

# Guilt, Grief, and The Good<sup>1</sup>

Dana Kay Nelkin  
February 16, 2018

## I. Introduction

Consider the following thesis:

***Desert***: The blameworthy deserve to suffer.

To some, Desert seems obvious; to others it is contentious, non-obvious, or even false.<sup>2</sup> It is possible that apparent disagreement about Desert ultimately rests on the fact that it can be interpreted in significantly different ways. One thing that is clear is that much is at stake, including the justification of punishment, the justification of many of our blaming practices and reactive attitudes such as resentment and indignation, and more. And it has been thought by many that we need to figure out whether anyone has free will in order to figure out whether anyone is blameworthy, precisely because of the truth of Desert.

To get clearer on what Desert means, we should ask two questions:

- (i) What do the blameworthy deserve to suffer?
- (ii) What does “deserving to suffer” mean or entail?

A variety of answers has been given to question (i), including all of the following: hard treatment, setback of interests, eternal torment, to feel guilty to the proper degree. In this paper, I focus on elaborating and evaluating the last of these:

**Desert-Guilt:** The blameworthy deserve to feel guilty to the right degree.<sup>3</sup>

There are several reasons for doing so. First, there has been a recent convergence of articles on this answer in both the moral and legal literature on responsibility.<sup>4</sup> Second, it has been thought to have a kind of intriguing advantage in responding to skeptical arguments concluding that no one is really blameworthy. As some have noted (e.g., Michael McKenna), one might be tempted by the idea that the bar of justification is lower to justify inducing guilt feelings for wrongdoing than eternal torment or even hard treatment.<sup>5</sup> That is, if eternal torment is at stake, it might seem that the conditions for blameworthiness must be extremely robust whereas if guilt feelings are at stake, the conditions could be more inclusive. (I should say that I have some doubts about this as a general thesis; feeling guilt might—for some anyway—be more painful in the relevant sense than, say, hard labor. But the idea that this thesis points to a lower bar is intriguing.) And finally, there is simply something very appealing about the thesis which no doubt explains the recent attention to it.

When it comes to (ii), the question about the nature of desert, following Feinberg, I take it that desert may very well be unanalyzable, but that there are things to say *about* it. For example, what is deserved is valenced and scalar in a way parallel to the objects of desert, that is, to what it is that renders one deserving. Part of what makes all of the above candidates for answers to (i) is precisely that they are negatively valenced. Two further implications that have been claimed to help *explicate* desert are the following:

**Good-Desert:** If x deserves y for a'ing, then it is non-instrumentally good that x gets y for a'ing.<sup>6</sup>

**Reason-Desert:** If x deserves y for a'ing, then there is some *pro tanto* (or defeasible) reason to promote x getting y for a'ing.

Now some use the consequents in both principles interchangeably with the claim that x deserves y. For them, showing that it is non-instrumentally good that x suffers for a'ing or, likewise, that there is some *pro tanto* reason to promote x's suffering for a'ing, would just amount to showing that Desert is true. But others who are quite explicit that these are not intended as analyses of desert nevertheless agree that showing the truth of the consequents would provide *support* for a thesis like Desert.<sup>7</sup>

Putting these together with the idea that what the blameworthy deserve to suffer is to feel guilty, we get the following theses:

**Good-Guilt:** It is non-instrumentally good that one who is blameworthy feel guilty.

**Reason-Guilt:** There is a *pro tanto* reason to induce feelings of guilt in the blameworthy.<sup>8</sup>

In sections II and III, I will assess what has so far seemed to me the best case for and the best case against Good-Guilt.<sup>9</sup> While I believe that the case against Good-Guilt is quite strong, I concede that it might fail in the end. Nevertheless, as I explain in section IV, even if Good-Guilt

is true, the very reasons that best support it undermine the idea that it supports Desert-Guilt, as opposed to a related thesis that I call “Fittingness-Guilt”. I then explore the implications of this fact for what is at stake in the debate about Desert-Guilt. Finally, in section V, I turn directly to Reason-Guilt, and argue that we should take seriously the idea that it is not true—again, even *if* Good-Guilt is true.

I conclude by sketching a view that denies Reason-Guilt, and denying on the same grounds the more general thesis that being deserving by itself provides a *pro tanto* reason to give the blameworthy what they deserve. The view is not entirely negative, however. For as I argue, it is still the case that being blameworthy can, in some circumstances, become part of a reason to do so. Thus, something like Desert is correct. The view is robustly anti-skeptical in recognizing a coherent (and often-instantiated) conception of desert, and at the same time, it is modest in rejecting a popular implication of desert, namely, Reason-Guilt.

## II. The Case for the Non-Instrumental Goodness of Guilt: The Caring Argument

We find what I think is the most compelling argument for Good-Guilt in a variety of different sources, united by an appeal to the idea that caring about the right things is manifested by guilt under the right circumstances. Consider the following account of the role of guilt by Hilary Bok:

“If we care about living by our standards and about the state of our wills, then we will find the thought that we have failed to act as we think we should painful. In so doing...[we are] responding in the only appropriate way to a fact about our conduct.”

And

“My feelings of guilt make sense only on the assumption that I see my friend, her autonomy, and our friendship as important enough that my violation of them pains me...”<sup>10</sup>

Bok is not defending Desert here, or even Good-Guilt explicitly. But she provides the ingredients for an argument that is taken up more directly by Randolph Clarke:

“If one cares about others, if one respects them, and if one cares about whether one’s behavior is justifiable to them, then typically, one will feel bad, to some extent, when one realizes that one is guilty of acting in a way that is not so justifiable. Feeling guilty on such an occasion is an expression of one’s moral concern. It is a good thing that one have this concern, and one might think it good as well that one who is guilty have the emotional response that expresses this concern. Again, both of these things might be thought to be non-instrumentally good, good even if they bring about no further good.”<sup>11</sup>

The basic idea is that if we care about the right things—other people and doing right by them—then we will (typically) feel guilty when we’ve wronged them.<sup>12</sup> To feel guilt in such circumstances is a manifestation of caring, which is clearly a good thing, and the manifestation itself is a closely related good thing.

And Michael McKenna points to an instructive analogy between guilt and grief, arguing that while both are painful they are both essential manifestations of care, which is itself a non-instrumental good.<sup>13</sup>

...what constitutes the goodness of caring about the moral community or others' concern for her conduct involves her liability to the emotional harms that come with compromising these things. As for the comparison, I have suggested that the noninstrumental good found in the harm of blaming one who deserves blame is similar to the good found in experiencing appropriate grief in response to losing a loved one. The grief no doubt counts as a harm, but as an expression of one's love it also counts as a kind of noninstrumental good. How so? It is a better world that one grieves a loved one lost as in comparison with a world in which one experiences that loss and yet does not grieve; the grieving is itself an expression of the goodness of that love.<sup>14</sup>

While the theorists arguing for this connection between guilt on the one hand and caring on the other do not explicitly take up accounts of caring, they gain support from the most prominent among them. Consider, for example, what I here call the Emotional Vulnerability Account of caring<sup>15</sup>:

"...caring has an even more complex structure than most ordinary emotions—it is best understood as a structured compound of various less complex emotions, emotional predispositions, and also desires, unfolding reliably over time in response to relevant

circumstances. Typical components of caring include: joy and satisfaction when the object of one's care is flourishing and frustration over its misfortunes; anger at agents who heedlessly cause such misfortunes; ...fear when the object of care is in danger and relief when it escapes unharmed; grief at the loss of the object, and the subsequent nostalgia....Such component emotions, emotional predispositions, and desires all construe the same object—be it a person, an animal, or an ideal—as a source of importance commanding emotional vulnerability...”<sup>16</sup>

The central idea in the reasoning about guilt is that if we care about the right things, then suffering as a response to harm to the object of our care is simply a manifestation of such care; and, arguably, if we are the agent of said harm by way of our own blameworthy wrongdoing, guilt would seem to be a particular form of suffering that manifests caring of an important sort. As such a manifestation, it is itself non-instrumentally good.

It is notable here that there are several different objects of care mentioned in the relevant passages: another individual, a relationship, one's own standards, justifying oneself to another. But in each case, it seems that there is a worthy object of care, and guilt or a painful feeling of some similar sort is the manifestation of such care when one violates the relevant standards.<sup>17</sup>

Does it follow that *actual* suffering of guilt associated with caring is non-instrumentally good independently of the specific circumstances of wrongdoing? It appears that the answer is no, since from the premise that it is good to be *disposed* to X if one were in circumstances C it does not follow that it would be good to X in circumstances C. But it might still be the case that

in particular circumstances, such as one's having wronged another, the actual suffering of guilt is non-instrumentally good.<sup>18</sup> In other words, one might argue in the following way:

The Caring Argument: Under circumstances in which one culpably does wrong, not to suffer guilt would be bad insofar as, and because, it would be a *manifestation* that a component of caring is missing.<sup>19</sup> Put positively, to suffer would be good insofar as it would be a manifestation of the good of caring in the circumstances. And it is not (merely) an instrumental good, one that is good only insofar as it brings about a further good. Suffering is then an essential component of a non-instrumentally good response to circumstances in which one has done wrong.

Does the Caring-Guilt Argument succeed? The answer I will defend is that it either fails in an interesting way that could prompt us to revisit the nature of caring, or it succeeds, but too well for the ultimate aim of vindicating Desert-Guilt. But first, turn to the case against Good-Guilt.

### III. The Case against

A natural starting point is to present candidates for counterexamples. Gilbert Harman presents himself as one such case, along with others whom he knows:

“There seem actually to be many moral people with moral principles but no susceptibility to non-trivial guilt feelings. To mention one example, as far as I can tell, I am not

susceptible to non-trivial guilt feelings, yet I have moral principles and seem (at least to myself) to be a relatively moral person.”

And he goes on to argue that

“Non-trivial guilt is of course a negative experience that can make people miserable. It might be worth paying this price if susceptibility to guilt made people act better. But there is no evidence that susceptibility to nontrivial guilt is needed to make people act morally.”<sup>20</sup> (Gilbert Harman 2009).

Now, we would need more detail to see the case as a counterexample to the emotional vulnerability account of caring as applied to guilt. (Harman says he is a moral person, but doesn't refer to caring in particular.) Further, Harman's own focus—at least in the quoted passage above—seems to be about instrumental good, so, again, we don't have a case with enough detail to assess here as a potential direct counterexample to the non-instrumental goodness of guilt. Is some non-instrumental good lost if we could be as Harman claims himself to be? The *implicit* suggestion seems to be “no”. But again, it would be helpful to have more detail about the case.

Consider, then, a thought experiment, which gives us the ability to stipulate some of the conditions that are left incomplete in Harman's self-description. Compare our world to a world in which otherwise very similar creatures (call them the “No Guilt Creatures”) do not feel guilt, but possess a syndrome of dispositions, desires, and cares similar to ours. So when they freely

and culpably act wrongly, they wish that they had not acted in that way, resolve not to act that way in the future, invest time and effort working on making the resolution meaningful, and so on. Is this world better than ours? Worse? Equally good? But before we can answer these questions, we need to make sure the thought experiment is coherent.<sup>21</sup>

We could have a number of different reactions to the thought experiment's claim to coherence:

- (1) The thought experiment isn't coherent. If you don't feel guilt or suffer in some closely related way in such circumstances, it isn't true that you care in the right way.
- (2) The thought experiment is coherent, but couldn't be true of (most) human beings. So the disposition to guilt and suffering is necessary after all for the right kind of caring insofar as we are human. (Or a variant: it is coherent if we replace "caring" with "creature-caring" and take it that creature-caring is no less valuable than caring.)
- (3) The thought experiment is coherent, but couldn't be true of human beings as we are, with the histories we have, and so on. So the disposition to guilt and suffering is necessary after all for the right kind of caring insofar as we are human beings with our actual histories.<sup>22</sup>
- (4) The thought experiment is perfectly coherent and humans could shed their dispositions to guilt without shedding the related dispositions.

If either (2), (3), or (4) is correct, then it is plausible that the hypothetical world is better for including less suffering, and retaining the goods that the suffering of guilt was initially supposed

essential for. If (2) or (3), then at least guilt would seem to be a perhaps unfortunate aspect of being human or of being human in our circumstances, but not essential to what is good about caring about the victims of one's wrongdoing, or about doing the right thing. I confess that in some moments, I think that the thought experiment is coherent, and in some not. If it is, then it suggests a revisiting of the Emotional Vulnerability Account of caring described earlier. In that case, emotional vulnerability to painful experiences is merely characteristic of, but neither constitutive of nor necessary to, caring.

Consistent with any of (2), (3), or (4) is the idea that guilt is excellent *evidence* of caring about the right things. In fact, it is such good evidence that we might confuse it with an essential manifestation of it. And this might explain why (1) might *appear* true. I believe that the thought experiment, if coherent, provides support for the idea that the disposition to guilt is not itself an essential aspect of caring about the right things as it seems that the argument requires. And it goes some way to undermining the idea that guilt is non-instrumentally good.

It does not go all the way because consistent with the idea that the world of the No Guilt Creatures is a better world than ours, all things considered, is the idea that it is nevertheless missing a non-instrumental good that we have in our world. Perhaps guilt is non-instrumentally good in one way, even though overall it is a better world in which guilt does not exist. At the same time, even if this were the case, the significance of Good-Guilt would be reduced. The reason is that any claimed non-instrumental goodness of guilt would seem to be outweighed by the overall badness of guilt, which it presumably possesses in virtue of its painfulness. Thus, even if we recognize that something good is missing in the No Guilt Creatures world, as long as we find their world better overall, then we must concede that there can be no net positive

amount of goodness added by promoting guilt feelings to a world that was otherwise the same as ours in all relevant respects, including caring and manifestations of it in action.

In the end, while I take there to be good reason to reject Good-Guilt on the strength of this thought experiment, I won't assume that it is false in the next two sections. The question for us now is whether it follows from Good-Guilt that we have thereby supported Desert-Guilt. Not necessarily. I turn to this question in the next section. And here I think the answer is no. Of course, Good-Guilt is not intended as an analysis of Desert-Guilt in the first place. But an axiological claim of this sort is often used to explicate a claim about desert, is sometimes itself taken to help us understand a "desert thesis", and is thought to support something like Desert.<sup>23</sup> In the next section, I will argue that Good-Guilt—particularly when supported (exclusively) by the Caring Argument—not only does not explicate Desert-Guilt; it goes some way to undermining it.

#### IV. **Goodness, Fittingness, and Desert: An Argument Against Good-Guilt as a Defense of Desert-Guilt**

Suppose, then, that emotional vulnerability is essential to caring. Even in that case, it is essential to note that the argument that guilt is a non-instrumental good depends ultimately on the fact that it is related (non-instrumentally) to the good of caring. But guilt is good then *not* because it is painful or negatively-valenced, but rather because of its connection to caring about the right things. The fact that guilt is painful or negatively-valenced does not then explain why it is on the candidate list of things deserved in virtue of one's being blameworthy.

And this casts some doubt on the idea that guilt is *deserved* even if there is something good about it.

Analogously, when it comes to grief, grief is good insofar as it is a manifestation of caring, but not in virtue of its being painful. And the problem for the project of supporting Desert-Guilt by means of the Caring Argument for Good-Guilt is that the value of guilt is *too much* like that of grief in this case. If its value does not lie in the blameworthy getting something negative in response to wrongdoing, but rather in its connection to caring, then it seems not to capture the idea of desert, as opposed to a more general kind of fittingness that applies equally to grief and guilt. To see this in another way, we wouldn't typically say that grief is deserved, and though I don't want to put too much stock in ordinary ways of talking, it seems that the concept of desert doesn't apply to grief in this case. Derk Pereboom argues that a painful kind of regret can be appropriate as a response to harming others, but that this, too, is not deserved.<sup>24</sup> Perhaps the same can be said of guilt even in response to culpable wrongdoing. We might say, then, that Good-Guilt better supports

***Fittingness-Guilt:*** It is fitting for the blameworthy to feel guilty to the right degree.

Now I think that the defender of Desert-Guilt has three kinds of response. The first is to deny the claim that when it comes to guilt—and grief for that matter—it is possible to separate out the painfulness and the relationship to caring. The second is to find a way of distinguishing guilt from grief so that the former but not the latter is deserved in a robust way that sets it aside from mere fittingness. (Another way to put this is to say that desert is a special *kind of*

fittingness that applies to guilt but not to grief.) The third kind of response is more concessionary, and it is to say that even if guilt is merely a fitting response to blameworthiness, fittingness can do the important *work* of desert.

The first response denies the starting point of the reasoning, and suggests that there is really no sense to be made of the claim that guilt is *good in virtue of* its connection to caring, while it is not *good in virtue of* its painfulness. McKenna, for example, argues that both guilt and grief are best thought of as “an organic unity whose whole has a value that is not analytically decomposable...”<sup>25</sup> There is something appealing in this idea. But I think I can make sense of the idea that even in cases in which painfulness is a constitutive part of an experience that is non-instrumentally good, it is not so in virtue of the painfulness. Take, for example, an empathetic response. It may be a non-instrumentally good thing for a doctor to have an empathetic response to the pain of her patient, but its goodness is arguably not in virtue of its painfulness, but in virtue of her richer understanding of his experience and as a manifestation of her willingness to treat him as a person and not as a symptom. Further, even if this turns out to be the wrong way to think about these things, and one cannot separate out the painfulness and the other features that make the experience good in the case of guilt, it is still the case that the guilt seems structurally quite similar to grief on the assumptions of the caring argument. Both are good insofar as they are manifestations of caring about the right things. So there remains a question of what makes guilt *deserved* while grief is fitting, but not *deserved*. Thus, the proponent of the first response might in the end need to appeal to additional resources that can be found in either the second or third response.

The second response is to distinguish guilt from grief in a relevant way. A natural first idea is to say that while it is not the case that grief is good in virtue of its being painful (but perhaps despite it), guilt is good in virtue of its being painful. But this is precisely what the Caring Argument places in doubt. If the good of guilt is really its connection to caring, then it appears to be in exactly the same situation as grief.

An alternative attempt to distinguish guilt from grief is to say, as Andreas Carlsson does, that guilt is “morally good” in contrast to grief, in that it gives others reasons to act so as to induce it, while grief does not.<sup>26</sup> And it might be argued that what I have called Reason-Guilt has conceptual priority, and itself supports Good-Guilt.<sup>27</sup> If Reason-Guilt is true, and there is no analog of Reason-Guilt for grief, then this would be a way of distinguishing between guilt and grief. It is worth noting that to do so, however, would be to offer a very different route to the defense of Good-Guilt than that of the Caring argument. Still, it will be important to take up the question of whether Reason-Guilt is true, and I will turn to this task in the next section.

But before doing that, it is important to consider the more concessionary response to the challenge that guilt and grief are too similar in respect of their non-instrumental goodness for Good-Guilt to support Desert-Guilt. According to that response, we don’t need a special *kind* of fittingness to do the work in question; we just need it that it is guilt that is fitting. Fittingness is sufficient to give us reason to treat people in guilt-inducing ways, say, and to still allow for a non-skeptical position about blameworthiness if it is true that it is (sometimes) fitting to induce guilt. Things get quite subtle here, and I will not be able to do this issue justice. I will just make one observation now, and leave somewhat open how much, if anything, is lost by moving from Desert-Guilt to Fittingness-Guilt.

There is what might seem a reason to resist the concession to abandon Desert-Guilt in favor of Fittingness-Guilt. It seems even less plausible that mere fittingness by itself yields reasons for action than it does that desert yields such reasons. As I have argued elsewhere, fittingness alone does not yield such reasons.<sup>28</sup> For example, the fact that a racist joke is funny seems to entail a certain fittingness of laughing. But such fittingness needn't provide any reason at all, it seems to me, not even a *pro tanto* one that is overridden. Perhaps it could be claimed that it is not the overt response, namely, laughing, that is fitting, but the emotion or feeling of amusement, and, thus, that one has reason to feel amusement insofar as it contains a true appraisal of the situation. I am not sure whether this is correct, for it might be that one simply has reason to *judge* the joke funny without having reason to feel amusement. But even if it were correct, this move would sever reasons for feeling from the reasons for acting in ways the feeling disposes one to act. This seems a potentially unstable position.

Perhaps another case is an even clearer one for casting some doubt on Fittingness-Guilt. It might be fitting to respond to something with surprise, but it isn't clear what it would mean for one to have reason to bring about one's own response of surprise. For this reason, it might seem that a special kind of fittingness is important in order to obtain a direct connection to reasons.

To sum up this section, then, we have arrived at the conclusion that feelings of guilt might be fitting, and even non-instrumentally good, but the very argument that we have seen support its non-instrumental goodness either fails to support, or counts against, the claim that it is deserved in a sense distinct from fittingness.

## V. Desert, Goodness, and Reasons

Now let us turn briefly to Reason-Guilt. It is important in its own right, but also, because, as we saw earlier, at least some have claimed that Reason-Guilt is the more fundamental claim, which in turn supports Good-Guilt. It seems to me this is even less intuitively plausible than Good-Guilt, particularly if we are not able to appeal to Good-Guilt in order to support it. A thought experiment might best bring this out.

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

I do not have the intuition here that you would be making this kind of mistake. That suggests to me that there is not a *pro tanto* reason to induce guilt that stems from blameworthy action alone. For by hypothesis, there were no *overriding* reasons in favor of

passing (no cost, no effort, and so on). So if there is indeed such a *pro tanto* reason (and of a sort that involves justice and that isn't overridden), it seems that you would be making a *kind* of mistake. Thus, I tentatively conclude that Reason-Guilt is false.<sup>29</sup>

Of course, in many actual cases, guilt plays a valuable instrumental role (and we can see this from the human developmental literature). And, again, guilt is an excellent piece of evidence for the kind of genuine resolve I stipulated in the thought experiment. But the idea behind the thought experiment is to strip away the existence of typical instrumental goods and to provide alternative sorts of evidence by stipulation.

Now one might object to the use of the case to draw a general conclusion about reasons. One might do so by appealing to a distinction between two roles for reasons: requiring and justifying.<sup>30</sup> According to Joshua Gert, a reason requires the performance of an action when it would be irrational not to perform it. In contrast, a reason justifies an action by making it not irrational not to perform it. In turn, corresponding to each role is a distinct measure of strength: justifying strength and requiring strength. As Gert elaborates, given two reasons, R1 and R2, R1 has more *requiring* strength if and only if R1 would make it irrational to do anything R2 would make it irrational to do and R1 would make some things irrational to do that R2 would not<sup>31</sup> (2003, pp. 15-16). On the other hand, given two reasons R1 and R2, R1 has more justifying strength if and only if R1 would make it rationally permissible to do anything that R2 would make it rationally permissible to do, and R1 would make some things rationally permissible to do that R2 would not make it rationally permissible to do. An illustration will help bring out the distinction. A reason might have high justifying strength without having high requiring strength. To set out an adapted example of Gert's, the fact that you could save forty

children from a burning orphanage might have high justifying strength (justifying even your risking your life to do so), while it might have relatively lower requiring strength, so that you are not rationally required to do it, even if the cost would be substantially lower than risking your life, such as risking an index finger.<sup>32</sup> Applying the distinction to “The Look”, one might say that while there is a reason with justifying strength for giving the look, giving the look lacks a reason of requiring strength and this fact entails that there is nothing irrational about not giving the look even though the opportunity to give the look is reason-giving in the sense of offering justification.

While the proposed distinction is an interesting one, I think that the objection can be answered without engaging in an evaluation of the view according to which there are two kinds of strength that come apart. For it is notable that even if dimensions of requiring and justifying come apart in cases such as the orphanage, nothing is implied about the special case in which one has (by stipulation) only one reason on one side and none at all on the other. In this case, to claim that there is nothing irrational about the choice not to give the look, despite the existence of a reason with justifying strength on the side of giving the look, the following claim would also have to be true: the requiring strength of the opportunity to give the look is not simply small, but, rather, it is *zero*; that is, it plays absolutely no role in the calculation of what one *ought* to do. The claim that there is justifying but not requiring strength in the reason at hand is doubtful, however, and, were it true that giving the look has zero requiring strength, this fact would considerably weaken Reason-Guilt. To explain: first, we would have to have reason to think that the requiring strength of reasons can ever be zero, let alone zero in this case. In the motivating examples, the requiring strength is non-zero. If one could save the

forty orphans at no cost at all, for example, it is hard to see that one would not be rationally required to do so.<sup>33</sup> It is plausible that if there is *no* reason against a course of action, and taking the action is supported by reasons, one would indeed be irrational in not taking that action. At the least, without a rationale for thinking that the requiring strength of reasons can be zero, it seems that there is no reason to doubt the lesson of this thought experiment.

A second important concern about deploying this strategy is that even if the claim could be defended, it would considerably weaken Reason-Guilt insofar as it would no longer offer guidance about what we ought to do. While it is true that Reason-Guilt could still play a justifying role of our treatment of others on this view, it could not play a guiding role of any kind, and this seems not only to countenance an odd (if not incoherent) kind of reason, but also to detract from at least one significant aspect of the thesis.

One additional consideration might be thought to allow the deployment of the distinction between justifying and requiring reasons without committing to a null requiring strength for giving the look. It is to point to the fact that guilt itself is painful and so there *is*, contrary to the explicit supposition of the thought experiment, a reason against giving the look. Insofar as the experience of guilt is painful and all suffering is bad, the experience of guilt is itself bad and offers reason not to bring it about; insofar as it is a manifestation of caring, it is good and offers reason to bring it about. There is certainly something to the idea that suffering is, in general, bad. But it is important that for the thought experiment to allow for rational permission to pass on giving the look, these reasons—the suffering of guilt and the rightness of guilt—would seem to have to precisely cancel each other out on the dimension of requiring strength. And if they always cancel each other out, while the painfulness of guilt is inextricable

from guilt, then bringing about guilt in wrongdoers would never be *net*-reason requiring. Thus, this move does not ultimately help in undermining the conclusion from the thought experiment that bringing about guilt is not reason-giving in a non-instrumental way.

To this point, I have offered a negative argument, suggesting that Reason-Guilt is false, and that there is not even a *pro tanto* reason to promote the experience of guilt in the blameworthy simply because they are blameworthy. Here I want to highlight one important implication of this conclusion, and one important limitation of it. First, the implication. If we accept the general thesis that the non-instrumental goodness of a state of affairs itself provides a *pro tanto* reason to promote the experience of guilt, then the absence of such a *pro tanto* reason would entail the *lack* of (net) non-instrumental goodness of such an experience. In that case, the thought experiment would serve to cast doubt on the claim of non-instrumental goodness (or at least the idea that there is a net amount of such goodness), independently of the details of positive arguments, such as the Caring Argument, in favor of Good-Guilt. Now the general entailment between goodness and *pro tanto* reasons might be questioned. And it is important to note that Good-Guilt does not include an explicit claim to net-positive goodness; so in principle, one could still accept the Caring Argument with its conclusion of Good-Guilt while rejecting the idea that guilt ever provides a net positive good. Still, it seems that at least some of the force of Good-Guilt would be undermined if this were the case. Thus, the conclusion of the argument based on The Look is significant, casting doubt on the idea that there can ever be a net non-instrumental good in experiencing guilt, as well as on Reason-Guilt.

Now for the limitation of the argument based on The Look. It does not follow that there is no connection at all between desert—even desert of guilt—and reasons. In fact, it is

plausible that desert offers a certain kind of *conditional* reason for bringing it about that someone gets what she deserves. Under certain conditions, the fact that one is deserving of something on account of one's blameworthy wrongdoing can make it the case that it would be better than the alternatives that one get what one deserve. And under certain conditions, it can also give one, as well as others, reason to bring about that state of affairs.<sup>34</sup> Let us abstract for a moment from the idea that what the blameworthy deserve is to feel guilty. The basic idea can be captured as followed:

***Conditional Reason-Desert:*** If x deserves y for a'ing, then there is a conditional reason to promote x getting y for a'ing.

Suppose, for example, that you are in a position in which you have no choice but to promote someone's being harmed, and you can harm someone deserving of it or harm someone else. This may give you a reason to promote the person who is deserving getting what she deserves. In such circumstances, it is not unfair to impose harm on the deserving whereas it would be to the undeserving. Nevertheless, in other circumstances, the fact that someone is deserving may provide no reason at all. Thus, there is a systematic connection between desert and liability to harm, even if it is not in place in all circumstances in which one is deserving of harm.<sup>35</sup>

Now let us offer a specific version of the thesis in terms of guilt as what is deserved:

***Conditional Reason-Guilt:*** There is a conditional reason to induce feelings of guilt in the blameworthy.

Conditional Reason-Guilt is consistent with the rejection of both Good-Guilt and Reason-Guilt. Importantly, if it is true, it does not gain support from Good-Guilt (or the Caring Argument in favor of it). It is also not clear that it is more plausible than the more general Conditional Reason-Desert thesis, and a full defense would require a special reason for restricting the thesis to a systematic relationship between blameworthiness and liability to the promotion of guilt in particular. But it is not undermined by the argument based on The Look, and in my view remains a thesis well worth further consideration.

## **VI. Conclusion**

In this paper, I have explored two particular claims that have been thought to explicate and support Desert-Guilt. I have presented some arguments against Good-Guilt that seem to me to be strong; but I have also conceded that it might be true, while maintaining that even in that case the best argument for it—the Caring Argument—is one that undermines its ability to support Desert-Guilt in particular. Some have argued instead that Desert-Guilt is best supported by Reason-Guilt, but it seems to me that this gets things the wrong way around at best, and in any case I have offered some reason to reject Reason-Guilt, at least as a fundamental principle. At the same time, I have argued that for all of the concerns about Good-Guilt and Reason-Guilt, Conditional Reason-Guilt could still be true and revealing of an important aspect of desert. Finally, there is no doubt much more to say about these theses,

their relation to Desert-Guilt, and about the more general claim of Desert which is itself untouched by anything I have argued here.

---

<sup>1</sup> I am indebted to Michael McKenna for editing this journal issue and organizing the associated conference, *Responsibility: The Next Generation*, at the University of Arizona. As a result, I received invaluable written comments from both Michael and Randy Clarke whose work helped spark my own thinking on the topic, very challenging and insightful feedback from Ellie Mason and David Shoemaker as first-questioners, and excellent questions from all of the participants. I am also very grateful to a number of people for comments and discussion that greatly improved the paper, including audiences at the *Summer Program for Women in Philosophy*, UC San Diego, June 2016, the *UC Riverside Graduate Student Conference on Agency*, May 2016, the *Moral and Political Philosophy Seminar*, UC San Diego, April 2017, the University of Mainz, June 2017, and the *Workshop on Responsibility* at the Gothenburg Responsibility Project, June 2017. For very helpful comments and discussion on the paper and related topics, special thanks to Julia Annas, Richard Arneson, Christopher Bennett, Gunnar Bjornnson, David Brink, Andreas Brekke Carlsson, Kathleen Connelly, Taylor Cyr, Cory Davia, Elizabeth Harman, Bob Hartman, Agnieszka Jaworska, Kathryn Joyce, Robert Kane, Keith Lehrer, Ben Mathieson, Joe Metz, Kristen Mickelson, Per-Erik Milam, Fred Miller, Derk Pereboom, Sam Rickless, Connie Rosati, Carolina Sartorio, Jeffrey Seidman, Houston Smit, David Schmidtz, Manuel Vargas, Steve Wall, Eric Watkins, and Monique Wonderly.

<sup>2</sup> See, for example, Shoemaker, David *Responsibility From the Margins* (Oxford: Oxford University Press, 2015).

---

<sup>3</sup> For this formulation, see Randolph Clarke, “Moral Responsibility, Guilt, and Retributivism,” *Journal of Ethics* 20 (2016): 121-137.

<sup>4</sup> For example, Randolph Clarke “Some Theses on Desert,” *Philosophical Explorations* 16 (2013): 153-64, Randolph Clarke, “Moral Responsibility, Guilt, and Retributivism,” *Journal of Ethics* 20 (2016): 121-137, Michael McKenna, “Punishment and the Value of Deserved Suffering” (in preparation), Michael McKenna *Conversation and Responsibility* (Oxford: Oxford University Press, 2012), and Andreas Brekke Carlsson, “Blameworthiness as Deserved Guilt,” *Journal of Ethics* 21 (2017): 89-115. For something like this idea applied in the context of criminal punishment, see Antony Duff (“Restoration and Retribution,” *Restorative Justice and Criminal Justice: Competing or Reconcilable Paradigms*, edited by Andreas von Hirsch, Julian Roberts, Anthony E Bottoms, Kent Roach, Mara Schiff (Oxford: Hart Publishing, 2003): “Talk of retribution conjures up in many minds the image of a vindictive attempt to inflict hardship—to ‘deliver pain’ ...--‘for its own sake’; and who could argue in favour of that? I will argue, however, that the retributivist slogan—that ‘the guilty deserve to suffer’—does express an important moral truth; and that in the case of the criminally guilty it is the state’s proper task to seek to ensure that they suffer as they deserve...First, he deserves to suffer remorse: he should come to recognize and repent the wrong that he did—which is necessarily a painful process. Second, he deserves to suffer censure from others...; this, too, if taken seriously must be painful. (p. 182).

<sup>5</sup> McKenna, Michael, “Punishment and the Value of Deserved Suffering” (in preparation).

<sup>6</sup> Sometimes this thesis is taken to be inspired by Kant: “Now inasmuch as virtue and happiness together constitute possession of the highest good in a person and happiness distributed in exact proportion to morality (as the worth of a person and his worthiness to be happy) constitutes the

---

*highest good* of a possible world, the latter means the whole, the complete good, in which, however, virtue as the complete condition is always the supreme good, since it has no further condition above it, whereas happiness is something that ... is not of itself absolutely and in all respects good, but always presupposes morally lawful conduct as its condition." (Immanuel Kant *Critique of Practical Reason*, edited by Mary Gregor (Cambridge: Cambridge University Press, 1788/1997), 5:111).

<sup>7</sup> See, e.g., McKenna, Michael (2012) *Conversation and Responsibility* (Oxford: Oxford University Press). McKenna takes something like the combination of the two principles to be a "desert thesis." More recently, McKenna describes similar principles as "an indirect way of defending a basic desert thesis for blameworthiness" (Michael McKenna "Defending conversation and responsibility: reply to Dana Nelkin and Holly Smith," *Philosophical Studies* 171 (2014): 73-84, p. 82) and "a way to understand a basic desert thesis" (Michael McKenna "Basically Deserved Blame and Its Value" (in preparation), p. 11). Christopher Bennett offers a defense of the goodness of the suffering of guilt as part of an overall defense of retributivism. For Bennett, the pain of guilt is one of at least three varieties of suffering that make up the retributive experience, which is non-contingently good. Explaining the particular varieties of suffering that are deserved in this way is thus part of the larger case for retribution, as he sees it. (See Bennett, Christopher "The Varieties of Retributive Experience," *The Philosophical Quarterly* 52 (2002): 145-163.)

<sup>8</sup> Randy Clarke has suggested (in correspondence) that once we have focused on guilt as the thing that is fundamentally deserved, we should instead focus on reasons to *feel*, rather than on reasons to induce feelings as in Reasons Guilt. I believe that we can do both, and indeed

---

that they are related. One reason for continuing to focus on Reason-Guilt is that it is an instantiation of Reason-Desert, while a related thesis about reasons to feel would not be. Since Reason-Desert has been taken to support Desert, and since reasons to feel do not have as direct a connection to the justification of overt blame and punishment, both of which involve actions, it seems important in the context of an exploration of alternatives to traditional retributive views to consider Reason-Guilt in particular. Further, it seems that the kinds of reasons in having reasons to feel are related to having reasons to bring about a feeling. While I am skeptical that there is a straightforward inference from the premise that one has reasons to feel E in a given situation to the conclusion that there is some *pro tanto* reason to promote one's feeling E, in many situations it will be the case that the reasons for feeling something will also serve as reasons to promote one's feeling it.

<sup>9</sup> In previous work, I argued against the more general claim of Good-Desert. Those arguments apply to the more specific Good-Guilt. However, there are special arguments *in favor* of Good-Guilt in particular that need addressing on their own terms, and I acknowledge that there is a special appeal of Good-Guilt such that the best argument for the more general claim, it seems to me, would be an argument in favor of Good-Guilt. (See Dana Kay Nelkin, "Desert, Fairness, and Resentment," *Philosophical Explorations* 16 (2003): 117-132.)

<sup>10</sup> Bok, Hilary *Freedom and Responsibility* (Princeton: Princeton University Press, 1998), pp. 174-75.

<sup>11</sup> Randolph Clarke, "Some Theses on Desert," *Philosophical Explorations* 16: (2013) 153-64, p. 156.

---

<sup>12</sup> See also David Shoemaker, who, while not defending Good-Guilt directly, also connects guilt and caring. In the context of rebutting a Kantian claim about the distinctness between motivation and cares, Shoemaker says specifically that “it is constitutive of guilt that one cares about the ideals or values one has betrayed, however reluctantly” (Shoemaker, David, “Caring, Identification, and Agency,” *Ethics* 114 (2003): 88-118, p. 99).

<sup>13</sup> See Dan Moller (“Love and Death,” *Journal of Philosophy* 104 (2007): 301-16) for a nuanced view that something important is regrettably lost when people quickly cease to grieve the death of those close to them, whom McKenna approvingly cites. But see Monique Wonderly (“On Being Attached,” *Philosophical Studies* 173 (2016): 223-42 for a compelling argument that grief is a manifestation, not of (mere) care, but of security-based attachment. Consistent with this conclusion is that care constitutively involves emotional vulnerability of some sort; just not (by itself) the sort associated with grief.

<sup>14</sup> Michael McKenna “Punishment and the Value of Deserved Suffering” (in preparation).

<sup>15</sup> This label is an oversimplification since there is more than emotional vulnerability to caring. But it captures a central element, and one that is especially important for our purposes.

<sup>16</sup> Agnieszka Jaworska “Caring and Internality,” *Philosophy and Phenomenological Research* 74 (2007): 529-68, p. 560.

See also Harry Frankfurt: “A person who cares about something is, as it were, invested in it. He *identifies* himself with what he cares about in the sense that he makes himself vulnerable to losses and susceptible to benefits depending on whether what he cares about is diminished or enhanced” (“The Importance of What We Care About,” *Synthese* 53 (1982): 257-272, p. 260).

---

<sup>17</sup> Note that one might care about some of these objects without caring about others, though they very often go together. Very young children might care about another person without having the capacity to care about not wronging them. In that case, they would be emotionally vulnerable to a state of suffering, but not, perhaps, to guilt in particular. Other such cases might include certain forms of dementia. It appears that one could care about a person without caring about not wronging them if one lacked the conceptual resources to do so. At the same time, it is an interesting question whether someone with the *capacity* to care about wrongdoing could care about other individuals without caring about not wronging them.

<sup>18</sup> At first glance, this account of caring, together with the plausible claim that caring is an important non-instrumental good, might seem to offer a kind of solution to the problem of evil. But precisely because one can care—and so have that good—without its manifesting in the particular way of suffering, this will not ultimately contribute even partially to a solution.

<sup>19</sup> This might need some qualification. Notably, Clarke's claim is that when we have done wrong and we care properly, we "typically" respond with guilt. And this qualification seems apt for the reason that we might have a disposition to suffer guilt, but we are too depressed or exhausted or distracted in the circumstances, say, to actually feel guilty. In such a case, the disposition essential to caring would be masked, but not missing. Still, we might say that it is a good thing when such dispositions *are* manifested, and I take it that this would be what is technically needed for the argument to go through.

<sup>20</sup> Gilbert Harman, "Guilt-Free Morality," *Oxford Studies in Metaethics* IV: (2009) 203-214.

---

<sup>21</sup> It is worth noting a thought experiment in many ways parallel to this one concerning a set of creatures who do not experience grief at the loss of loved ones, presented by Dan Moller (“Love and Death,” *Journal of Philosophy* 104 (2007): 301-16).

<sup>22</sup> See Michael McKenna for a defense of the idea that what is essential to get clear on is not what is conceptually possible, but what is possible given life as we know it (*Conversation and Responsibility* (Oxford: Oxford University Press, 2012), pp. 110—12).

<sup>23</sup> See, e.g., *ibid.* and Michael McKenna “Basically Deserved Blame and Its Value” (in preparation).

<sup>24</sup> Derk Pereboom (2017) “Responsibility, Regret, and Protest,” *Oxford Studies in Agency and Responsibility* (Oxford: Oxford University Press).

<sup>25</sup> Michael McKenna “Basically Deserved Blame and Its Value” (in preparation).

<sup>26</sup> Carlsson, Andreas Brekke, “Blameworthiness as Deserved Guilt,” *Journal of Ethics* 21 (2017): 89-115.

<sup>27</sup> Clarke suggests that something similar to Reason-Guilt, namely, the claim that we have reason to *feel* guilt has conceptual priority over its being good to feel guilt, where the reason in question is a reason of justice (Randolphe Clarke “Some Theses on Desert,” *Philosophical Explorations* 16 (2013): 153-64).

<sup>28</sup> Nelkin, Dana Kay, “Accountability and Desert,” *Journal of Ethics* 20 (2016): 173-89.

<sup>29</sup> The Look is in some ways similar to Kant’s case of the last murderers in a state that is dissolving and where no further good effects can be delivered by executing them. (See Immanuel Kant, *Metaphysics of Morals* 6:33.) Kant famously answers that they ought to be executed, and that this would require a parallel response in the Look, namely, that one ought to give the Look. Note, first, that Kant offers his verdict as an implication of the conclusion of a

---

subtle argument; it is not presented as a thought experiment. Second, Kant's case concerns state punishment, and while I believe that there are important connections between Desert and the justification of punishment, questions about whether one has reasons to promote the experience of guilt and whether the state has reasons to punish (and, more specifically, punish with death) are quite distinct. For these reasons, I set aside Kant's case here, while noting the interesting parallels.

<sup>30</sup> See Joshua Gert, "Requiring and Justifying: Two Dimensions of Normative Strength," *Erkenntnis* 59 (2003): 5-36.

<sup>31</sup> *Ibid.*, pp. 15-16.

<sup>32</sup> This is controversial. But the illustration (or one like it) supports the objection to my use of the thought experiment, and my response to the objection does not require settling this controversy here.

<sup>33</sup> Notably, Gert shares this assessment; in his view, there is a non-zero value of requiring strength in the orphan case. And more importantly, he explicitly recognizes that his argument does not have implications for the sort of case at hand in which there are reasons on one side and none on the other. See Joshua Gert, "Requiring and Justifying: Two Dimensions of Normative Strength," *Erkenntnis* 59 (2003): 5-36, p. 7.

<sup>34</sup> See Dana Kay Nelkin, "Desert, Fairness, and Resentment," *Philosophical Explorations* 16 (2013): 117-132 and "Accountability and Desert," *Journal of Ethics* 20 (2016): 173-89.

for development of this view, which shares some features with the view that Feinberg calls "Fault Forfeits First" (Joel Feinberg, "Sua Culpa," in *Doing and Deserving: Essays in the Theory of Responsibility* (Princeton: Princeton University Press, 1970: 187-221). See also Arneson,

---

Richard “Resolving the Responsibility Dilemma,” In *The Ethics of War: Essays*, eds. Saba Bazargan-Forward and Samuel C. Rickless. Oxford: Oxford University Press (2017), and “Just Warfare Theory and Non-Combatant Immunity”. *Cornell International Law Journal* 39 (2006): 663–688.

<sup>35</sup> Randy Clarke (in correspondence) has suggested a possible connection between reasons and desert that appeals to a distinction between what he calls “reasons that permit” and “reasons that require”. An example of the first is given by my giving you permission to do a thing. I thereby acquire a reason favoring doing it, despite not being required to do it. In the case of The Look, there might be reasons that permit, but non-requiring reasons. And such reasons to permit are such as could enter into deliberation about what one should and should not do. This is an interesting suggestion, and it might in fact be compatible with the suggestion in the text, depending on how exactly we are to understand a “permissive” reason. As I understand the suggestion, the permissive reason is not a positively weighted reason to perform an action, but instead merely “cancels” a negative reason that might otherwise be in place. So it is not a *pro tanto* reason. It seems a bit odd to my ear to call this a reason favoring the action. But insofar as it cancels a reason against a certain action, then I believe I can accept the suggestion as consistent with the proposal in the text.