Guilt, given that it presupposes basic desert, would then involve pain that one regards as appropriately intentionally imposed by oneself or by appropriately situated others for its own sake. The pain of grief lacks this feature. Despite grief and the pain that it involves being basically appropriate for those who have experienced loss of a loved one, no one is permitted to intentionally impose the pain of grief in such circumstances for its own sake. One may inform the bereaved that she has undergone the loss of a loved one, as a result of which it's evident that she will feel the pain of grief, but this is not a case of intentionally imposing the pain of grief for its own sake. The pain of grief is never basically deserved despite at times being basically appropriate, and so here we have a potential indicator for distinguishing basically deserved from non-basically-deserved appropriate pain.

In the passage from Hilary Bok, she makes two claims regarding her analogy of the pain of regret to the pain of heartbreak. The first is that the pain of regret is not a pain that one inflicts on oneself as a punishment. The second is that it is not a pain that we undergo because we deserve it. This suggests that she conceives of the claims as linked: the pain of heartbreak's not being appropriately imposed is connected with its not being deserved. My related proposal is that it is impermissible to intentionally impose the pain of regret non-instrumentally and for its own sake, and this is what

distinguishes it from the putative basically deserved pain of guilt (cf. McKenna 2012, 2019). This allows that one may issue a moral protest against a wrongdoer for the reason that it stands to result in moral formation, foreseeing that it will result in the pain of regret. Here it is not the case that pain is imposed for a non-instrumental reason. It may be intuitive that in certain circumstances it is permissible to intentionally impose the pain connected with recognition of wrongdoing on others and on oneself for its own sake. But the retraction of basic desert has its costs, and this is one of them.

Dana Nelkin (2020) provides a thought experiment that in her view supports conclusion that the pain of guilt/regret is not appropriately imposed for its own sake:

Imagine that you have a special power (call it “The Look”). By looking at another person in the right way, you can bring about feelings of guilt. The other person culpably wrongs another—it is not a trivial offense, but neither is it the worst possible. Imagine that she betrays the confidence of a friend and as a result the friend has a bad day. You now have the chance, by looking at the offender in that way you have mastered so well, to bring about guilt feelings in her. It would be the easiest thing, requiring nothing in the way of effort or sacrifice. But now also imagine that there is no further good to come from your exercising this power you have. The offender has already resolved not to do the same sort of thing again, no one else is around to experience the results, the relationship is either already irreparably damaged no matter what, or all has been forgiven, so that there is no benefit to the relationship to be had, and so on. Would you be making a mistake, or leaving a reason on the table, so to speak, by taking a pass

on inducing this painful feeling? Would there be a (non-instrumental) moral good that would have been costless to achieve that you failed to promote? I do not have the intuition here that you would be making this kind of mistake. That suggests to me that there is not a pro tanto reason to induce guilt that stems from blameworthy action alone.

I agree with Nelkin; the thought experiment provides intuitive reason, independent of the arguments for free will skepticism and the general concerns about basic desert, to deny the value of what I'm calling guilt. But it does not yield a challenge to regret.¹

There are other justifications for regret on the supposition that it doesn't presuppose basic desert. Ben Vilhauer (2008) 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 motivate repentance and moral reform, for reconciliation with those one has wronged, and restoration of 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.

¹ Here see McKenna's (2019) discussion according to which the pain of guilt and the value of guilt as an expression of moral concern make up an organic unity whose whole has a value that is not analytically decomposable, and Nelkin's (forthcoming) response to McKenna.

Final words

I've proposed a conception on which to blame others is to take on a stance of moral protest toward them in virtue of an action one regards as morally wrong. The reasons for doing so are forward-looking, such as the wrongdoer's moral formation or reconciliation in a relationship that has been impaired as a result of the wrongdoing. By extension, to blame oneself is to take on a stance of moral protest toward oneself in virtue of an action one has performed and regards as morally wrong, and to do so for similar forward-looking reasons. Regret, a painful response to one's own wrongdoing that does not presuppose basic desert, may appropriately accompany self-blame. Guilt and regret can both be classified as basically appropriate pained responses to one's own wrongdoing. But the pain of guilt, and not that of regret, counts as basically deserved because it is prima facie permissible for those who are appropriately situated to intentionally impose it on a wrongdoer for non-instrumental reasons.²

² For excellent comments and discussion I'm grateful to the participants at a workshop at the University of Oslo in September 2019, organized by Andreas Carlsson. Special thanks are due to Dana Nelkin for valuable comments on several drafts.

References

- Alicke, Mark D. (2000). "Culpable Control and the Psychology of Blame, *Psychology Bulletin* 126, pp. 556-74.
- Alicke, Mark D., T. L. Davis, and M. V. Pezzo. (1994). "A Posteriori Adjustment of A Priori Decision Criteria," *Social Cognition* 8, pp. 286-305.
- Alicke, Mark D. Rose, and D. Bloom. (2012). "Causation, Norm Violation and Culpable Control," *Journal of Philosophy* 106, pp. 587-612.
- Bok, Hilary. (1998). *Freedom and Responsibility*, Princeton: Princeton University Press.
- Brink, David and Dana Nelkin. (2013). "Fairness and the Architecture of Responsibility," *Oxford Studies in Agency and Responsibility*, v. 1., David Shoemaker, ed., pp. 31-54.
- Carlsson, Andreas Brekke. (2017). "Blameworthiness as Deserved Guilt." *The Journal of Ethics* 21, pp. 89-115.
- Caruso, Gregg D. (2021). *Rejecting Retributivism: Free Will, Punishment, and Criminal Justice*, Cambridge: Cambridge University Press.

Chislenko, Eugene. (2019). "Blame and Protest," *The Journal of Ethics* 23 (2), pp. 163-81.

Clarke, Randolph. (2013). "Some Theses on Desert," *Philosophical Explorations* 16, pp. 153-64.

D'Arms, Justin, and Daniel Jacobson. (2003). "The Significance of Recalcitrant Emotion (or, Antiquasijudgmentalism)," *Royal Institute of Philosophy Supplements* 52, pp. 127-145.

Dennett, Daniel. (1984). *Elbow Room*, Cambridge: MIT Press.

Dennett, Daniel. (2003). *Freedom Evolves*, New York: Viking.

Dennett, Daniel C., and Gregg D. Caruso. (2020). *Just Deserts*, Cambridge: Polity Press.

Duggan, Austin P. (2018). "Moral Responsibility as Guiltworthiness," *Ethical Theory and Moral Practice* 21, pp. 291-309.

Duggan, Austin P. (ms.). "A Genealogy of Retributive Intuitions."

Feinberg, Joel. (1970). *Doing and Deserving*, Princeton: Princeton University Press.

Fischer, John. M., and Mark Ravizza. (1998). *Responsibility and Control*, Cambridge: Cambridge University Press.

Goldberg, Julie, Lerner, J., Tetlock, P. (1999). "Rage and Reason: The Psychology of the Intuitive Prosecutor," *European Journal of Social Psychology* 29, pp. 781-95.

Haji, Ishtiyaque. (1998). *Moral Accountability*, New York: Oxford University Press.

Hanser, Matthew. (2005). "Permissibility and Practical Inference," *Ethics* 115, pp. 443-70.

Hieronymi, Pamela. (2001). "Articulating an Uncompromising Forgiveness," *Philosophy and Phenomenological Research* 62, pp. 529-54.

Kant, Immanuel. (1797/1963). *The Metaphysical Elements of Justice*, John Ladd, tr., New York: Bobbs-Merrill.

Lenman, James. (2006). "Compatibilism and Contractualism: The Possibility of Moral Responsibility," *Ethics* 117, pp. 7-31.

Lerner, Jennifer S., Julie H. Goldberg, and Philip E. Tetlock. (1998). "Sober Second Thought: The Effects of Accountability, Anger, and Authoritarianism on

Attributions of Responsibility.” *Personality and Social Psychology Bulletin* 24 (6), pp. 563–74.

Litvak, Paul M., Jennifer S. Lerner, Larissa Z. Tiedens, and Katherine Shonk. (2010). “Fuel in the Fire: How Anger Impacts Judgment and Decision-Making.” In *International Handbook of Anger*, edited by Michael Potegal, Gerhard Stemmler, and Charles Spielberger, New York: Springer, pp. 287–310.

Macnamara, Coleen. (2015a). “Reactive Attitudes as Communicative Entities,” *Philosophy and Phenomenological Research* 90, pp. 546–569.

Macnamara, Coleen. (2015b). “Blame, Communication, and Morally Responsible Agency,” in *The Nature of moral Responsibility: New Essays*, Randolph Clarke, Michael McKenna, and Angela M. Smith, eds., New York: Oxford University Press. pp. 211–35.

Markovits, Julia. (2010). “Acting for the Right Reasons,” *The Philosophical Review* 119 (2), pp. 201-42.

McKenna, Michael. (2012). *Conversation and Responsibility*, New York: Oxford University Press.

McKenna, Michael. (2019). "Basically Deserved Blame and its Value," *Journal of Ethics and Social Philosophy* 15, pp. 255-82.

Moore, Michael. (1998). *Placing Blame*. Oxford: Oxford University Press.

Nelkin, Dana. (2011). *Making Sense of Moral Responsibility*, Oxford: Oxford University Press.

Nelkin, Dana K. (2020). "Guilt, Grief, and the Good," *Journal of Ethics and Social Philosophy* 36 (1), pp. 173-91.

Nussbaum, Martha. (2016). *Anger and Forgiveness*, Oxford: Oxford University Press.

Pereboom, Derk. (2001). *Living without Free Will*, Cambridge: Cambridge University Press.

Pereboom, Derk. (2014). *Free Will, Agency, and Meaning in Life*, Oxford: Oxford University Press.

Pereboom, Derk. (2017). "Responsibility, Regret, and Protest," *Oxford Studies in Agency and Responsibility*, v. 4., David Shoemaker, ed., pp. 121-140.

Pereboom, Derk. (2020). "Forgiveness as Renunciation of Moral Protest," in *Forgiveness*, Michael McKenna, Dana Kay Nelkin, and Brandon Warmke, eds., New York: Oxford University Press.

Pereboom, Derk, and Michael McKenna. (2021). "Manipulation Arguments," in *The Oxford Handbook of Moral Responsibility*, Dana Nelkin and Derk Pereboom, eds., New York: Oxford University Press, forthcoming.

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. (2013). "Giving Desert Its Due," *Philosophical Explorations* 16, pp. 101-16.

Schwitzgebel, Eric. (2013). "A Dispositional Approach to Attitudes: Thinking Outside the Belief Box" in Nikolaj Nettleman, ed., *New Essays on Belief: Constitution, Content and Structure*, Palgrave-Macmillan, pp. 75-99.

- Shabo, Seth. (2012). "Where Love and Resentment Meet: Strawson's Interpersonal Defense of Compatibilism," *The Philosophical Review* 121, pp. 95-124.
- 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.
- Shoemaker, David. (2017). "Response-Dependent Responsibility," *The Philosophical Review* 126, 481-527.
- 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.

Talbert, Matthew. 2012. "Moral Competence, Moral Blame, and Protest," *Journal of Ethics* 16, 89-101.

Vargas, Manuel. (2013). *Building Better Beings*, New York: Oxford University Press.

Vilhauer, Benjamin. (2004). "Hard Determinism, Remorse, and Virtue Ethics," *Southern Journal of Philosophy* 42, pp. 547-64.

Waller, Bruce. (1990). *Freedom without Responsibility*, Philadelphia: Temple University Press.

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