

Desert, Control, and Moral Responsibility

Douglas W. Portmore

I follow Strawson (1962) and others in thinking that a subject is morally responsible for having ϕ -ed if and only if she's the appropriate target of reactive attitudes in virtue of having ϕ -ed. But, of course, there are different sorts of attitudes that one could appropriately have in reaction to a subject's ϕ -ing. Consequently, there are different types, or faces, of moral responsibility. On one type, the relevant reactive attitudes paradigmatically include shame in the case of blaming oneself and disdain in the case of blaming others. Such attitudes are appropriate in reaction to a subject's ϕ -ing whenever she has, in ϕ -ing, revealed that her values, commitments, and/or character traits are flawed in some way. To illustrate, take the case of *The Misogynistic Professor*. He calls on several men but on no women during his three-hour-long seminar even though there were numerous occasions when just as many women had their hands raised. And let's assume that this behavior reveals that the professor's character is indeed misogynistic. It is, then, appropriate for the professor to be ashamed of, and for others to have disdain for, his misogynistic character. For, in having these reactive attitudes, they are (and he is) attributing a flawed character to him (or himself), thereby blaming him (or himself) in what's known as the attributability sense.

But there are other sorts of reactive attitudes that are appropriate in the case of *The Misogynistic Professor*. For, in behaving misogynistically, the professor not only reveals his flawed character but also violates certain legitimate demands. As a consequence, he seems liable to sanction, at least to the sort of sanction that comes from the disapproval of his own conscience. Thus, it seems appropriate for the professor to be held to account for having violated these legitimate demands. And he can hold himself accountable by feeling guilt while others can hold him accountable by resenting him and/or being indignant with him. And this is to blame him in what's known as the accountability sense.

So, there seem to be at least two types, or faces, of moral responsibility and, hence, at least two types of blameworthiness (WATSON 1996).¹ For one, a subject is blameworthy in the *attributability sense* for having φ -ed if and only if she is the appropriate target of reactive attitudes such as shame and disdain in virtue of having φ -ed and thereby having revealed her flawed values, commitments, and/or character traits. For another, a subject is blameworthy in the *accountability sense* for having φ -ed if and only if she is the appropriate target of reactive attitudes such as guilt, resentment, and indignation in virtue of having φ -ed and thereby having violated some legitimate demand.² These two types of blameworthiness differ not only in the sorts of reactive attitudes that they take to be paradigmatically relevant, but also in their control conditions.³ For, as noted by Gary Watson (1996), it seems that avoidability is a requirement for a subject's being blameworthy in the accountability sense but not for her being blameworthy in the attributability sense. He says, "the pressure for some kind of...principle of avoidability comes entirely...from accountability and its corresponding notion of blame" (1996, p. 229). For whereas it seems appropriate for the professor to be ashamed of his character whether he could have avoided developing that sort of character or not, it seems appropriate for him to feel guilty about having developed such a character only if he could have avoided doing so.

In this paper, I take it for granted both that there are these two types of blameworthiness and that avoidability is necessary for only accountability blameworthiness. My task, then, will be to explain why avoidability is necessary for accountability blameworthiness but not for attributability blameworthiness. I'll argue that what explains this is both the fact that these two types of blameworthiness make different sorts of reactive attitudes fitting and that only one of these two types of attitudes requires having been able to refrain from φ -ing in order for them to be fitting.

1. Accountability, Attributability, and Avoidability

¹ Some think that there are more than two. For instance, David Shoemaker (2015) argues that there are three distinct types: attributability, answerability, and accountability.

² See CARLSSON FORTHCOMING, NELKIN 2015, and WATSON 1996.

³ For a defense of the view that these are the relevant reactive attitudes for each type of blameworthiness, see CARLSSON FORTHCOMING.

A subject is accountable for having φ -ed if and only if she can appropriately be held to account for having φ -ed and, thus, appropriately held liable to reward or sanction in virtue of her having φ -ed. The reward or sanction needn't come from the law, society, or common opinion, but it must at least come from the approval or disapproval of her own conscience—see MILL 1861, chap. V. Thus, a subject is blameworthy in the accountability sense for having φ -ed if and only if it would be appropriate for her to feel guilt for having φ -ed. Yet, guilt is inherently unpleasant, and no one deserves to suffer any unpleasantness unless they had the opportunity to avoid it. So, if the appropriateness of one's feeling guilt depends on one's deserving to suffer its associated unpleasantness, then a subject can be blameworthy in the accountability sense for having φ -ed only if she could have refrained from φ -ing. And, so, we get the following argument for why avoidability is necessary for accountability. I call it the Guilt Argument.

- (G1) A subject is blameworthy in the accountability sense for having φ -ed only if it's appropriate for her to feel guilt in virtue of having φ -ed. [Analytic]
- (G2) Given both that guilt is inherently unpleasant and that no one deserves to experience any unpleasantness unless they had the opportunity to avoid it, it's appropriate for a subject to feel guilt in virtue of having φ -ed only if she could have refrained from φ -ing. [Assumption]
- (G3) Therefore, a subject is blameworthy in the accountability sense for having φ -ed only if she could have refrained from φ -ing. [From G1–G2]

Now, it may seem strange that I've formulated this argument only in terms of guilt and not also in terms of other relevant reactive attitudes, such as resentment and indignation. But although blameworthiness in the accountability sense can make all three of these reactive attitudes appropriate, it's only guilt that is inherently unpleasant for its target. One can be the target of resentment and/or indignation without experiencing any unpleasantness. For one can be unaware of the fact that one's the target of resentment and/or indignation. Or, alternatively, one can be aware of this fact but not care about it and, so, suffer no unpleasantness as a result.

By contrast, it's impossible to be the target of guilt and be unaware of it. To be the target of guilt, one must feel guilt. And to feel guilt is to be aware of it. Moreover, it's impossible to feel guilt without suffering any unpleasantness. For guilt is inherently unpleasant.⁴ If the feeling one is experiencing isn't unpleasant, it can't be guilt. Thus, given that the Guilt Argument relies on the assumption that no one deserves to suffer any unpleasantness unless they had the opportunity to avoid it, guilt is the only one of these three attitudes that's relevant. For it's the only one that's necessarily unpleasant for its target.

It may also seem strange that the above argument focuses solely on blame rather than on both praise and blame. After all, moral responsibility concerns praiseworthiness just as much as it does blameworthiness. But I make blameworthiness my exclusive focus only so as to simplify our discussion. Indeed, I believe that the above argument would be just as plausible if I were to make the following substitutions throughout: (1) 'pleasant/ness' for 'unpleasant/ness', (2) 'praiseworthy' for 'blameworthy', and (3) 'pride' (or whatever the positive analogue of guilt is) for 'guilt'. But although I could easily discuss and defend such an argument along with the Guilt Argument, I don't see any benefit to doing so. It seems that this would just lengthen our discussion without providing any additional insights. So, I leave it as exercise for the reader to think about how I could make analogous, and equally plausible, claims about pride, pleasantness, and praiseworthiness.

Lastly, it may seem strange that I've relied on the Guilt Argument to explain why avoidability is necessary for accountability when I deny that avoidability is necessary for attributability. For just as guilt is inherently unpleasant, so is shame. And, thus, it may seem that the Guilt Argument is no more plausible than the following analogous argument for why avoidability is necessary for attributability. I call it the Shame Argument.

- (S1) A subject is blameworthy in the attributability sense for having φ -ed only if it's appropriate for her to feel shame in virtue of having φ -ed. [Analytic]

⁴ I'm not alone in thinking that guilt is inherently unpleasant. See, for instance, CARLSSON 2017 (p. 91), CLARKE 2016 (p. 122), MORRIS 1976 (p. 101), ROSEN 2015 (p. 67, n. 6), and WOLF 2011.

- (S2) Given both that shame is inherently unpleasant and that no one deserves to experience any unpleasantness unless they had the opportunity to avoid it, it's appropriate for a subject to feel shame in virtue of having φ -ed only if she could have refrained from φ -ing. [Assumption]
- (S3) Therefore, a subject is blameworthy in the attributability sense for having φ -ed only if she could have refrained from φ -ing. [From S1–S2]

Given that the point of this paper is to explain why avoidability is necessary for accountability blameworthiness but not for attributability blameworthiness, I must hold that, while the Guilt Argument is sound, the Shame Argument is unsound. And since both arguments are clearly valid and since both G1 and S1 are analytically true, I'll need to argue that, while G2 is true, S2 is false.⁵ To do so, I'll need to get clear on what specific notion of appropriateness is at work in these two arguments. For there seem to be various different notions of appropriateness that could be at work, depending on what the relevant norm is. Here, then, are three possible candidates for being the relevant norm of appropriateness.

- *The Norm of Fairness*: Some treatment is fair and, therefore, appropriate in the sense governed by this norm if and only if it doesn't entail treating differently those who are alike in all the relevant respects.
- *The Norm of Desert*: Some subject morally deserves X such that it's appropriate that she gets X in the sense governed by this norm if and only if, as a matter of justice and in virtue of her possessed characteristics or prior activities, she merits X in the sense that entails that it would, in some respect, be non-instrumentally good that she gets X.

⁵ This set up of the problem involving both the Guilt Argument and the Shame Argument is inspired by the same sort of set up found in CARLSSON FORTHCOMING.

- *The Norm of Fittingness*: Some attitude is fitting and, therefore, appropriate in the sense governed by this norm if and only if its representations are accurate.⁶

Let's consider each in turn, starting with the norm of fairness. We may think that "it is unfair to impose sanctions upon people unless they have a reasonable opportunity to avoid incurring them" (WATSON 1996, p. 237). And we may additionally think that in having an unpleasant feeling like guilt or shame one imposes a sanction upon oneself. And if both of these thoughts are correct, it will be appropriate (in the sense that's governed by the norm of fairness) for one to feel guilt or shame in virtue of having φ -ed only if one could have refrained from φ -ing. So, if the norm of appropriateness that's being appealed to in the above two arguments is the norm of fairness, both G2 and S2 are going to be true. But we should reject this line of reasoning. First, to have an unpleasant feeling such as guilt is not to impose a sanction upon oneself. To impose guilt upon oneself, one would need to be able to voluntarily choose to have this feeling. But although we can feel guilt in response to reasons, we cannot voluntarily choose whether to feel guilt—that is, we cannot feel guilt simply by intending to feel guilt. Thus, guilt is not something we impose upon ourselves and, so, can't be something that we unfairly impose upon ourselves.

Second, it seems mistaken to think that the relevant notion of appropriateness is the one governed by the norm of fairness. For if this were correct, whether it would be appropriate for us to blame someone would depend on whether we would also be blaming those who are alike in the relevant respects. But it seems that whether it's appropriate for us to blame, say, Hitler for certain German atrocities has nothing to do with whether we also blame Pol Pot for certain Cambodian atrocities. It might be unfair for us to blame Hitler and not Pol Pot, but that wouldn't make it any less appropriate for us to blame Hitler—at least, not in the sense that's relevant to the above two arguments. For it seems that whether it is, in the relevant sense,

⁶ Admittedly, there are different interpretations of the notions of desert, fairness, and fittingness. For instance, some deny that desert has any axiological implications. And some take desert to be just a kind of fittingness. But whether you agree with my labels for these norms is not what's important. What's important is that there are these different norms of appropriateness.

appropriate for us to blame Hitler depends solely on whether he's blameworthy and not at all on whether we would be treating him or others unfairly in doing so.

Perhaps, then, the above arguments are appealing to the norm of desert. If so, we should accept that it's appropriate for someone to feel guilt or shame for having ϕ -ed only if she could have refrained from ϕ -ing. For, as we've seen, both guilt and shame are inherently unpleasant and no one deserves to suffer any unpleasantness unless they had the opportunity to avoid it. So, if the relevant notion of appropriateness were the one governed by desert, G2 and S2 would both be true.

But this isn't the relevant notion of appropriateness. For one can blame someone by having reactive attitudes such as shame or disdain. And it seems that someone could be worthy of such blame without deserving to be the target of such attitudes. Of course, we need to be careful here, for sometimes 'deserve' is used as a synonym for 'fitting' rather than as I'm using it: as a word referring exclusively to what's *morally* deserved. And it's only when 'desert' is being used to refer to what's morally deserved that the claim that a subject deserves X entails that it would, in some respect, be non-instrumentally good for her to get X.⁷ So, I concede that it would be felicitous to claim, say, that Southwest Airlines deserves a five-star customer-approval rating given its exceptional customer satisfaction. Yet, even if Southwest Airlines deserves a five-star rating in some sense, it doesn't *morally* deserve this rating. After all, it is not, in any respect, non-instrumentally good that Southwest Airlines gets this rating. If it's good, it's only instrumentally good. For there is nothing inherently morally good in Southwest Airlines' getting a customer-approval rating that accurately reflects the satisfaction of its customers. So, even if someone who is worthy of shame and/or disdain is necessarily deserving of these reactive attitudes in the sense of its being fitting for them to be the target of such attitudes (just as it's fitting for Southwest Airlines to receive a five-star customer-approval rating), it isn't non-instrumentally good for them to be the target of such attitudes. So, they don't morally deserve to be the target of such attitudes. And since the norm of desert refers to moral desert and not

⁷ Others who think that the claim that a subject deserves X entails that it would, in some respect, be non-instrumentally good that they get X include Carlsson (FORTHCOMING), Clarke (2013), and McKenna (2012). But Nelkin disagrees with this—see her 2016.

fittingness, we should reject the idea that the norm of appropriateness at work in the above two arguments is the norm of desert.

This leaves us with the norm of fittingness. And for an attitude to be fitting is for it to accurately represent the way things are. Thus, it's important to note that many attitudes have a representational component in addition to their affective and motivational components.⁸ And it's the attitude's representational component that sets its fittingness conditions, such that an attitude is fitting if and only if its representations are accurate.⁹ In this sense, the fear of X is appropriate if and only if X poses a threat, for the fear of X (at least, typically) represents X as posing a threat.¹⁰ So, in order to determine which of G2 and S2 are true, we'll need to determine which representations are constitutive of feeling guilt and which representations are constitutive of feeling shame. For it's only if shame and guilt have different constitutive representations, that G2 will be true while S2 is false. So, in the next section, I'll investigate which representations are constitutive of each of these two attitudes.

⁸ I'm not committed to the view that all attitudes (or that all emotions) necessarily have some sort of representational content. For instance, it may be, for all that I claim here, that anger needn't have any representational content. But I will insist that each of the following are essentially representational: guilt, pride, shame, disdain, resentment, and indignation. So, my view is compatible with the sort of view offered in D'ARMS AND JACOBSON 2003, where they distinguish between natural emotional kinds such as anger and what they call *cognitive sharpenings*. Cognitive sharpenings are a proper subset of instances of some natural emotional kind that are identified by their essential representational content. For instance, resentment is, on their view, a cognitive sharpening of anger that's identified as a species of anger that represents the world as being one in which the subject has been wronged. I should note, however, that they provisionally suggest that guilt is a natural emotional kind and not a cognitive sharpening, contrary to what I'll claim here (see 2003, p. 138).

⁹ Perhaps, this is too quick. For it doesn't seem appropriate for a worm to fear a bird unless the worm is capable of representing the bird as a threat to itself. So, I should probably qualify the above as follows: it's appropriate for a subject to form an attitude only if she has the option both to form this attitude and to form the representations essential to it. Of course, I'm not committed to fear having some essential representational content; it could, as D'Arms and Jacobson suggest (2003), be a natural emotional kind. But if it is, it's fitting when, and only when, it has some representational content and that representational content is accurate.

¹⁰ The idea that it would be fitting/unfitting for a subject to ϕ is distinct from the idea that it would be fortunate/unfortunate that she ϕ s. Consequently, we must allow that it could be fitting, say, to fear an animal even if this would be unfortunate given that the animal would then sense this fear and become even more of a threat as a result. Despite its being unfortunate to have this fear, it would, nevertheless, be fitting so long as it correctly represents its object as a threat. In general, attitudes represent their objects as being a certain way and are, therefore, fitting (that is, correct) to the extent that their representations are accurate. By contrast, an attitude is *fortunate* if and only if good consequences would result from one's having that attitude. For more on this distinction, see CHAPPELL 2012.

2. Fitting Guilt, Fitting Shame, and Avoidability

It seems to me that, in feeling guilt for having φ -ed, one represents the unpleasantness of this feeling as something that one deserves to experience in virtue of having violated a legitimate demand in φ -ing — that is, one represents this unpleasantness as something that it is morally and non-instrumentally good for one to be experiencing given that one has violated this demand (PORTMORE FORTHCOMING).¹¹ Note, then, that guilt isn't just any unpleasant feeling associated with certain motivational tendencies. To feel guilt, one must additionally represent the unpleasantness of this feeling as something that one deserves to experience in virtue of one's having violated a legitimate demand.¹² Thus, a woman with Tourette's may feel bad for having involuntarily uttered some obscenity, but this feeling won't amount to guilt unless she experiences its unpleasantness as at least partially deserved. What's more, she could both feel bad about having uttered the obscenity and experience this unpleasantness as at least partially deserved (perhaps, because she wrongly believes that people with her affliction deserve to suffer) and still not count as feeling guilt. For it will be guilt that she's feeling only if she represents its associated unpleasantness as something that she deserves *in virtue* of her having failed to live up to some legitimate demand (MORRIS 1976, pp. 61 & 104).

Of course, none of this is to suggest that it's impossible to feel guilt unless one *believes* that one deserves to experience the unpleasantness of this feeling in virtue of having violated a legitimate demand. Clearly, one can feel guilt without having this belief just as one can fear a large hairy spider without believing that it poses a threat. For although fearing something implicates what I'll call the "thought" that it poses a threat, one needn't *believe* that it poses a

¹¹ Likewise, it seems to me that when one resents, or is indignant with, a subject for having φ -ed, one represents the world as being one in which that subject deserves to suffer the unpleasantness of feeling guilt in virtue of her having violated a legitimate demand in φ -ing. Of course, in the case of resentment, one will additionally represent the world as being one in which one was a victim of this violation.

¹² Here, I concur with Darwall and Mill: "Mill calls guilt a kind of 'internal sanction', but it is important to appreciate that guilt is not merely painful, or the (painful) fear of further (external) sanctions (MILL 1998: Ch. III). It is the painful sense of having done wrong, having violated a legitimate demand that comes, not just from someone else, say God, but also that one implicitly makes of oneself, through blaming oneself in feeling guilt" (DARWALL 2013, p. 16).

threat to have this thought.¹³ Having the thought that it's dangerous necessitates only experiencing it as dangerous. And to experience it as dangerous is just for it to strike one as dangerous in the same way that the lines in a Müller-Lyer illusion can strike one as unequal even while at the same time believing that they're equal.¹⁴ Likewise, to have the thought that one deserves to experience the unpleasantness of guilt in virtue of having failed to live up to some legitimate demand, one needn't believe this. Rather, one need only have it strike one as being so. Thus, I'll be using the word 'thought' as a term of art such that one has the thought that p if and only if one represents the world as being such that p —that is, represents the world as being such that p in the same way that one can believe that the lines in a Müller-Lyer illusion are equal while representing them to oneself as unequal.¹⁵ But I take no stand on what exactly it is to have such a representation other than to say both that it's not a belief and that it's the sort of thing that we're having when the lines in a Müller-Lyer illusion strike as unequal. I will, though, assume that a representation that p (what I'm calling the thought that p) is accurate if and only if p . Thus, the thought that those lines are unequal—a thought that we have only while under the force of the illusion—is inaccurate because the lines are in fact equal.¹⁶

¹³ Here, I follow CLARKE 2016 and ROSEN 2015 in distinguishing *thoughts* from *beliefs* such that having the latter, but not the former, necessitates assenting to the attitude's representational content.

¹⁴ For more on this idea, see both CLARKE 2016 (pp. 122–123) and ROSEN 2015 (pp. 71–72).

¹⁵ Thus, my view is compatible with the possibility of experiencing recalcitrant guilt—that is, with the possibility of experiencing guilt while at the same time believing that one doesn't deserve to feel its unpleasant affect. But although this is possible, it is, on my view, no more possible for a subject to feel guilt without it seeming to her that she deserves to suffer its unpleasant affect than it is for a subject to feel resentment without it seeming to her that she's been wronged. And, contrary to D'Arms and Jacobson (2003, p. 143), I contend that there is no difficulty or instability in a subject's continuing to resent someone while believing that that someone didn't wrong her. Take, for instance, this real-life example. My wife once dreamed that I cheated on her. That morning, it was clear to me that she resented me. Initially, she denied it. But when I explained how she was treating me as if I had wronged her, she admitted that she had this dream and that, although she knew that I hadn't wronged her, it still seemed to her as if I had. Consequently, it was recalcitrant resentment that she was feeling. And this recalcitrant resentment continued stably for as long as it continued to seem to her that I had wronged her.

¹⁶ Note, then, that I'm not saying that an attitude is fitting if and only if it wouldn't be morally bad to have such an attitude. That would be to commit what Justin D'Arms and Dan Jacobson (2000) call the *moralistic fallacy*. For fittingness is a matter of accurate representation rather than moral goodness. And, thus, a reaction can be fitting even if morally bad. For instance, amusement can be the fitting response to a funny joke in that it accurately represents that joke as amusing even if that response would be morally bad given both that it's a cruel joke and that any visible or audible expression of amusement in response to it would cause harm to others.

It's because the unpleasantness of guilt strikes us as being at least partially deserved that we experience it quite differently than we do, say, the unpleasantness of a headache. Unlike the unpleasantness of a headache, the unpleasantness of guilt strikes us as something that it would, in some respect, be morally problematic to be rid of.¹⁷ After all, to say that a subject deserves X is to say that, as a matter of justice and in virtue of her possessed characteristics or prior activities, she merits X in the sense that entails that it would, in some respect, be non-instrumentally good that she gets X. Thus, taking a pill to alleviate one's appropriate guilt seems morally problematic in a way that taking a pill to alleviate one's headache does not. For whereas there is nothing non-instrumentally bad about ridding oneself of a headache, there does seem to be something non-instrumentally bad about ridding oneself of one's appropriate guilt.¹⁸ And this difference isn't just due to the fact that guilt, but not a headache, can be appropriate. For fear can be appropriate and yet there is nothing non-instrumentally bad about taking a pill to get rid of one's appropriate fear.¹⁹

Note that the thought implicated by guilt is not the thought that the wrongdoer deserves to suffer in general.²⁰ Rather, the thought is only that she deserves to suffer the specific

¹⁷ This thought is connected to several of Herbert Morris's thoughts: "the man who feels guilty often seeks pain and somehow sees it as appropriate because of his guilt; indeed, the feelings of guilt may disappear and the man may connect their disappearance with the pain he has experienced. When we think of what it is to feel guilty then, we think not only of painful feelings but of something that is owed; and pain is somehow connected with paying what one owes" (1976, pp. 89–90). Thus, "what is sought out is the pain of feeling guilty as punishment for wrongdoing" (1976, p. 104).

¹⁸ Clearly, one reason that it would often be morally problematic to take a pill to alleviate one's guilt is that experiencing guilt can often be instrumentally valuable in making one less likely to commit future wrongs. Likewise, shame can be instrumentally valuable in helping one to regulate one's conduct. But it seems to me that it would be morally problematic to take a pill to alleviate one's guilt even if experiencing that guilt would be of no instrumental value. It would be morally problematic in that one deserves to feel bad for violating a legitimate demand and it is morally and non-instrumentally good for people to get what they deserve. In this respect, then, guilt seems unlike shame.

¹⁹ Appropriate grief is a bit more complicated. There does seem to be something morally problematic about taking a pill to get rid of one's appropriate grief. But I suspect that this is because it may count as disrespectful to the one lost and/or as a form of denial that hinders one's ability to heal from that loss. In any case, the fact that there is nothing problematic about taking a pill to rid oneself of one's appropriate fear in instances where having that fear would be of no instrumental value shows that the mere appropriateness of a feeling doesn't determine whether it's something that it would be morally problematic to get rid of.

²⁰ Many find this idea unacceptable. For instance, T. M. Scanlon rejects the idea that "it is good that people who have done wrong should suffer" (2013, p. 102). Likewise, R. Jay Wallace rejects the "problematic thought that

unpleasantness involved in having this feeling (that is, the feeling of guilt) and that she deserves this in virtue of having failed to live up to some legitimate demand. To illustrate, suppose that my wife is accountable for having mistreated me. It would, then, be fitting for me to want her to feel at least some guilt for having mistreated me, and to want this even if such feelings would be of no instrumental value. Nevertheless, it would not be fitting for me to want her to feel lonely. Nor would it be fitting for me to want her to suffer more than she deserves to suffer.²¹ And, importantly, the thought that my wife deserves to feel guilt for her mistreatment of me implies only that it would, *in some respect*, be non-instrumentally good that she feels this way, not that her feeling this way would be *overall* good. Thus, the implication is only that the world in which she feels guilt for having accountably mistreated me is, other things being equal, non-instrumentally better than the world in which she likewise feels guilt for having non-accountably mistreated me.²²

wrongdoers positively deserve to suffer" (1994, p. 108). But rejecting this idea doesn't entail rejecting the idea that someone accountable for some wrongdoing deserves to suffer the unpleasantness of feeling guilt for having committed that wrongdoing. Indeed, Scanlon now accepts that wrongdoers deserve to feel guilt for their wrongdoing—see SCANLON 2008, p. 188. So, even if we reject the idea that wrongdoers deserve to suffer generally, we shouldn't necessarily reject the idea that wrongdoers deserve to suffer the specific unpleasantness of feeling guilt for their wrongdoing—see, for instance, MCKENNA 2012 (chaps. 6–7).

²¹ Why think that she deserves to feel bad at all? Here's my argument: (P1) Given that she's accountable for having mistreated me, it's appropriate to want her to feel guilt for having mistreated me and to want her to have this experience even if her having it wouldn't be instrumentally valuable. (P2) If it's appropriate to want X even if X wouldn't be instrumentally valuable, then X must be non-instrumentally valuable. (C1) Thus, her feeling guilt for having mistreated me is non-instrumentally valuable. (P3) What most plausibly accounts for C1 is that she deserves to feel guilt for having mistreated me. (C2) Therefore, she deserves to feel guilt for having mistreated me. And, in defense of P3, I would add, first, that what explains the non-instrumental value of her feeling guilt for having mistreated me is not that her having this experience is itself non-instrumentally valuable. It isn't. After all, her having this experience wouldn't be non-instrumentally valuable if she weren't accountable for her mistreatment of me. Second, the fact that it is *fitting* for her to feel guilt for having mistreated me is not what explains why her feeling guilty is non-instrumentally valuable. For, in general, there's nothing non-instrumentally valuable about having a fitting attitude. There's nothing, for instance, non-instrumentally valuable about fearing that which poses a threat even though it is fitting to fear that which poses a threat. Thus, it seems that what explains the fact that her feeling guilt is non-instrumentally valuable is both that she deserves to feel guilt and that it is, in general, non-instrumentally valuable that people get what they deserve.

²² If you think that there's nothing non-instrumentally good about her feeling guilty for having accountably mistreated me, then you would have to think (implausibly) that the world in which she feels guilty for having accountably mistreated me is, other things being equal, no better than the world in which she likewise feels guilty for having non-accountably mistreated me.

This, then, is a very minimal claim about desert; it claims only that feeling appropriate guilt is, other things being equal, less non-instrumentally bad than feeling inappropriate guilt.²³ And, yet, even this rather minimal claim is quite explanatorily useful. First, it explains why there is essentially something morally problematic about someone's taking a pill to alleviate her fitting guilt, whereas there need be nothing morally problematic about someone's taking a pill to alleviate her fitting fear. Second, it explains why there is nothing morally problematic in our preferring the guilty feeling guilt to the innocent feeling guilt. Third, it explains the way in which feeling guilt for having φ -ed differs from other ways of feeling bad about having φ -ed (such as shame and embarrassment): only feeling guilt necessitates the thought that one deserves to feel this way. Fourth, it explains why there is nothing morally problematic in our expressing our resentment and indignation in the hopes of getting the guilty to feel guilt. And, thus, it explains why we are appropriately angered and frustrated when the guilty respond to such expressions with only a sincere promise to do better next time, and without any hint of guilt or remorse. And, fifth, it explains why the sorts of excuses that exculpate us from blame (e.g., ignorance and a lack of control) are precisely those that diminish the degree to which we are deserving of the unpleasantness of guilt.

It seems, then, that the thought that's constitutive of a subject's feeling guilt for having φ -ed is the thought that she deserves to experience the unpleasantness of this feeling in the recognition that she has violated a legitimate demand in φ -ing. Call this *the desert view*. One reason to accept this view is that it is explanatorily useful in the five ways described above. But another reason is that, additionally, it helps us to explain why, after a sufficient amount of self-reproach, it ceases to be appropriate to feel guilt anymore.²⁴ For, on this view, the relevant

²³ Note that my very minimal claim about desert is even more minimal than what others consider to be a relatively minimal claim about desert—see, for instance, Carlsson's claim that "if an agent deserves some harm, it will be non-instrumentally good that this harm occurs" (2017, p. 99). These others are committed to the view that the world in which my wife feels guilty for having accountably mistreated me is, other things being equal, better than the world in which she doesn't feel guilty for having accountably mistreated me. On this view, her feeling guilty is not just, *in some respect*, non-instrumentally good, but is, *overall*, non-instrumentally good. I'm not committed to this stronger claim.

²⁴ As Morris notes: "feelings of guilt may disappear and the man [who used to feel guilty] may connect their disappearance with the pain he has experienced" (1976, p. 90). The idea, I take it, is that punishment (even self-punishment in form of guilty feelings) can undercut the appropriateness of feeling further guilt, for if we've suffered

thought is that one deserves to suffer the unpleasantness of this feeling, and this thought ceases to be true after a sufficient amount of self-reproach has occurred. After all, one deserves to suffer only so much self-reproach for having violated a legitimate demand. So, after a sufficient amount of self-reproach has occurred, one will cease to deserve to feel guilt anymore. And, consequently, it will cease to be appropriate for one to continue to feel guilt. So, where one has yet to suffer a sufficient amount of guilt, the thought constitutive of guilt—that is, “I deserve to suffer the unpleasantness of *this* feeling in virtue of having violated a legitimate demand”—will be true.²⁵ But, where one has already suffered a sufficient amount of guilt, this thought will be false. Thus, the desert view nicely explains why the appropriateness of one’s feeling guilt depends (in part) on whether one has already suffered a sufficient amount of self-reproach.

To better understand the merits of the desert view, it will be helpful to contrast it with the following potential rivals: (1) the quality-of-will view, according to which the thought implicated by guilt is that one manifested ill will in having φ -ed, (2) the blameworthy view, according to which the thought implicated by guilt is that one is blameworthy for having φ -ed, and (3) the wrong-doing view, according to which the thought implicated by guilt is that one was wrong to have φ -ed. I’ll take each in turn, and then, at the end, consider the possibility of combining one or more of these views with the desert view.

Consider, first, the quality-of-will view. For one, it seems possible to feel guilt for having φ -ed without having the thought that one manifested ill will in having φ -ed—that is, without it even seeming to one that one φ -ed out of ill will. For instance, a woman could feel guilt (although inappropriately so) for running over a boy who unexpectedly darted in front of her car without it seeming to her that her conduct manifested ill will. After all, the woman didn’t

enough for our wrongdoing, it is no longer appropriate for us to continue to suffer (1976, p. 62). As Brad Cokelet has pointed out to me, once you have suffered enough guilt, the wronged party may forgive you and rightly tell you that you shouldn’t feel guilty anymore. Of course, even if you’ve already suffered a sufficient amount of guilt, that doesn’t mean that those who you’ve wronged must forgive you, nor does it mean that you don’t still owe it to them to make amends and express your contrition.

²⁵ Note that the presence of the indexical ‘this’ in this proposition is what allows for the fact that the thought given in quotes can be true when entertained at a time prior to one’s having suffered a sufficient amount of self-reproach and then false when entertained at later a time—a time after which one has suffered a sufficient amount of self-reproach.

even have time to consider the boy before her car plowed into him. So, it won't seem to her that she bore him any ill will. Still, it could be guilt that she's feeling so long as it strikes her both that it is legitimate for others to demand that she not have killed the boy and that she deserves to suffer the unpleasantness of guilt in virtue of having violated this seemingly legitimate demand. Of course, it won't be fitting for her to feel guilt since these representations are clearly inaccurate, for, as we'll assume, it was impossible for her to have stopped in time. But it will strike her this way given that the mechanisms that she has for producing such mental states were trained, and/or evolved, under a range of contexts that differ significantly from the unusual one in which she now finds herself.

For another, if the quality-of-will view were right, there would be no reason to think that it would, after a sufficient amount of self-reproach, cease to be appropriate for one to feel guilt for having ϕ -ed out of ill will. For no matter how much one has already reproached oneself for having done so, it will never cease to be true that one ϕ -ed out of ill will. And, so, it will never, on this view, cease to be appropriate to feel guilt for having done so. Yet, intuitively, it seems that, after a sufficient amount of self-reproach, one should cease to rebuke oneself any further.

Next, consider the blameworthy view. Here, too, it seems possible for one to feel guilt for having ϕ -ed without having the thought that one was blameworthy for having ϕ -ed. For it seems that one needn't have the concept of blameworthiness to feel guilt. And this makes sense of our thought that children can feel guilt even if they lack the rather sophisticated concept of blameworthiness. This is because, as anyone with much experience with children knows, children have the considerably less sophisticated concepts of merit, desert, demands, and goodness, which is all the conceptual apparatus that, on the desert view, they need. Thus, it seems that what's needed to feel guilt is not the sophisticated concept of blameworthiness, but the less sophisticated concepts of merit, desert, demands, and goodness.

And, again, we see that, like the quality-of-will view, the blameworthy view has trouble explaining why, after a sufficient amount of self-reproach, it ceases to be appropriate to feel guilt anymore. For, again, the relevant thought—that is, “I was blameworthy for having ϕ -ed”—never ceases to be true. And, so, we must on this view counterintuitively hold that no

matter how much one has already reproached oneself for one's previous failure, it never ceases to be appropriate to continue to feel guilt for having φ -ed.

Lastly, consider the wrong-doing view. Here, too, it seems that one can feel guilt for having φ -ed without thinking that one has done wrong in having φ -ed. To illustrate, imagine that Huck Finn (Mark Twain's famous fictional character) had, contrary to the actual story, turned his friend Jim over to the authorities for being a runaway slave. In such a case, it seems possible for Huck to feel guilt for having turned his friend in even if it doesn't strike him as wrong to have done so. This is because whereas guilt seems to track the violation of any legitimate demand (be it moral or non-moral and be it overridden or not), moral wrongness seems to track only the violation of some non-overridden moral demand. So, even if, as Huck supposes, it wasn't wrong for him to turn Jim in, he did violate a legitimate demand of friendship in doing so. Thus, so long as it strikes him that he deserves to feel bad for having turned Jim over to the authorities in virtue of his thereby violating Jim's demand for the protection of friendship, it could be guilt that he's feeling. And, so, we should reject the wrong-doing view, for it incorrectly assumes that we must have the thought of having done wrong to feel guilt. What's more, we should reject the wrong-doing view, because, like the other two potential rivals, it fails to account for the fact that, after a sufficient amount of self-reproach, it ceases to be appropriate to feel guilt anymore. And it fails for the exact same reason that the other two failed: the thought in question—in this case: "I was wrong to have φ -ed"—is not one that ceases to be true after one has suffered a sufficient amount of self-reproach.

Of course, even if, unlike me, you think that some of these other thoughts (e.g., the thought that one was wrong to have φ -ed) are constitutive of one's feeling guilt for having φ -ed, this shouldn't cause you to reject my claim that at least one of the thoughts constitutive of feeling guilt is that one deserves to feel the unpleasantness that's inherent to it. For, in that case, you should just amend the desert view and hold something like the following view: the thought constitutive of feeling guilt for having φ -ed is the thought that one deserves to suffer the unpleasantness of this feeling in the recognition that one has done wrong (or has done something blameworthy or has manifested a poor quality of will) in φ -ing. For my purposes, all that matters is that the idea that one deserves to suffer the unpleasantness of this feeling in

virtue of having violated a legitimate demand is at least part of the thought (or one of the thoughts) that's constitutive of feeling guilt.

But what about shame? Shame is also unpleasant. But unlike guilt one needn't feel that one deserves to experience its associated unpleasantness in order to have this feeling. And this is why one can appropriately be ashamed of things over which one had no control. For whether it is appropriate for a subject to be ashamed of having ϕ -ed depends only on whether her having ϕ -ed reveals some flaw in her values, commitments, and/or character traits. Whether she could have avoided ϕ -ing is irrelevant, because whether she could have avoided ϕ -ing or not, it still reveals that she is flawed. Thus, we can appropriately be ashamed of our racist thoughts and feelings even if they come to us unbidden and unwanted. So long as they accurately reflect the fact that we're racist, it's appropriate for us to feel ashamed of having had such thoughts and feelings given what they reveal about us. And this points to another difference between shame and guilt. Unlike guilt, the thought of having violated a legitimate demand is not constitutive of shame. Rather, what's constitutive of shame is the thought that one fails to meet some standard in a way that could potentially result in a loss of honor, respect, or esteem. And the failure to meet such a standard could result in a loss of honor, respect, or esteem even if it would be illegitimate to demand that one meets it. And this is why someone can appropriately be ashamed of, say, having racist implicit attitudes even if it's illegitimate to demand that one not have such implicit attitudes given that there was nothing one could have done to avoid acquiring them through one's childhood experiences. Thus, the thought that's constitutive of being ashamed of X seems to be the thought that X reveals that one is substandard in a way that could potentially result in a loss of honor, respect, or esteem.

Now, although I think that this is the thought that's constitutive of shame, I'm not going to try to defend it in all its details. For my purposes, the only two claims about shame that I need to insist upon are that (1) shame is inherently unpleasant and that (2) no thought that's constitutive of shame necessitates that one must have been able to avoid that which one is appropriately ashamed of. Of course, my account of shame's constitutive thought certainly allows that one can be appropriately ashamed of things that one couldn't have avoided. But it doesn't matter whether my account is correct so long as the correct account also allows for this.

And it seems to me that it is widely recognized that one can be appropriately ashamed of things that one couldn't have avoided.²⁶ And this means that whether an account of the thought that's constitutive of shame is plausible or not is going to depend in large part on whether it allows that one can be appropriately ashamed of things that one couldn't have avoided.

So, it seems that we can safely assume that no thought constitutive of shame requires that one must have been able to avoid that which one is appropriately ashamed of. And that means that S₂ is false. For S₂ holds that, given both that shame is inherently unpleasant and that no one deserves to experience any unpleasantness unless they had the opportunity to avoid it, it's appropriate (that is, fitting) for a subject to feel shame in virtue of having ϕ -ed only if she could have refrained from ϕ -ing. But, as we've just seen, that's false. Yet, G₂ is true. For G₂ holds that, given both that guilt is inherently unpleasant and that no one deserves to experience any unpleasantness unless they had the opportunity to avoid it, it's appropriate (that is, fitting) for a subject to feel guilt in virtue of having ϕ -ed only if she could have refrained from ϕ -ing. And, given both that an attitude is fitting only if the thoughts constitutive of it are true and that the thought that's constitutive of guilt is the thought that one deserves to experience the unpleasantness of guilt in virtue of having violated a legitimate demand in ϕ -ing, it follows that G₂ is true.

So, although many attitudes—including, guilt, grief, and shame—are inherently painful such that no one deserves to suffer them unless they could have avoided them, this doesn't mean that they're never fitting. Whether they're fitting just depends on whether they accurately represent the way things are. So, to illustrate, even if you don't deserve to suffer the unpleasantness of grief in virtue of the death of your loved one, it can still be fitting for you to grieve in virtue of that death, for your grief may well accurately represent that death as being a significant loss to you. Likewise, it can be fitting for you to experience shame in virtue of your implicit attitudes even if you were powerless to avoid acquiring them. This shame will be fitting so long as it accurately represents you as being substandard in a way that could potentially result in a loss of honor, respect, or esteem. But, by contrast, it can never be fitting for you to feel

²⁶ See, for instance, PRINZ & NICHOLS 2010.

guilt for having φ -ed unless you could have refrained from φ -ing. For the thought that's constitutive of your feeling guilt is, I've argued, the thought that you deserve to experience the unpleasantness of this feeling in virtue of having violated a legitimate demand. And no one deserves to experience any unpleasantness unless they had the opportunity to avoid it. So, you don't deserve to feel guilt for your implicit attitudes unless you could have avoided them.

I've argued, then, that what explains why avoidability is required for accountability blameworthiness but not for attributability blameworthiness is that, whereas the former requires that guilt be appropriate, the latter requires that shame be appropriate. And I've argued that, given the different thoughts that are constitutive of these two attitudes, only appropriate guilt requires avoidability. Thus, the Guilt Argument is sound, and the Shame Argument is unsound. For G2 is true, and S2 is false.

3. Objections

I've argued that what explains why avoidability is required for accountability but not for attributability is the fact that these two types of responsibility make different sorts of reactive attitudes fitting. Whereas accountability blameworthiness makes guilt fitting, attributability blameworthiness makes shame fitting. And I've argued that whereas feeling guilt for having φ -ed is fitting if and only if one deserves to suffer the unpleasantness of this feeling in virtue of having thereby violated a legitimate demand, being ashamed of X is fitting if and only if X reveals that one is substandard in a way that could potentially result in a loss of honor, respect, or esteem. This explains why avoidability is required for accountability blameworthiness but not for attributability blameworthiness. For although one can't deserve to suffer in virtue of having φ -ed unless one could have refrained from φ -ing, the fact that one has φ -ed could reveal that one is substandard even if one couldn't have refrained from φ -ing. This is, I think, a plausible explanation, but I will now consider three potential objections to it.

3.1 If guilt and shame have different constitutive thoughts, then how could pride be the positive analogue of both guilt and shame? It's essential to the explanation that I've provided that guilt and shame have different constitutive thoughts. But one may object that if guilt and shame had different constitutive thoughts, pride wouldn't be the positive analogue of both attitudes. And it

is quite common to suppose that pride is the positive analogue of both guilt and shame. But I don't think that pride—at least, not pride generally—is the positive analogue of guilt. To see why, consider that the thought that's constitutive of being proud of X is the thought that X reveals that one is particularly good (or exceptional) in a way that could potentially result in a gain of honor, respect, or esteem. But if that's right, then pride is too broad to be the positive analogue of guilt. For if pride were the positive analogue of guilt, it would be appropriate to feel guilt about the negative analogue of anything that we could appropriately take pride in. But that's not the case. For instance, it's appropriate to take pride in one's exceptional innate intelligence, but not to feel guilt for lacking such intelligence. Likewise, it's appropriate to take pride in one's exceptional innate physical prowess, but not to feel guilt for lacking such prowess. It seems that one can appropriately feel guilt only for something that one did or failed to do, and, thus, not for one's innate characteristics. Of course, one could appropriately feel guilt for having employed one's innate characteristics in the service of evil ends, but it's inappropriate to feel guilty about merely possessing the characteristics.

So, if some sort of pride is the positive analogue of guilt, it must be the sort that can be taken only in one's good deeds. And, to keep this specific sort of pride distinct from pride more generally, I'm going to call it *deontic pride*. Given that we can appropriately take deontic pride only in our good deeds, it seems that the thought constitutive of it is the thought that one deserves to experience the pleasantness of this feeling in virtue of having responded in some particularly good (or exceptional) way. Thus, in taking deontic pride in having performed some act, one needn't have the thought that this deed reveals that there is something particularly good (or exceptional) about oneself, but only that there is something particularly good (or exceptional) about one's deeds.

So, this objection fails. We should not think that guilt and shame both have pride as their positive analogue. At best, the positive analogue of shame is general pride, whereas the positive analogue of guilt is a specific sort of pride: namely, deontic pride. At worst, the positive analogue of guilt is some other attitude entirely: perhaps, being gratified that one has ϕ -ed. But even if the positive analogue of guilt is deontic pride, the objection fails. For deontic pride and

general pride have different constitutive thoughts, and, so, we have every reason to think that guilt and shame will as well.

3.2 *Perhaps, guilt and shame have different appropriateness conditions.* In arguing for my favored explanation, I've assumed that, when it comes to moral responsibility and the appropriateness of reactive attitudes, the relevant norm of appropriateness is going to be the same regardless of which specific reactive attitudes we're focused on. But not everyone agrees. For instance, Andreas Brekke Carlsson (FORTHCOMING) has argued "that accountability blame and attributability blame are governed by different notions of appropriateness: an agent S is accountability blameworthy for X only if S *deserves* to feel guilty; an agent S is attributability blameworthy for X only if it is *fitting* that S feels shame for X." This allows him to accept G₂ while denying S₂, and it allows him to do so without holding, as I do, that the thought that's constitutive of a subject's feeling guilt for having ϕ -ed is one that can be true only if she could have refrained from ϕ -ing.

But Carlsson's view is problematic for at least two reasons. First, Carlsson's view leaves unexplained why guilt and shame are governed by different norms of appropriateness. For why is it that the appropriateness of guilt is governed by the norm of desert when the appropriateness of every other attitude—hope, fear, envy, grief, shame, disgust, surprise, contempt, amusement, etc.—seems to be governed by the norm of fittingness? After all, it seems that guilt has a representational component just as much as these other attitudes do. And, thus, just as these other attitudes can be accurate or inaccurate in their representations of the world, so can guilt. Indeed, Carlsson concedes this. But, unlike myself, he doesn't hold that the thought that's constitutive of a subject's feeling guilt for having ϕ -ed is one that can be true only if she could have refrained from ϕ -ing. So, in order to explain why avoidability is necessary for accountability but not attributability, Carlsson has to assume that whereas one is blameworthy in the attributability sense only if shame is *fitting*, one is blameworthy in the accountability sense only if guilt is *deserved*. But in doing this Carlsson is forced to accept a less unified view of moral responsibility. For whereas I hold that a subject is morally responsible for having ϕ -ed if and only if she's *the fitting target* of reactive attitudes in virtue of having ϕ -ed, Carlsson must hold that a subject is morally responsible for having ϕ -ed if and only if she *either deserves to be*

the target or is the fitting target of reactive attitudes in virtue of having φ -ed, and he has to hold that whether its desert or fittingness that's relevant depends on which reactive attitudes are at issue. Yet, why think that moral responsibility concerns not only the fittingness of reactive attitudes, but also whether one deserves to experience their pleasant or unpleasant affect? And this disunification seems to come at a cost, a cost that we needn't bear. For if I'm right, we can account for the fact that avoidability is necessary for accountability but not for attributability simply by appealing to the different sorts of reactive attitudes that the two make fitting and the different constitutive thoughts associated with these different sorts of reactive attitudes.

Second, since Carlsson admits that guilt has representational content and can, therefore, be fitting or unfitting, he must allow for the possibility that guilt is like fear in that it can be fitting even if its subject doesn't deserve to suffer its unpleasant affect. But although fear can be both fitting and undeserved, I don't believe that guilt can. In this respect, guilt seems unlike fear. For although it seems that there are many instances in which it is fitting for me to feel fear even though I don't deserve to suffer any unpleasantness, there don't seem to be any instances in which it is fitting for me to feel guilt even though I don't deserve suffer any unpleasantness. It seems that if it's fitting for me to feel guilt, it must be because I deserve to suffer its unpleasant affect.

3.3 What explains why guilt ceases to be appropriate after one has suffered a sufficient amount of self-reproach is just the general fact that attitudes cease to be appropriate after a period of time, such as the period of time it takes to suffer a sufficient amount of self-reproach. I've assumed that the thought constitutive of feeling guilt must be able to explain why guilt ceases to be appropriate after one has suffered a sufficient amount of self-reproach, but it may be objected that it's just a general feature of attitudes that they cease to be appropriate after a sufficient amount of time and that this has nothing to do with their constitutive thoughts.

I admit that there does seem to be a sense of appropriateness in which the passage of time can make it appropriate to "move on" and to, consequently, feel attitudes such as guilt, grief, and anger less intensely and/or less frequently.²⁷ But I believe that the relevant sense of

²⁷ For two very interesting discussions of how, after the passage of time, it can be appropriate to feel attitudes such as guilt, grief, and anger less intensely and/or less frequently, see MARUŠIĆ FORTHCOMING and MOLLER MANUSCRIPT.

appropriateness is not the one governed by the norm of fittingness. For it seems that it never ceases to be fitting to feel guilt, grief, or anger (or to feel them any less intensely) simply as a result of the passage of time. Rather, it ceases to be fitting to feel such an attitude only if the world changes in such a way that the attitude in question no longer accurately represents the world as it once did. Thus, I think that guilt can cease to be fitting only if the world changes from being one in which the subject deserves to suffer more of the unpleasantness of guilt to being one in which she doesn't deserve this. For given what I take to be the thought that's constitutive of guilt, this thought will be true in the former but not the latter.

But why, then, does it seem that guilt, grief, and anger cease to be appropriate in some sense after the passage of time? Why does it seem that I should eventually just get over it and move on? Consider, for instance, that if I'm just as angry now in 2018 as I was back in 1974 when my mom wrongly punished me for something that my friend did, something is certainly amiss. Likewise, if, in 2018, I'm still grieving as much about my high-school buddy's death as I did right after it happened in 1985, then, again, something seems amiss. Yet, these attitudes don't seem any less fitting now than they were then. After all, their representations seem just as accurate now as they did then. For instance, my righteous anger for having been wrongly punished by my mom represents what my mom did as an injustice, and that this is no less true now than it was then. Likewise, the grief that I originally felt over the loss of my friend represents his death as being the loss of someone whom I deeply cared about. But this is no less true now than it was then. And if my intense grief in 1985 accurately represented the fact that I had lost a dear friend, why would it not be just as fitting to experience intense grief now? Yet, it seems inappropriate for me to grieve now as intensely as I did back then. But that, I'll argue, is not because it's any less fitting for me to feel intense grief, but because the relevant sense of appropriateness is one that's not governed by the norm of fittingness. I take the norm of appropriateness that accounts for why it is no longer appropriate for me to feel intense grief over the loss of my high school buddy to be the following.

- *The Norm of (Coherence) Rationality*: Some attitude is (coherence) rational and, therefore, appropriate in the sense governed by this norm if and only if it coheres well with one's other mental states.

So, whereas an attitude is fitting if and only if it accurately represents the way the world is, an attitude is rational in the coherence sense of rationality (and, hereafter, I'll often leave this qualification implicit) if and only if it coheres well with one's other mental states. So, whereas fittingness is beholden to a mind-independent reality, rationality isn't. Rationality (at least, coherence rationality) is beholden only to one's mind and its internal states. To illustrate, take belief. The belief that p represents the world as being one in which p is true. Thus, the belief that p is fitting if and only if p . So, if the number of atoms in the universe is some prime number, then it's fitting for you to believe that this is so. Nevertheless, it may not be rational for you to believe this. For whether it's rational to believe that the number of atoms in the universe is some prime number depends, not on what the world is like, but on what your other mental states are like. Thus, even if the number of atoms in the universe is some non-prime number, you would be rationally required to believe that it's some prime number if you had the following beliefs: (1) God created the universe, (2) God loves prime numbers, and (3) if the being who created the universe loves prime numbers, the number of atoms in the universe is going to be some prime number.

The same holds for other attitudes. So, whether it's rational for me to feel guilt, grief, or anger depends on what my other mental states are like and not on whether they accurately represent the way the world is. Take grief, for instance. Whether it is, at present, rational for me to grieve over the death of my high-school buddy depends, not on whether this grief would accurately reflect the fact that his death represents the loss of someone whom I deeply cared about, but on what other mental states I have at present. Specifically, it depends on what my occurrent hopes, wishes, desires, memories, and beliefs are. Thus, it would be irrational for me to be at present grieving over his loss if all my occurrent hopes, wishes, desires, memories, and

beliefs were focused upon others.²⁸ Indeed, if all my occurrent mental states revolve around happy events involving living people and none of them revolve around anyone who has passed away, it would be quite irrational for me to be grieving. Moreover, even if I did irrationally grieve, it's unclear how this grief could constitute grief over the death of my high-school buddy given that none of my occurrent mental states take him as their focus. By contrast, consider that during the days immediately following his death my mind was often consumed with thoughts of him: wishing he were still around, remembering the good times that we had had together, missing the jokes he would tell in the locker room after practice, thinking about how much I will miss him, and wanting there to be something I could say to his family to help console them, etc. But, as time passed, my memories of him have faded, my desires and thoughts have shifted away from him and towards those who I currently interact with regularly, and my hopes and plans for the future no longer include him, but include only others (Nussbaum 2001, p. 80). Consequently, although my grieving over his death would be no less fitting now than it was then, it would be a lot less rational given the way my desires, thoughts, and other mental states have shifted their focus away from him.

So, I believe that this objection also fails. Although I concede that with the passage of time it often ceases to be rational to feel guilt, grief, and anger with the same intensity and frequency that one initially felt, this is not because it becomes any less fitting to have these attitudes. Rather, with the passage of time, past events often tend to occupy less of one's mental attention. And, as one's mind shifts to more recent events, it becomes less and less rational (and, thus, less and less appropriate in the associated sense) to feel the way that one once did about these past events. For the rationality of an attitude is simply a function of how well it coheres with the rest of one's mental states. And, so, what explains why guilt ceases to be appropriate after one has already suffered a sufficient amount of self-reproach has nothing to do with the

²⁸ Of course, I will still have the non-occurrent belief that his death is bad and constitutes a significant loss for me. But I tend to think that this non-occurrent belief doesn't put much rational pressure on me to grieve just as the mere fact that there are certainly some logical inconsistencies among all my many non-occurrent beliefs puts little to no rational pressure on me to revise my beliefs. Such rational pressure arises only after, say, some philosophical inquiry brings the inconsistency to the forefront of my conscious mind in the form a set of jointly inconsistent occurrent beliefs.

passage of time or with one's other mental states and everything to do with the thought that's constitutive of guilt: the thought that one deserves to suffer the unpleasantness of guilt. After all, someone might be perfectly rational in feeling guilt even though she has already suffered enough. For someone may mistakenly feel that she deserves to suffer more than she does. And, so, her conscious mind could very well be preoccupied with thoughts about some past misdeed such that it is perfectly rational for her to feel guilt. But it won't be fitting for her to feel guilt unless she still deserves to suffer more of the unpleasantness of guilt.

It seems, then, that none of these objections succeed and so we have no good reason to reject the explanation that I've provided. We should accept, then, that the explanation for why avoidability is required for accountability blameworthiness but not for attributability blameworthiness is that whereas accountability blameworthiness makes guilt fitting, attributability blameworthiness makes shame fitting. And whereas it can be fitting to feel guilt for having φ -ed only if one could have refrained from φ -ing, it can be fitting to feel shame for having φ -ed even if one couldn't have refrained from φ -ing.²⁹

Conflict of Interest Statement

On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

Carlsson, A. B. (FORTHCOMING). "Shame and Attributability." *Oxford Studies in Agency and Responsibility*. Downloaded on September 2, 2018, from https://www.academia.edu/36133447/Shame_and_Attributability.

²⁹ For helpful comments and discussions, I thank Andreas Brekke Carlsson, Brad Cokelet, Josh Glasgow, Pat Greenspan, Berislav Marušić, Shyam Nair, David Shoemaker, Sergio Tenenbaum, and the organizers and participants of the 2018 Bled Philosophy Conference on Ethical Issues.

- . (2017). "Blameworthiness as Deserved Guilt." *Journal of Ethics* 21: 89–115.
- Chappell, R. Y. (2012). "Fittingness: The Sole Normative Primitive." *The Philosophical Quarterly* 62: 684–704.
- Clarke, R. (2016). "Moral Responsibility, Guilt, and Retributivism." *Journal of Ethics* 20: 121–137.
- . (2013). "Some Theses on Desert." *Philosophical Explorations* 16: 153–164.
- Darwall, S. (2013). "Morality's Distinctiveness." In S. Darwall (ed.) *Morality, Authority, and Law: Essays in Second-Personal Ethics I*, pp. 3–19. Oxford: Oxford University Press.
- D'Arms, J. and D. Jacobson (2003). "The Significance of Recalcitrant Emotion (or, Anti-Quasijudgmentalism)." *Philosophy* 52 (Supp): 127–145.
- . (2000). "The Moralistic Fallacy: On the 'Appropriateness' of the Emotions." *Philosophy and Phenomenological Research* 61: 65–90.
- Marušić, B. (FORTHCOMING). "Do Reasons Expire? – An Essay on Grief." *Philosophers' Imprint*.
- McKenna, M. (2012). *Conversation and Responsibility*. Oxford: Oxford University Press.
- Mill, J. S. (1991). [1861]. *Utilitarianism*. Reprinted in J. M. Robson (ed.), *Collected Works of John Stuart Mill*. Vol. 10, pp. 203–59. London: Routledge.
- Moller, D. (MANUSCRIPT). "Rationality and the Function of the Emotions."
- Morris, H. (1976). *On Guilt and Innocence: Essays in Legal Philosophy and Moral Psychology*. Berkeley: University of California Press.
- Nelkin, D. K. (2016). "Accountability and Desert." *The Journal of Ethics* 20: 173–189.
- . (2015). "Psychopaths, Incurable Racists, and the Faces of Responsibility." *Ethics* 125: 357–390.
- Nussbaum, M. (2001). *Upheavals of Thought: The Intelligence of Emotions*. Cambridge: Cambridge University Press.
- Portmore, D. W. (FORTHCOMING). *Opting for the Best: Oughts and Options*. New York: Oxford University Press.

- Prinz, J. J. and S. Nichols. (2010). "Moral emotions." In J. M. Doris (ed.) *The Moral Psychology Handbook*, pp. 111–146. Oxford: Oxford University Press.
- Rosen, G. (2015). "The Alethic Conception of Moral Responsibility." In R. Clarke, M. McKenna, and A. M. Smith (eds.) *The Nature of Moral Responsibility: New Essays*, pp. 65–87, Oxford: Oxford University Press.
- Scanlon, T. M. (2013). "Giving Desert Its Due." *Philosophical Explorations* 16: 101–116.
- . (2008). *Moral Dimensions. Meaning, Permissibility and Blame*. Cambridge: Harvard University Press.
- Shoemaker, D. (2015). *Responsibility from the Margins*. Oxford: Oxford University Press.
- Strawson, P. F. (1962). "Freedom and Resentment," *Proceedings of the British Academy*. 48: 1–25. Reprinted in P. F. Strawson (2008) *Freedom and Resentment and Other Essays*. London: Routledge.
- Wallace, R. J. (1994). *Responsibility and the Moral Sentiments*. Cambridge: Harvard University Press.
- Watson, G. (1996). "Two Faces of Responsibility." *Philosophical Topics* 24: 227–248.
- Wolf, S. (2011). "Blame Italian Style." In R. J. Wallace, R. Kumar, and S. R. Freeman (eds.), *Reasons and Recognition: Essays on the Philosophy of T. M. Scanlon*, pp. 332–347. New York: Oxford University Press.