What is blame’s nature, and what are its norms? In what follows, I shall answer these questions by reference to a conversational theory of moral responsibility. According to this theory, blame—at least a distinctive kind of blame—can be accounted for in terms of its communicative role in an interpersonal exchange between members of a moral community. The view I defend is broadly Strawsonian insofar as a restricted class of emotions plays a pivotal role. Ultimately I shall offer an account of blame’s norms that makes room for the familiar thought that blame is deserved by one who is blameworthy.

1. The Challenge of Theorizing about Blame

Despite the pervasiveness of the phenomenon in ordinary life, blame is an elusive notion. It is maddeningly hard to nail down a theory that gets the extension even close to right. This is shown by the diversity of strikingly different views about its nature. On some views, to blame is just to engage in a kind of punishment (Smart 1961). On others, to blame is most fundamentally to register a criticism of an agent’s free conduct (Zimmerman 1988) or, instead, of how she exercised her judgment-sensitive capacities (Smith 2005). On yet others, it is a matter of altering one’s relationship with the blamed party in light of that party’s wrongful impairment to the relationship (Scanlon 2008). Another option has it that to blame is to believe that a person has acted badly or has a bad character and to desire that this not be so (Sher 2006). On a similar proposal, to blame is to perceive that a person acted wrongly and with ill will, and to do
so from the standpoint of caring about morality (Arpaly 2006). Then there is
the view that blaming is a matter of reacting to the one blamed with a perti-
nently charged emotion, such as resentment or indignation (Strawson 1962).
This list is far from exhaustive.

The task of accounting for blame is made all the harder given that there are
different notions of responsibility. It’s not just that there are different theories
of what responsibility is; it’s that there are different kinds of responsibility to
theorize about and a multiplicity of theories directed at these different kinds.
Blame, as one member in the family of responsibility concepts, is caught up in
this mix. So it is especially challenging to get focused on the proper range of
phenomena and the attendant concept(s) of blame.

There is, furthermore, a wide variability to the permissible modes of mani-
festing blame. This complicates matters even more. You and I might both be
entitled to blame a colleague for some bit of scandalous conduct, and yet we
might fittingly do so in ways that are wildly divergent. How, then, might one
get an account of blame up and running?

I shall proceed by focusing on prototypical cases of blame, which I shall
characterize in terms of directed blame. I’ll seek a means of explaining just
these and will forgo any attempt at an exhaustive account of blame. Neverthe-
less, the central cases of blame I will focus on are, I think, representative of
what many have in mind when reflecting on the nature and norms of blame.
My hope is that other cases of blame—cases within, so to speak, the extension
but closer to the periphery of the concept—can be explained at least in part by
reference to my treatment of the prototypical cases. I’ll begin by calling atten-
tion to several points on a conceptual map.

2. Some Preliminary Considerations regarding the
Nature and Norms of Blame

Consider first responsibility. I am interested in blame only as it bears on ques-
tions of moral responsibility, as opposed to, say, legal responsibility. As for moral
responsibility, many philosophers have argued that there is more than one sort
(e.g., Haji 1998; Fischer and Tognazzini 2011; Scanlon 1998; Shoemaker 2011;
Watson 1996; Zimmerman 1988). Here, I wish to focus just on moral responsi-
bility in the accountability sense. Moral responsibility in this sense involves the
possibility of holding an agent to account for her conduct, and thereby seeing
her as properly responsive to our demands and sanctions. Such an agent is one
who can be expected to acknowledge and comply with others’ moral expecta-
tions. When she falls short, she is taken to be justly liable to burdensome modes
of response from those who are warranted in holding her to account.

Now consider blame. Moral blame as it bears on accountability involves a
negative evaluation of, for instance, an agent’s action in a manner that is in
some way linked to the appropriateness of holding her to account for so acting. The natural thought, when focusing on this sort of blame, is to attend to cases in which a person is held directly to account by another who openly blames her as a direct means of making moral demands, expressing expectations, reprimanding, or something similar. But care needs to be taken at this point, since accountability-blame is not limited to such cases. We often blame others—in the pertinent sense—in their absence. Also, we sometimes conceal our blame from everyone.

It is useful to distinguish between private blame, overt blame, and directed blame. Private blame involves adopting a blaming attitude toward someone but concealing the outward behavioral manifestations. Overt blame involves adopting such an attitude and making it manifest in one's conduct. This can be done in the absence of the blamed; obvious examples involve blaming the dead. Directed blame is a form of overt blame manifested in the presence of the blamed party. Indeed it is outwardly directed at the blamed party.\(^1\)

Instances of directed blame are relatively rare by comparison with instances of private blame and overt blame that is not directed. Or, at any rate, that is my impression. Nevertheless I wish to focus just on directed blame. In my estimation, it's the more fundamental notion. Hence, despite the relative rarity of directed blame, I believe it is best to treat such cases as prototypical and then seek to explain other cases, such as cases of private blame, by reference to our understanding of the more fundamental cases.

Thus far I have engaged in preliminary conceptual spadework in the service of getting clear on the question of blame's nature. But what of its norms? Here we can distinguish three questions. The first concerns the normative warrant for an agent's being to blame. What justifies its being the case that an agent ought to be blamed? The second concerns the standing or license one has who is rightly positioned to blame the party who is blameworthy. What makes it permissible (or obligatory?) for this person or that, but not some other, to do the blaming? While the first of these two questions naturally points to the agent who is to blame, the second points to the one who does the blaming. A third question concerns the normative force of blame—the sting it putatively ought to have when directed at one who is blameworthy. Why and how ought a blamed person regard as burdensome the blame directed at her? In my estimation, this question naturally points at the (potential) relationship between the blamer and the blamed. The first of these questions we can call the question of normative warrant; the second we can call the question of normative

\(^1\) Note that directed blame, in the technical sense I have assigned to it here, is second-personal. So it is most fitting to focus on cases in which one blames another. Self-blame seems ill-suited for this model. Despite this appearance, I do not think it is. As I understand it, self-blame, like other forms of blame that are not directed, is to be accounted for by reference to cases of one directly blaming another. In this essay I'll not attend to self-blame. But see chapter 3, section 3 of my 2012.
standing; and the third we can call the question of normative force. In this paper I’ll direct my attention just to the topic of normative warrant.

As to the question of normative warrant, it will be granted on all sides that one pertinent norm is a matter of veracity. Is the agent who is blamed for something or other in fact blameworthy for it? Did she, for example, perform the action, or did someone else do it? Was what she did morally wrong or instead morally objectionable along some other dimension, or is this instead a matter of dispute? Did she do it under duress or in some other manner that compromised her freedom? Was she nonculpably ignorant in doing it? In considering whether an agent ought to be blamed, these questions of fact need to be settled. The relevant norm of veracity is that, at least as a pro tanto consideration, an agent ought to be blamed only if she is in fact blameworthy. The more challenging topic as regards the question of normative warrant is why, once it is settled that an agent is blameworthy, there is reason to blame that agent. This will figure prominently in subsequent discussion.

3. Blame as a Response to Quality of Will

I turn in this section and the next two to an account of blame’s nature. P. F. Strawson (1962) remarked that we care a great deal for the regard that others have for us, especially when that regard is revealed in how they act toward us. He then proceeded to account for the phenomenon of blaming in terms of responding to such regard. On the Strawsonian view I endorse, blaming another for something she has done is primarily, albeit not exclusively, a matter of responding in a distinctive fashion to the perceived morally objectionable quality of an agent’s will as manifested in her blameworthy behavior.

What is meant by the expression ‘quality of will’? On my view, it is not about identifying a distinct faculty, a will, and attending to some particular features of it. It’s rather a matter of the value or worth of an agent’s regard for another, or other salient considerations (McKenna 2012, pp. 57–63). This value can be good, ill, or indifferent. It can also be further qualified as moral or nonmoral. So understood, the moral or nonmoral worth of an agent’s regard for another can be manifested in the content of her intentions or her reasons for action. But it also can be revealed by her failure to show due regard for someone who, or something that, she should have, and thus in her failure for this to figure in her intentions or reasons at all.

One important factor in assessing the moral quality of an agent’s will has to do with the moral status of her action—for instance, whether the agent acted

---

2 The qualification “other salient considerations” is meant to allow for morally objectionable behavior that involves not necessarily harm to another person but instead to, say, nonhuman animals, or the environment, or perhaps a value or ideal.
morally wrongly and so violated a moral obligation. Another is whether she did so freely. A third is whether she did so either knowingly or from nonculpable ignorance. Note that we have here three variables. One concerns some moral evaluation of the nature of the agent’s action. A second concerns a control or a freedom condition. And a third concerns an epistemic condition. Each can contribute to an assessment of the moral quality of an agent’s will. Although controversial, I do not think these factors are always sufficient to discern the quality of an agent’s will as it bears on blameworthiness. In many cases they will be. But it is possible to imagine cases in which an agent freely and knowingly does something morally wrong and yet she does not act from a morally objectionable quality of will. Suppose, for instance, that she is acting in the context of a moral dilemma. (Think of Sophie from Sophie’s Choice.) Or she might be acting under some sort of duress that does not exempt her from her obligations or take her freedom from her. In these kinds of cases, she might very well not act from any objectionable lack of regard for moral considerations. Indeed she might be pained by the thought of acting as she does. If so, I contend, she is not blameworthy (McKenna 2012, pp. 14–20).

As the preceding discussion shows, blameworthiness involves different sorts of evaluative considerations. As a form of moral appraisal, one evaluative ingredient of blameworthiness is focused on the action performed by the blameworthy agent. This is settled by, for instance, determining that the agent acted morally wrongly. Here the evaluative focus is on the object of moral responsibility. But a distinct evaluative ingredient, I contend, is focused on the agent in relation to her action. Her regard for others and for salient moral considerations (such as the fact that she does moral wrong) is a further evaluative ingredient. In this case, the focus is on the subject of moral responsibility. Note that modifiers of ‘good,’ ‘ill,’ or ‘indifferent’ as pertaining to an agent’s will are axiological not deontological terms. And of course, insofar as the value or worth of an agent’s regard for her manner of acting also casts light on the nature and character of the agent, there is, at least potentially, an aretaic dimension as well. Hence, appraisals of blameworthiness can be evaluatively complex, encompassing deontic, axiological, and aretaic judgments.

This is an essay about the nature of blame, not the nature of blameworthiness. So, granting the preceding point, why is it important as regards blame? Because blaming is most fundamentally a response to perceived blameworthiness. It is useful to get a clear sense of what it is to which blame is sensitive. Indeed there is an important lesson about blame close at hand: variability in the fitting modes of blaming is liable to be a function of the variability in the evaluative cocktail that is embedded in a judgment of blameworthiness.

Of course, omissions or consequences of actions or omissions are also candidate objects of blameworthiness. I restrict attention here just to action solely for ease of exposition.
As for the evaluative component of blameworthiness involving the object of responsibility, some contend that moral blameworthiness is limited in its potential objects to morally wrong action (and omissions), and so to violations of moral obligations. In this way, it is thought that what an agent can be blameworthy for must be limited to the sphere of deontological evaluations (e.g., Darwall 2006; Wallace 1994). Naturally, on this view blame is limited accordingly. Although I wish to remain officially neutral on this point, my impression is that this is unnecessarily restrictive. Consider, for instance, the category of the supererogatory. Why are actions that can only be negatively evaluated in axiological or instead aretaic terms not even potential objects of blameworthiness? I fail to see why exactly a person cannot be blameworthy for acts that involve no violations of any moral obligations and so are not, strictly speaking, morally wrong, but are nevertheless morally bad, or instead unvirtuous. It may be that the mode of blaming that is fitting as a response to such acts is different. Perhaps it is weaker than that which is called for when an agent does something that is straightforwardly morally wrong. But as I see it, blameworthiness for such acts, and thus the prospect of blame directed at the agents who perform them, should be left open as a theoretically live option.

4. The Mode of Response Constitutive of Directed Blame

Grant, as I have argued in the previous section, that blaming is primarily a matter of responding to the morally objectionable quality of an agent’s will as manifested in her behavior. Grant also that blameworthiness is evaluatively complex and that blaming is liable to be sensitive to that complexity. The question remains, what sort of negative response is constitutive of blaming? More specifically, what sort of negative response is constitutive of directed blame? Is it primarily a matter of belief? Desire? Some combination? Along with other Strawsonians, I contend that it is most useful to understand moral responsibility and its cognate notions, blame being among them, by reference to a range of morally reactive attitudes. A reactive attitude is an attitude in response to the perceived attitude of another. The reactive attitudes at issue are not (merely) cognitive or conative but are affective; they involve emotions, in this case morally reactive emotions, and they pertain to the stance of holding morally responsible. Central to the current topic are those morally

---

4 Some (e.g., Haji 1998; Zimmerman 1988) deny that blameworthiness requires moral wrongdoing, but they do so in a manner that preserves a conceptual connection between blameworthiness and wrongness. On this sort of a view, moral blameworthiness involves acting with the belief that one is doing something that is objectively morally wrong.

5 There are as well reactive attitudes that are self-referential, such as guilt. Here I shall focus on reactive attitudes had by one that are directed at others.
reactive attitudes directly implicated in blaming. These particular reactive emotions are best understood as a species of moral anger, picked out by the terms of sentiment ‘resentment’ and ‘moral indignation.’

R. Jay Wallace (1994, pp. 25–33) made an important contribution to the reactive-emotions account of blaming by developing a normative interpretation whereby we can make good sense of a reactive emotion’s being appropriate or inappropriate or, perhaps put more cautiously, fitting or unfitting. The idea is to understand the emotional dimension of a phenomenon like blaming not merely in terms of a disposition to have and act from an emotion of a certain sort, but in terms of when it is apt or fitting to have one, and so when it is apt or fitting for one to respond to another in a blaming manner motivated by such an emotion. As Wallace noted, these attitudes have certain propositional objects at which they are directed. This allows us to make sense of how it is that these attitudes might be misdirected.

Wallace limited the relevant propositional objects to beliefs involving failures to comply with obligations or expectations. A blaming attitude of moral indignation, for instance, is inappropriate when the one at whom it is directed did not after all fail to comply with any obligation to which the one blaming held her. Or instead it is inappropriate when it is unfair to hold such a person to an expectation of this kind (perhaps she is severely mentally retarded and is just incapable of grasping or complying with such demands). While I agree with Wallace’s strategy regarding the normative interpretation of these reactive emotions and their place in blaming, I prefer a more inclusive and open-ended treatment of the range of propositional objects that bear on considerations of propriety or fittingness. By restricting the relevant morally reactive attitudes to beliefs about obligations, Wallace explicitly limits the domain of moral responsibility, and so the domain of blame’s target, to the deontic sphere of moral right and wrong (1994, p. 63). On my view, it is better to characterize the reactive attitudes of resentment and moral indignation so that they are sensitive to the belief that a person acted in a manner that is subject to some sort of moral criticism in deontological terms, but also, at least possibly, in axiological or aretaic terms.

On the view I propose, the species of moral anger distinctive of resentment and moral indignation is directed at a more complex sort of propositional object as in comparison with Wallace’s fairly lean, and admittedly more elegant, proposal. This species of moral anger, I contend, is aptly responsive to all of the following: the belief that an agent acted freely; the belief that an agent acted knowingly or from nonculpable ignorance; the belief that an agent’s act

---

8 There is some disagreement as to how exactly resentment and indignation differ. As I see it (McKenna 2012, p. 66), resentment is at play when one is oneself the person wronged or harmed or is in some manner directly affected or targeted by the objectionable behavior eliciting the emotion. Indignation is at play when it is another who is wronged, harmed, or in some way affected by the eliciting behavior.
was morally criticizable in some manner (was morally wrong, morally bad, or vicious); and, *most prominently*, the belief that an agent acted from a morally objectionable quality of will, either by acting with ill will or by acting in the absence of a sufficiently good will.

Strawson himself (1962) discussed the reactive attitudes in the context of their central role in adult interpersonal life, and most notably in our complex web of social practices. This emphasis is highlighted by Gary Watson’s (1987b) observation that for Strawson, holding someone morally responsible means something in practice. Blaming via resentment or indignation reveals itself in the altered means of interacting with the one blamed. Normal courtesies are withheld, patterns of conduct are changed, expected social plans and arrangements are altered, and particular means of expressing one’s moral anger in word and deed are found to be common in ways that are sometimes fitting and sometimes not. Although those discussing Strawson’s work often note this point, in my opinion it has not been fully appreciated. A focus *just* on the range of emotions and their aptness invites the misleading thought that what matters most fundamentally in the affective dimension of a phenomenon like blaming is the feeling part that can be privately experienced and concealed. In my estimation, this is backwards. As I have argued elsewhere (McKenna 2012, pp. 69–72), the conceptually more fundamental cases of emotions like resentment and moral indignation are those that encompass their characteristic behavior manifestations in ways that reveal their place in our interpersonal lives. We are able to understand privately experienced resentment, for example, by reference to the public cases.

To help illustrate the way these emotions are manifested in practice, I have focused on a simple case of blame between two friends and coworkers, Daphne and Leslie, who meet at a local shop for an afternoon coffee (McKenna 2012, pp. 68–69). In response to what she took to be an offensive racist joke about Hispanics, Daphne becomes morally indignant toward Leslie, angrily tells her she does not welcome the remark, and then storms out of the coffee shop. Daphne then alters future plans with Leslie, not inviting her to a lunch date with another coworker, who is Hispanic, and in various other ways makes fitting alterations to her means of interacting or not interacting with Leslie. Further, as I set up the case, Daphne takes there to be limits to the modes of blaming Leslie. For example, she does not let it affect her judgments about whether to offer Leslie some extra work. The point of calling attention to these details is to help show that these different ways that Daphne acts, from her angry verbal confrontation to her altered lunch plans, are not distinct from and so simply caused by her blaming Leslie; they’re constitutive of her blaming Leslie. They express her moral indignation and so are not *merely* motivated by her moral indignation. The role and presence of the emotion is infused in the altered social practices themselves, giving them a salience or meaning that they would otherwise not have.
Watson (1987b) has claimed that the morally reactive attitudes are incipently forms of communication. This is dead right. It is shown especially by focusing on cases of directed blame like the case of Daphne and Leslie. When the focus on the morally reactive attitudes is on their manifestation in our social practices and as directed at those blamed, we can see how it is that an episode of resentment or indignation can communicate moral demands and expectations. It can, as well, communicate an altered regard for the one held to blame and an indication of the likely means of future treatment, further expectations about means of redress, moral atonement, compensation, and so on. 7

5. Directed Blaming and Conversation

On the conversational theory of moral responsibility I defend, the actions of a morally responsible agent are potential bearers of a species of meaning, agent-meaning (McKenna 2012, pp. 92–94). As bearers of agent-meaning, such actions function as (fallible) indicators expressive of the quality of will with which agents act. When one directly blames another, she responds to the agent in light of this species of meaning. Her blaming response, via an overt manifestation of a pertinent reactive attitude such as resentment, communicates her regard for the blamed agent and that agent’s action in light of the quality of will that is presumed to be manifest in it.

Just now I characterized actions as being bearers of a species of meaning and of expressing quality of will. I also wrote of directed blaming via the morally reactive attitudes and altered interpersonal practices as playing a communicative role. But I have proposed a bolder thesis (McKenna 2012, ch. 4). The notions of meaning, expression, and communication do not fully capture the distinctive interanimation between blameworthy agent and those blaming, at least when the kind of blame is directed. More mileage can be gotten out of Watson’s insight that the reactive emotions as manifested in practice have a communicative role. The relation between a morally responsible agent and those who hold her to account for her blameworthy conduct, I contend, can be usefully illuminated on an analogy with a conversation. When a blameworthy agent acts, she understands that her actions are liable to interpretations in which members of the moral community might assign a salience or meaning to her mode of acting as indicative of one kind of quality of will or other. As a morally responsible agent, when she acts, she is aware that her actions can take

7 One wrinkle I have not taken in this section but have elsewhere (McKenna 2012, p. 25) has to do with the prospect of one’s holding morally responsible another who is blameworthy in cases in which one does not experience an episode of a morally reactive emotion at all. As Wallace has noted (1994, pp. 76–77), we can still hold a person morally responsible for a blameworthy act in those cases where, for whatever reasons, we just cannot muster the emotion.
on meaning. Hence, when her acts are pertinently morally charged, we can understand her as if she were introducing, or risking the possibility of introducing, a meaningful contribution that is a candidate for a conversational exchange with others. I have called this stage *Moral Contribution*. When one holds another morally responsible by blaming her, on the theory I have proposed, what she communicates can be understood on analogy with engaging in a conversation with the agent who initiated that exchange. This second stage, the stage wherein one blames, I have called *Moral Address*. It is then open to the blamed agent to extend the “conversation” further by means of offering an excuse, a justification, an apology, and so on. This third stage I have labeled *Moral Account*. It is then open to those holding morally responsible by blaming to extend the conversational analogue further, by, say, forgiving, or punishing, or simply ending the exchange and moving on, and so on. This dynamic, which we might call a *moral responsibility exchange*, is modeled on analogy with a conversational exchange between competent speakers of a natural language.

Understood in this light, moral responsibility is shown to be a dynamic process that is *essentially* interpersonal. It relates those who are and those who hold morally responsible in a fashion structurally analogous to the sort that unfolds between competent speakers of a natural language when involved in a conversational transaction. The nature of blaming, and specifically directed blaming, is a distinctive move in that kind of social practice. The conversational role of particular instances of blaming will arise against the backdrop of, or put differently, within the framework of, patterns of social life wherein variation from expectations will have a certain salience; for example, were Leslie not typically involved in lunch outings with Daphne, Daphne’s failure to invite her could not take on the meaning it does in the case as I have set it out.

Theorizing about moral responsibility in general and blame in particular as I have here has a payoff in terms of extending the analogy with features of literal, linguistic conversational exchanges. Speakers of a natural language engaged in the give and take of conversational transactions rely on complex expectations of shared background assumptions of a sort that allow for successful implicatures and related modes of conveying meaning that cannot be captured simply by attending to the strict semantic content of what is said between interlocutors. Innuendo, sarcasm, things discretely not said or not even indirectly mentioned, all figure into our understanding of the delicate interactions of individuals engaged in felicitous (and infelicitous) conversations with each other. Pragmatic context takes center stage here. A structurally analogous point applies to the nature of blaming understood as functioning like a move in a conversation. An altered pattern of behavior by one person as a means of manifesting her indignation could very well be taken to have a salience by a blamed party that it would not and could not have for another person. That other person’s interactions and relations with the blamed party
might be quite different, or might involve different roles (say, boss or spouse rather than mere acquaintance or casual friend). Hence, there is a sufficiently complex web of social interrelations between persons to make sense of something analogous to phenomena like implicatures in the modes of blaming that are available to members of a moral community. This helps to explain an observation made early on: that two individuals might very well blame another in ways that are equally warranted and fitting, but do so in wildly divergent ways.

Before turning in the next two sections to the norms of blaming, I pause to comment on two points about blame's nature which show where I part company with other Strawsonians. The first has to do with the modal status of claims about the relation between blaming and the morally reactive emotions. What I take to be essential to moral responsibility is that it is deeply interpersonal. It relates those who are morally responsible agents to those who hold morally responsible. This relation is essential to the nature of being morally responsible. This is in contrast to, for example, a ledger theory (e.g., Zimmerman 1988; Haji 1998). On a ledger theory, being morally responsible is, so to speak, metaphysically settled irrespective of any of the phenomena related to holding morally responsible. So I contend (McKenna 2012, pp. 80–88), along with other Strawsonians, we cannot make good sense of being morally responsible in the absence of some understanding of the standpoint of holding morally responsible. (Just as we cannot understand competent speakers of a natural language without understanding the nature of the audience to whom such speakers address themselves.) But are emotions required for this? Well, that depends on how strongly we understand the modal force of 'required.' Is it a metaphysical or a conceptual necessity that blaming involves some reference to an episode of a reactive attitude like resentment or indignation? I think this goes too far and is unnecessary to account for the phenomenon at hand. It seems to me to be at the very least a conceptual possibility that there could be beings who engage in social practices that are like ours, who hold each other to account for their morally significant conduct, and so on, but are emotionless beings altogether. Of course, qua agents who act at all, they would be motivated beings; they’d have a conative structure but would be devoid of affect.

It could be countered on behalf of Strawsonians that any essentialist claim about the place of the reactive attitudes is indexed to beings like us, to human persons. Rigidly designating humans, the claim might be, the morally reactive emotions are essential to accounting for moral responsibility in general, and the nature of blaming in particular. A yet weaker claim involves rigidly designating not humans, but instead just our practices—moral responsibility as we understand it—and then claiming that it is essential to those practices and this way of life that the moral emotions play a role. Perhaps this is correct, but I see no motivation to make any of these fairly strong modal claims, even the weakest of them. Why can’t the relation between these emotions and the social practices I have highlighted here be contingent but nevertheless deeply
embedded, so much so that we just would not get a credible theory of blame without attending to the role of these emotions? At any rate, as I see it, the morally reactive attitudes need only be thought of as contingently related to the nature of blaming (McKenna 2012, pp. 110–14). Regardless, they are deeply embedded in our blaming practices, and we would not understand our blaming practices, such as they are, without reference to them.

Here is the second point. Strawson (1962) and most Strawsonians (e.g., Bennett 1980; Wallace 1994; Watson 1987b) contend that the nature and norms of holding morally responsible are more fundamental than the nature of being morally responsible. The latter is not to be thought of as having a metaphysical standing to which the former must comply, but rather vice versa. I disagree. This is shown by how I have approached the topic of blame's nature in this and the two preceding sections. To get clear on blame, I have first attended to what it is to which blame is a response. Thus, I have focused on features of the agent who is morally responsible (by attending to quality of will considerations) as a way of helping to understand blame as a mode of holding morally responsible. It might be thought that I am forced into the opposing camp of those who would regard being morally responsible as the more basic or fundamental thing. But on my view, the mistake is to think that either is more basic than the other (McKenna 2012, pp. 50–55). There is a relation of mutual dependence that cannot be eliminated, much as we cannot make sense of what it is to be a competent speaker of a natural language without reference to the standpoint of a speaker's audience of potential interpreters who seek to understand what the speaker means to convey.

6. Directed Blaming and the Question of Normative Warrant

I turn now to the question of the normative warrant for directed blaming. What justifies its being the case that a blameworthy agent ought to be blamed? Also, how ought an agent to be blamed? There are at least two identifiable norms that bear on these questions. One, as previously noted (section 3), has to do with veracity. Is it true to the facts that the alleged blameworthy agent acted freely? Did she act with knowledge or culpable ignorance in doing what she did? Was what she did morally wrong, bad, or unvirtuous? Did she act from a morally objectionable quality of will?

A second norm arises naturally out of a conversational theory of moral responsibility. It has to do with meaningfulness or intelligibility—what makes sense as a sort of conversational response to the agent in light of her blameworthy action (McKenna 2012, p. 90). Just as in literal conversational transactions between competent speakers of a natural language, so too with blaming: there can be more and less sensible or intelligible responses to a meaningful contribution. In literal conversations, some conversational replies are infelicitous,
complete nonstarters. This happens when, for instance, there is a breakdown in communication, perhaps because one misheard what another said or did not pick up on an intended implicature. But other contributions can be felicitous and still be, on a scale, more fitting or apt than others. Sometimes, as in comparison with another conversationalist, one knows “just what to say,” and even here, there's no reason to think that there was, literally, just one correct way to contribute to or continue the conversation. I contend likewise for blaming. Some modes of blaming are infelicitous; they miss their target altogether and just do not make sense as a fitting response to one who is taken to be blameworthy. But as for those that do make sense, that are meaningful ways of communicating with a blameworthy party, there is no regimented way this must be done. Daphne blamed Leslie in, we can grant, a fitting manner by manifesting her moral indignation as she did. But we can equally well imagine that she could have done so by other means that made just as much sense—that is, that were just as conversationally meaningful. Here again we have a further explanation of why quite different modes of blaming can be regarded as equally suitable.

Norms of conversational intelligibility or meaningfulness also help call to attention a further feature of the permissible variability of blaming. In literal linguistic conversations, a contribution by an interlocutor is often pregnant with meaning so that there are different aspects or dimensions to what the speaker means. It is open to another conversing with the speaker to carry the conversation in a range of different ways, depending on which aspects of the original contribution this interlocutor wishes to focus on. Likewise for blaming as a “conversational” response to the perceived agent-meaning of a blameworthy party's act. Recall that judgments of blameworthiness are evaluatively complex. They can encompass deontological, axiological, and aretaic ingredients. If so, then it is open to one who replies to the blameworthy agent to attend more to one than another element in this evaluative cocktail.

As I have argued elsewhere (McKenna 2012, pp. 150–54), the two norms just countenanced—the norms of veracity and conversational intelligibility—are not adequate to account for the normative warrant of blaming. They're necessary but not sufficient. Directed blaming is liable to harm, and the norms canvassed thus far do not account for why it is that it is at least permissible to harm by directly blaming those who are blameworthy. What further norm will fill this bill? To answer this question, we need a clearer sense of just what about blaming constitutes the harm in it. Here we first must settle on what is meant by ‘harm’ and then return briefly to the topic of (directed) blame's nature.

As for the matter of harm, I follow Joel Feinberg in thinking of harm in terms of a setback to one's interests (1986, pp. 33–34). Feinberg held that the many different kinds of interests available to persons can be sorted into two general categories: welfare interests and ulterior interests. Ulterior interests concern aspirations and ultimate goals of the sort that figure in writing a novel or seeking a cure for cancer. Welfare interests are those more fundamental
interests that serve as a foundation for one's pursuing ulterior interests. Among these welfare interests Feinberg listed continuance of one's life into the future, physical health, integrity and functioning of one's body, absence of absorbing physical pain and grotesque disfigurement, emotional stability, the capacity to engage normally in social intercourse and enjoy and maintain friendships, minimal income and financial security, a tolerable social and physical environment, and a certain amount of freedom from interference from others (p. 37).

Now return to directed blame's nature. Due to blame's communicative and conversational role, it is liable to impose on the one blamed in particular ways. Anger, shunning, and alienation as expressions of morally reactive attitudes are often conveyed in blaming, and there is an expectation that the one blamed ought to reply by offering an apology or an explanation, revising modes of behavior, and so on. This can be emotionally taxing. It can compromise a person's welcomed relations with others. It can also cost her in terms of psychic energy and the freedom to live her life without having to pay the interpersonal costs of altered relationships and unpleasant demands and expectations put to her by those holding her to account. These conversational burdens placed upon her—even if she wishes to shrug them off—come at a cost to her, or at least they are liable to do so for all but the most hardened and indifferent among us. But notice that these costs are restricted to just some of the welfare interests Feinberg mentioned, in particular:

- The capacity to engage normally in social intercourse and enjoy and maintain friendships
- Freedom from others' interference
- Emotional stability

So, in directed blaming, a distinctive, albeit limited, class of welfare interests is liable to be compromised. A setback to these interests constitutes the unique harm(s) in directed blaming (McKenna 2012, pp. 134–41).

One further point about blame's nature before returning to the matter of blame's norms: On the conversational theory I endorse, and in opposition to how many others understand the issue, blame is distinct from punishment in a number of ways; it is not merely something like informal or "mini-punishment." For one thing, on the conversational theory I endorse, punishment is better thought of as having a distinct communicative and conversational role as in comparison with blaming, and so as coming at a distinct stage in a moral responsibility exchange. Blame occurs at the stage Moral Address and precedes the stage Moral Account. Punishment is suited for a stage that follows the stage Moral

---

8 My treatment in what follows of the harm and also the normative warrant for blaming shares affinities with an especially penetrating treatment of the topic by Christopher Bennett (2002).
9 See, for example, Bennett (2002), Feinberg (1970, p. 68), and Sidgwick (1874, bk. 3, ch. 5), who think of blame as just a species of punishment.
Account. At the stage Moral Account, the blamed agent is afforded the opportunity to offer some explanation, justification, apology, and so on. Punishment follows as a further response. But more to the immediate point of concern, the harms constitutive of punishment include a much greater range of setbacks to a person’s welfare interests, including a compromised physical environment and a loss of greater freedoms (confinement), on some views, physical harm (corporal punishment), loss of life (capital punishment), and certainly the prospect of financial insecurity (fines and penalties). The harm in blaming, limited to the distinctive class of welfare interests as I have set it out, is also limited in how much it can harm. Unlike the potential range of harms in punishment, the harm in blaming does not expose the blamed to the kind and degree of suffering that is at least available for various modes of punishment. The point is important since it bears keeping in mind when reflecting on what sort of warrant is called for to justify the harms in blaming. Some are against the practice of blaming because they conflate the harm in blaming with the harm in punishing, and then, thinking that the potentially extreme harm in punishing cannot be justified, draw the same conclusion about any normative warrant for the harm in blaming.\textsuperscript{10}

So in light of these considerations, what further norm might account for the warrant to harm in ways that are distinctive of directed blaming? One way to distinguish between different approaches to this question is in terms of whether it is good to harm a blameworthy person by blaming her. Some contend that it is no part of their account of the normative warrant for blaming that it is in any way good that the blameworthy person be harmed by way of blaming her.\textsuperscript{11} Others contend that it is. Call the former \textit{nondesert theorists} and the latter \textit{desert theorists}. Of those advancing desert theories, some contend that the ground for this goodness issues from more basic normative considerations.\textsuperscript{12}

This would be analogous to what in the sphere of distributive justice Rawls (1971, p. 104) called postinstitutional desert. On such a view, the goodness of the harm in blaming would flow from some sort of consequentialist or contractualist proposal, perhaps. Others advancing a desert theory contend that the ground for this goodness issues from more basic normative considerations.\textsuperscript{12} This would be analogous to what in the sphere of distributive justice Rawls (1971, p. 104) called postinstitutional desert. On such a view, the goodness of the harm in blaming would flow from some sort of consequentialist or contractualist proposal, perhaps. Others advancing a desert theory contend that the pertinent

\textsuperscript{10} There is yet another distinction between blame and punishment that appears to make justifying blame easier than doing so for punishment. In order to punish, one must intend to harm. But this is not so for blaming (McKenna 2012, pp. 144–46). All one must intend, in cases of directed blaming, is that one communicate one’s altered regard for the blamed agent in light of the quality of her will. One might thus intentionally harm the one blamed or instead just do so knowingly without intending to so harm her. (For a convincing treatment of the distinction between intending, intentional, and knowing, see Mele and Sverdlick [1996, pp. 273–74,].) Some object that certain theories of punishment are objectionable due to their connection with the vengeful intention to cause harm to those who are to be punished. Even if this were grounds for indicting those theories of punishment, it would miss its mark if targeted to similar theories of blame.


\textsuperscript{12} This seems to be the view that some, such as Ayer (1954), Hobart (1934), and Schlick (1939), endorsed. A more refined version, and an especially plausible one, as I read him, is found in Lenman (2006).
ground does not derive from more basic normative considerations; its desert claim is basic. The goodness of harming one who is blameworthy by the very activity of blaming her is a matter of its intrinsic or noninstrumental value. It is grounded just in the very fact of its being a distinctive kind of fitting response to the one blamed. Call these basic desert theories and nonbasic desert theories.

I find promise in efforts to account for the normative warrant to blame either by way of a nondesert theory or by way of a nonbasic desert theory (McKenna 2012, pp. 154–64). Each has merit and is worth exploring. In my view, those inclined to dismiss these strategies, taking them to be on their face inadequate to the task of accounting for moral responsibility, are mistaken to do so. Others, such as Pereboom (2009), far more cautiously acknowledge that these approaches offer accounts of bona fide senses of moral responsibility. They take these senses to be worth cultivating, but they reject the accountability senses of responsibility and blame that are in dispute in areas such as the free will debate. I disagree (McKenna 2012, pp. 171–72). I think it is a live option to argue that one can get the normative warrant for the pertinent sort of blaming and responsibility (the accountability sense) by way of some strategy other than appeal to basic desert. Nevertheless, I also think that, relying on the conversational theory, formulation of a defensible basic desert thesis for blame’s warrant is a live option (p. 172). Space does not permit consideration of the full range of theoretically live options. Because many take a basic desert thesis to capture a widely held conviction about moral responsibility’s normative warrant, in what follows I’ll devote my attention just to the articulation of a basic desert thesis for blame.

For example, see Bennett (2002), Pereboom (2001), and Sher (2006). Of course, as is well known, Kant endorsed a version of this view.

I prefer the term ‘noninstrumental’ rather than ‘intrinsic’ despite the fact that many, such as Wallace (1994, p. 60, n. 13), are careful to use the term ‘intrinsic.’ Why the more inclusive formulation? On a narrow understanding of ‘intrinsic,’ in contrast with ‘extrinsic,’ something does not have intrinsic value but instead extrinsic value if its value is dependent on its relation with other things. (Example: a flower’s being valuable because rare depends on whether there are others like it.) Things with extrinsic value can still, however, have value that is not instrumental—is not in the service of other things. I wish to leave it as an open theoretical possibility that on a basic desert theory for blame, the harm in blaming has extrinsic value (insofar as it depends on its relation to “conversational” dynamics of moral responsibility relations), but that this value is nonetheless noninstrumental. (I am indebted to conversations with Piers Rawling and Mark Timmons on this point.)

Justin Coates and Neal Tognazzini have raised an especially insightful question about my proposal thus far. I granted above that the norms of veracity and conversational intelligibility are insufficient to account for the normative warrant to harm by way of blaming. And I have turned to the question of whether an appeal to desert can bridge that gap. But why doesn’t the conversational norm already have a consideration of desert built into it? The norm would be of the following sort: It’s not conversationally fitting to blame one who does not deserve it. This is a nice question. And I suppose one could contend that we could get at least this much commitment to desert from my conversational proposal. But note here that the appeal to desert is meant to do no more than function as a defeater to the claim that it would be warranted to harm a blameworthy person. Appeal to desert is not here doing the work of providing a positive justification for why, in responding in a conversationally meaningful way, it would be at least permissible to engage in a practice that is liable to harm the one blamed. I am interested in what might account for this more robust justificatory question.
7. Directed Blaming and Basic Desert

In the previous section, to draw the distinction between desert-based and nondesert-based theories, I formulated the desert thesis in axiological terms. But this was slightly misleading. There is some discrepancy as to whether a basic desert thesis ought to be formulated in axiological terms or instead in deontological terms. Some, such as Feinberg (1970), Sher (2006), and Zimmerman (1988), characterize a basic desert thesis for blame in terms of what is required, right, or instead permissible. Others, such as Bennett (2002) and Scanlon (1998), do so in terms of what is good. A useful proposal, which I will adopt here, joins the two. Here is a formulation of a generic and modest basic desert thesis that draws on both axiological and deontological ingredients:

GD: Because it is a noninstrumental good that a wrongdoer is harmed, it is permissible to harm her.

Why is GD generic, and why is it modest? It is generic because it is not tailored to blame in particular and so could be used in a justification for punishment as well. It is modest because it does not add further content that many retributivists take to be essential to basic desert, such as a requirement of proportionality for deserved punishment, as well as some consideration of answering a harm with a deserved like harm (e.g., an eye for an eye). Its modesty is also a function of its deontic status, which commits to no more than permissibility, as opposed to a stronger version that would be expressed in terms of obligation. To the extent that these further desert theses make justifying basic desert more difficult because they command harsh(er) treatment for wrongdoers, or instead impose greater burdens on others, GD is by comparison more modest. It makes room for the possibility that it might be good and so permissible to harm a wrongdoer by harming her in ways that are in no way as great as was the sort of harm she caused. It avoids any connection to barbarisms of the sort associated with taking an eye for an eye. And it also does not require of others that they harm wrongdoers.

On my view, the claim of noninstrumental goodness is to be understood in terms of cross-world differences. Consider two worlds exactly alike in every respect as regards their histories and as regards two qualitatively indistinguishable agents, each of whom do something morally wrong. The worlds now differ in just the following way: In one world the wrongdoer is harmed in response for her wrongdoing. In the other she is not. In the former, due to her wrongdoing, she is made worse off than she otherwise would be. In the latter she is not. According to those who would advocate a thesis like GD, or some stronger variant, the former world is noninstrumentally a better world in

---

Wallace formulates a desert thesis in this manner, treating retributivism as “the view that it is intrinsically good that wrongdoers should suffer harm, and that we therefore have a positive moral obligation to inflict such harms on them” (1994, p. 60, n. 13).
comparison with the latter. Put simply, in the former world, the wrongdoer gets something that she deserves.

As it is, GD needs developing. It does not distinguish between culpable and nonculpable wrongdoing. For example, some who are wrongdoers might not have done so freely, or instead might have done so from nonculpable ignorance. Furthermore GD restricts the desert thesis to wrongdoing, which some might contend is problematic. But set these concerns aside since our interest here is in a more refined version of a desert thesis that is more specifically about blame and the harms distinctive of blame. The reason to begin with the more generic GD is just to call to attention that, as a general point, one that is not unique to the advancement of a theory like my conversational theory of blame, desert theses of the sort that apply to both blame and punishment can be fairly modest and so can avoid the more excessive commitments that are often the target of critics' most damning objections to claims of basic desert.

So now let us apply these considerations to the distinctive characteristics of directed blame as I have characterized it. Consider, then, this principle:

DB: Because it is a noninstrumental good that one who is blameworthy is harmed by the communicative practices constitutive of directed blaming, it is permissible to harm her by directly blaming her.

Like GD, DB also needs further development. For instance, some qualification is needed to accommodate considerations of normative standing. It might be amended to read, “it is permissible for those with proper standing to blame her.” We might also amend DB by making perspicuous that the mode of directed blaming that harms is by way of a conversationally intelligible or meaningful reply to the agent's blameworthy conduct. But considerations such as these aside, DB will suit our current purposes.

Apply the same thought experiment about cross-world differences to the claim of noninstrumental goodness in DB. Consider the case of Daphne and Leslie. Would it have been a better world had Daphne not blamed Leslie as she did when Leslie made that racist remark? Would it have been a better world if Daphne had not responded in a conversationally apt way to Leslie about the meaning of her (Leslie's) remark, and so had not expressed her moral demands and expectations? And would it have been a better world had Daphne's taking that stance and engaging in those practices not harmed Leslie in just the following specific ways? (1) Leslie was emotionally upset and hurt by Daphne's withdrawal of her goodwill toward Leslie and her (Daphne's) expressed anger. (2) Leslie's personal freedom from interference was strained by virtue of the social pressures Daphne placed on her. These pressures were to acknowledge the morally objectionable meaning of her remark and the quality of her will reflected in it and to respond to Daphne's mode of blaming her. And (3) Leslie's capacity to maintain and enjoy friendships with Daphne and other
coworkers was diminished by virtue of how Daphne communicated her regard toward Leslie in her (Daphne’s) withdrawal from normal, friendly personal interactions.

It is open to those advocating a desert thesis for blame as expressed by DB to contend that these distinctive harms to Leslie are noninstrumentally good. It would be a worse world were Leslie not to have been harmed in just these very special ways. Of course, it is open to those who argue against a desert thesis to dispute this. They might ask why it would not be an even better world if Daphne and Leslie were to be engaged in the sort of conversational analogue I have highlighted but Leslie were not harmed in that process. As regards this manner of resisting a basic desert thesis for blame, first, bear in mind that all I have attempted to do in this and the preceding section is articulate a basic desert thesis for blame. I have not attempted to argue for it. But, second, note also that this claim about cross-world comparisons has to be assessed in terms of what sort of agents we can expect the likes of Daphne and Leslie to be. Were we all closer to angels, then we might hope for modes of communicating our blaming regard in ways that are not liable to harm at all, and we might hope that those blamed are not disposed to experience these sorts of harms while nevertheless remaining open to the moral entreaties of others. But given that we are the kinds of creatures we are, it is not clear how we could engage in these practices at all and how they could have the kind of meaning for those blamed they have without their also carrying the characteristic sting associated with them. Nor is it clear how those doing the blaming could with sincerity communicate and engage with those blamed while fully suppressing the emotional springs that often motivate the likes of mere mortals like us. So in making the cross-world comparisons currently under consideration, great care has to be taken in thinking about whether, at pertinent worlds, we are departing far too much from the kinds of creatures that in fact we are, and in fact we are able to be.

The modest desert thesis, DB, when developed within the framework of a conversational theory of moral responsibility, has a considerable theoretical payoff. Some philosophers committed to one species or other of moral responsibility skepticism have made their task far too easy. They have presumed that the moral responsibility at issue concerns claims of basic desert, and then they have ramped way up the dramatic nature of what is or potentially could be deserved, especially as regards the blameworthy. Galen Strawson (1994), for instance, executes his impossibilist argument for moral responsibility by contending that genuine responsibility of the sort that, he contends, is not metaphysically possible but is nevertheless presupposed by our moral responsibility practices, requires that we at least be able to make sense of heaven and hell

17 For such an argument, see my 2012 (pp. 164–71).
responsibility—which includes hell's eternal torments and suffering. Others, such as Derk Pereboom (2009), make far less dramatic claims but still maintain that the kind of moral responsibility that implicates basically deserved blame is liable to license excessive means of acting on one's moral anger in ways that can lead to violence and other forms of extremely harsh treatment. As should be clear, the basic desert thesis DB, as applied to the conversational theory of moral responsibility, has no such implications. If the accountability sense of moral responsibility, the sort at issue in the free will debate, really does require a commitment to basically deserved blame, then these moral responsibility skeptics might have made clear why excessive versions of desert theses should be rejected, but they have not told us enough to explain why the claims of goodness and permissibility constituting DB, constrained as they are by a conversational theory, are to be rejected.

8. Concluding Remarks

In this essay I have drawn on the conversational theory of moral responsibility to offer an account of the nature and the norms of blame. I have done so by limiting the relevant notions of blame and of moral responsibility just to the accountability sense, and then to prototypical cases of blaming in the form of what I have called directed blame. I shall close by commenting briefly on a few points that I was not able to attend to in detail in the preceding discussion.

One dimension of directed blame's nature has to do with its relation to the voluntary. Curiously, few critics of emotion-based theories have raised worries about the control a blamer can be expected to exercise over her blaming. But there is a worry here that needs addressing. For the most part, except in fairly esoteric contexts, the mere experiencing of an emotion is not thought to be under an agent's direct voluntary control. If so, it looks like emotion-based theories of blaming risk the objection that one who blames is not in a position to settle freely on whether or not to blame. This objection need not be developed in terms of a blamer's having no control over whether she experiences an episode of a blame-constituting emotion. There is the notion of nonvoluntary control of the sort competent rational agents exercise in belief formation and retention. An agent who resents someone for wrongdoing can, for instance, exercise rational control over her emotion. She can do so to the extent that she can assess the truth as to whether, for instance, that agent really did wrong. But this being the case, she is not able to simply choose or decide directly whether, with these factual matters settled, she undergoes the emotion. Agents can't control their emotions directly in the requisite way.

18 Of course, there are indirect methods. One can avoid conditions in which she might have occasion to experience the emotion, or can have a stiff drink ready at hand, and so on.
If those advancing an emotion-based theory of blame were not able to say more about the control an agent can exercise in blaming, this would create problems accounting for the relation between the nature of blaming and its norms. Some of these norms concern questions of whether an agent ought to be blamed, and not just whether it is true to the facts that she is blameworthy. But if merely by having an episode of a pertinent reactive emotion, one does blame, it appears that one cannot comply with relevant judgments about whether, all things considered, an agent ought to be blamed. As might be expected, the way out for the emotion-based theorist is by tending to features of blaming that are within an agent’s voluntary control. The outward manifestation of blaming practices in modifications to interpersonal relations involves acts and omissions that, it can be argued, are within a blamer’s control. Here we have further reason to attend to the (I say) more fundamental prototypical cases of blame outwardly directed at the blamed. Doing so makes clear that the relevant norms of blame—and more particularly the expectation that blamers can comply with them—are not defeated merely by the fact that blame’s nature has an emotional and so nonvoluntary dimension.

I close with one final comment about the methodology I have adopted. Some will be disinclined to a conversational theory, or more generally any communication-based theory, because it seems too easy to generate counterexamples to it. There are cases in which an agent is morally responsible for some act, being either praiseworthy or blameworthy for it, and there is no conversational analogue that takes place. Or instead, it will be protested, in cases of private blame or blaming the dead, we have blame but no communication at all. So a theory such as the one I have advanced and a treatment of blame by reference to it is just inadequate. It tends to some special range of cases but not others. Scanlon, for instance, has objected to Watson’s contention that blame is to be understood in terms of its communicative role. Why? Because a person can blame while not communicating anything (Scanlon 2008, pp. 233–34, n. 54).

My reply to this objection is that it assumes that a proper theory of moral responsibility should be cast in terms of necessary and sufficient conditions à la proper classical analysis. But I reject this way of proceeding. Phenomena such as blaming are too diffuse in their extensions, and our intuitions about them are too contested in a wide range of cases, to hope that we can get anything like a proper analysis of the full extension of these concepts. Rather, it is best to proceed by focusing on central cases, ones that are of immediate concern as regards questions of normative warrant, standing, and force. If we can account adequately for the nature and norms of these, then the other cases can be explained by reference to these prototypical cases. My conversational theory of moral responsibility attends most directly to our adult interpersonal relations with each other, and the ways these do alter, and give good reason to
alter, our moral standing with those who are capable of morally significant action. It does so by finding meaning in the actions of morally responsible agents and in the means of conversationally engaging with them in ways that communicate moral demands and expectations. I propose that blame, at least the sort that bears on moral responsibility in the accountability sense, is best understood by adopting this strategy.