Deserved Guilt and Blameworthiness over Time

Andreas Brekke Carlsson, Inland Norway University of Applied Sciences


1. Introduction

The literature on moral responsibility is ripe with accounts of what it takes for an agent to become blameworthy. Typically, a responsibility theorist will say that the agent must act wrongly and display an objectionable quality of will. Moreover, she must have generated this action, attitude, or omission knowingly, voluntary, and under control. When this happens, she is worthy of blame. But for how long? Very little is written about what it takes for an agent’s blameworthiness to cease or diminish.¹

One reason for the lack of interest in this question might be that most philosophers share the assumption that blameworthiness is forever: Once you’re blameworthy for something, you’re always blameworthy for that thing.² But it is by no means obvious why this should be the case. Most philosophers working on moral responsibility also accept that an agent is blameworthy for something if and only if it is appropriate to blame her for that thing. This generic biconditional entails that if blaming S for x is no longer appropriate, then S is no longer blameworthy for x. Suppose that blame is equated with the reactive attitudes resentment, indignation, and guilt. The question of whether an agent remains blameworthy then boils down to the question of whether resentment, indignation and guilt remain appropriate. Once we frame the question in this way, there seem to be reasons to doubt that blameworthiness is forever. The diminution of blame, towards oneself, as well as towards other people, is a familiar phenomenon. The intense and painful guilt we might feel in the hours, days and weeks after we have done something wrong rarely persist. Similarly, resentment or indignation can change and disappear in time. Of course, many of the reasons for why blame diminishes seem to have little to do with the agent’s blameworthiness. It would be puzzling if the mere passage of time could

---

¹ Notable exceptions are Tognazzini (2010), Coleman and Sarch (2012), Khoury and Matheson (2018) and Clarke (manuscript). Shoemaker (forthcoming) provides an interesting and relevant discussion, although the paper is framed in terms of forgiveness rather than blameworthiness. Callard (2018) and Marusic (2020) provide interesting discussions of the rationality of anger over time, without linking this discussion to blameworthiness.

² Randolph Clarke (manuscript) call this “the standard view.”
render an agent less blameworthy for what he has done. A victim might decide to let go of blame because he needed to move on with his life. Such prudential reasons might outweigh other reasons the victim has for continuing to blame the wrongdoer. But it would not make the agent any less blameworthy.

However, there seems to be certain things a victim or a wrongdoer can feel or do that might change the wrongdoer’s status as blameworthy. A wrongdoer might experience guilt, atone, apologize and make reparations. A victim might forgive. Such actions and emotions seem to be reasons for why the victim should stop feeling resentment, or at least for resentment to diminish and for why the wrongdoer should stop feeling guilt for their action, or at least for guilt to diminish. Moreover, these reasons seem relevant to the agent’s blameworthiness, and not merely to the overall justification of continued blame.

My aim in this paper is to answer the question of how and why blameworthiness can cease or diminish. In order to investigate this question, we need to consider what it is to be blameworthy. I will begin by discussing several ways in which a theory which understands blameworthiness in terms of the fittingness of resentment and indignation might attempt to answer this question. I will argue that all of the most plausible candidates face serious difficulties (sections 3 – 5). The problem is that it is very hard to develop a view of the representational content of blame that can account for the ways in which, intuitively, blameworthiness might diminish or disappear. The solution, I will argue, is to adopt a different account of blameworthiness, which does not rely merely on the fittingness of other-directed blame, but rather on the desert of self-directed blame. If we understand blameworthiness in terms of deserved guilt, we can give a plausible account of how blameworthiness can change over time. In section 6 and 7 I present this account. The thesis I will defend is this: whether an agent remains blameworthy will depend of whether she has experienced the guilt she deserved to feel for her action, attitude or omission.

2. Blameworthiness and the Reactive Attitudes
The starting point for my discussion is the following, commonly accepted, generic biconditional:

**Blameworthiness:** An agent S is blameworthy for X iff it is appropriate to blame S for X.\(^3\)

---

However, sometimes we use the word “blameworthiness” in a different way. We talk of an agent being *guilty of a wrongdoing*, meaning simply that S did some wrongful action X at some point t and that they satisfied the control and knowledge conditions for being morally responsible for X at t. No matter what happens after, it remains the case that the agent did x under certain conditions. In this sense, it will remain true that they are guilty of the wrongdoing.\(^4\) Note however that this conception of blameworthiness does not say anything about how we should treat or feel towards agents who are blameworthy in this sense. Consider an analogy.\(^5\) An offender who serves his time is nevertheless still guilty of his crime. It will remain true that he has committed the crime, acting with the relevant capacities. However, it would no longer be appropriate to keep him in prison. After all, he has gotten what he deserved. Whereas punishment was appropriate when he was sentenced, it no longer is when he has served his time. Similarly, although a culpable wrongdoer will always remain guilty of his wrongdoing, it is by no means obvious whether it will always be appropriate to blame him for his wrongdoing. What gives the debate about moral blameworthiness its urgency is the question of whether our practices of feeling and expressing blame are appropriate at a particular time. Therefore, I will be concerned with blameworthiness in the sense that is identified by the appropriateness of blame, and not with blameworthiness in sense of being guilty of a wrongdoing.\(^6\)

The accounts of blameworthiness I will focus on in this paper understand blame as an emotion.\(^7\) It is typically understood either as resentment or indignation (Hieronymi 2001; Graham 2014; Menges 2017; Strabbing 2018). Appropriateness is understood as fittingness. Emotions are commonly taken to represent their objects has having evaluative properties. Envy portrays one’s rival as having something that one lacks, and casts this circumstance in a negative light. Regret represents one’s action as a mistake. Emotions are fitting only when such representations are correct or accurate. As D’Arms and Jacobson (2000: 72) put it, “[i]n this respect, “the fittingness of an emotion is like the truth of a belief.” This conception of fittingness

---

\(^4\) Khoury and Matheson (2018:207) call this the trivial sense of blameworthiness, since it follows trivially from the fixity of the past.

\(^5\) I borrow this analogy from Clarke (manuscript) who argues that blameworthiness ought to be understood analogous to legal guilt rather than analogous to the appropriateness of punishment.

\(^6\) Thanks to Michael McKenna for pressing me on this point

\(^7\) I will also focus on moral blameworthiness as accountability in this paper. For discussions of other kinds of blameworthiness, see Watson (2004), Shoemaker (2015), and Carlsson (2019)
is sometimes called “alethic”. (Rosen 2015) Once we specify the relevant notions of blame and appropriateness, we get the following account of blame:

**Blameworthiness as fitting resentment**: An agent S is blameworthy for X iff it is fitting to resent S for X. (Hieronymi 2001; Graham 2014; Strabbing 2018)

According to **Blameworthiness as fitting resentment**, an agent is blameworthy to the extent that resentment is fitting. Whether blame is fitting will thus depend on what resentment represents. Different views give different accounts of the representational content of blame. Graham (2014), for example, takes resentment to be constituted only by a thought about the wrongdoer’s insufficiently good will. Strabbing (2018) adds to that a thought about the wrongdoer’s capacity to have acted better than she did. Such differences will yield different accounts of the conditions of blameworthiness.

However, to answer the question of whether blameworthiness can change over time, we also need to know how these thoughts are indexed to time. Does resentment represent the wrongdoer’s quality of will or capacity to act better at the time of the wrongful action, or does it rather represent these attributes at the moment of blaming?

### 3. Resentment Indexed to the Past

Let us begin with the first option. Although few accounts have addressed this issue explicitly, it is possible to draw conclusions from how the representational content is formulated by different philosophers. Graham writes: “the content of a blame emotion felt toward a person for phi-ing is that, in phi-ing, that person has violated a moral requirement of respect.” (Graham 2014: 408, my italics). Similarly, Strabbing defines the two constitutive thoughts as follows: “(1) In doing A, S expressed insufficient good will (toward me), (2) S could have done better.” (2018: 3136, my italics). Both of these formulations define the representations in the past tense. Resentment represents something in the past. This gives us:

---

8 Note that this sense of fittingness is different from the ones typically employed in the X-first literature. The alethic sense of fittingness is neither analysed in terms of reasons nor value, nor is it a sui generis normative concept. For discussion of different senses of fittingness, see Howard (2018) and Macnamara (2020)

9 Shoemaker (2015) shares this framework but argues that we should rather understand blaming emotions in the accountability sense as what he calls “agential anger.”

10 More specifically, he takes the blaming emotions to represent the wrongdoer as having violated “a moral requirement of respect”.

11 Coleman and Sarch (2012: 14) note that this may be way that reactive attitudes accounts of moral blameworthiness might attempt to deal with the issue of blameworthiness over time.
**Resentment indexed to the past:** Resentment represents the blameworthy agent as having acted wrongly, with ill will, sufficient control etc. *at the time of wrongdoing.*

This emotion will be fitting if what the emotion represented is correct: An agent is blameworthy if and only she acted wrongly with control, knowledge, insufficient quality of will, etc. at the time of action. As a matter of phenomenology, this way of indexing the representational content seems plausible. Blame is typically backward-looking, so it seems natural to take its representational content to be about the past: about the wrongness of the action and the agent’s mental states, at the time of action, or prior to the action. When we blame other people, it is directed toward them as they are now, but what we blame them for is typically not about who they are now, but rather about what they did in the past.

An implication of this view is that blameworthiness, like diamonds, is forever.\(^{12}\) On this view, resentment represents what happened in the past, and nothing that happens in the present can change the past. Resentment will always be fitting, no matter what happens afterwards, since the content of this representation will always be correct. Therefore, the agent will always be blameworthy. This is a problematic implication of the view. Guilt, atonement, reparation and forgiveness seem able to lessen or extinguish one’s blameworthiness. But if resentment is indexed to the past, those actions and emotions would not make resentment any less fitting.

The most principled reply on behalf of **Resentment indexed to the past** is to maintain that resentment is always fitting when its representational content is correct, so the wrongdoer would always be blameworthy. We should, however, distinguish between fitting blame *and an all things considered justification* of blame. Although resentment is always fitting, on this view, there are often reasons against feeling resentment, and these reasons can gain strength over time. These reasons will not bear on the blameworthiness of an agent. In this sense, they are reasons of the wrong kind. However, that doesn’t make them any less important when considering whether to continue to blame. First, one may perhaps lack the standing to blame. Over time, for example in a long relationship, one might have come to commit many of the wrongs for which one initially blamed the other. This provides reasons against blaming. Second, continued resentment over long stretches of time will often be damaging to our relationships and detrimental to our own well-being. This fact gives us both moral and prudential reasons to

\(^{12}\) I borrow this phrase from Khoury and Matheson(2018).
let go of blame. Third, and perhaps most importantly, the *significance* of the wrongdoing in one’s life will change over time. The reason why we resent someone who wronged us depends on many factors: a comment might be hurtful because of one’s vulnerability at a certain point in life, or because the role the wrongdoer played in our life at that moment. These things might change and this may change one’s reasons to feel resentment, although resentment will nevertheless continue to be fitting. Given the prevalence of strong, albeit wrong kinds of reason not to feel resentment for wrongs that happened long ago, the proponent of *Resentment indexed to the past* could provide a debunking explanation for why one might think that blameworthiness can diminish or disappear over time: Blameworthiness is forever, but our (wrong kinds of) reasons for continued blame are not.

I don’t think this reply is successful. Although many of our reasons to stop blaming do not have any bearing on whether an agent remains blameworthy, there are many reasons which do. Consider self-blame. As Douglas Portmore has argued, the fact that an agent already has blamed herself by feeling guilty seems to matter to whether it is appropriate for her to continue blaming herself (2019b:15). Unlike the reasons for ceasing to blame discussed above, the fact that you already have blamed yourself does seem to affect your blameworthiness. We can illustrate this with a comparative case. Consider two agents who both acted wrongly. Suppose that their wrongdoings are similar in all relevant respects. Their actions were equally wrong, committed with the same quality of will, knowledge and control. One of the wrongdoer experienced guilt for what he did: he was pained by the recognition of his wrongdoing. The other wrongdoer did not experience guilt. Although both were equally blameworthy at the time of action, it seems, at the very least, that the former wrongdoer is *less* blameworthy than the latter at some later point in time. Yet, according to *Resentment indexed to the past*, it follows that they are equally blameworthy, since by stipulation, they were similar in all relevant aspects at the time of action. Similar cases could be constructed where one but not the other wrongdoer not only experiences guilt, but also apologize, atones and is forgiven. Intuitively, these kinds of actions will make the wrongdoer less blameworthy than he otherwise would have been, but *Resentment indexed to past* has no way of explaining why this should be the case.

4. Resentment Indexed to the Present
To avoid these problems, a natural solution would be to think that resentment is not indexed to the past, but rather to some objectionable quality that the agent displays at the present. Call this

---

13 For an interesting discussion of letting go of resentment, see Milam and Brunning (2018).
view **Resentment indexed to the present.** Recently, Khoury and Matheson (2018) have defended this kind of view. The main point of their paper is to show that blameworthiness can diminish over time, depending on the wrongdoer’s psychological connectedness to her past self. Even though a wrongdoer is psychologically continuous with her past self (through overlapping chains of strong psychological connectedness), it does not follow that she now shares any of the objectionable psychological features of her past self. It is possible to be psychological continuous with one’s past self, but at the same time be psychologically disconnected to it. If that is the case, the agent is no longer blameworthy. They illustrate their claim with the example of Leon, who drinks from the well of immortality. At the age of 30 Leon commits a terrible crime. However, at the age of 530 Leon “is a perfect time-slice psychological twin of your favourite moral saint. At 530 Leon shares no distinctive psychological features with Leon 30 but remain psychologically and biologically continuous with him.” (Khoury and Matheson 2018: 214). Khoury and Matheson argue that at 530, Leon is not blameworthy for the atrocity he committed when he was 30, because he no longer shares any of the psychological features that were essential to his performing the atrocity.

They rely on the following version of **Blameworthiness as fitting resentment:** “we claim that the intentional content of blame involves the attribution of a flaw to the subject. That is, blaming involves thinking or representing the subject to have a flaw. And a necessary condition of the subject being a worthy target of blame is the truth of this thought or accuracy of this representation” (Khoury and Matheson 2018: 222). It is clear from the context that they take resentment to represent the wrongdoer as having a flaw at the moment of blaming. They use this account to explain how someone whose psychology changes radically can cease to be blameworthy: if the agent no longer displays the flaw, the representational content is no longer correct. This means that resentment is no longer fitting, and thus, that the agent is no longer blameworthy. This view can better account for the comparative case presented above. This is not because guilt or apologies as such lessen the agent’s blameworthiness, but rather because guilt and apologies can serve as evidence for the fact that the flaw is gone.

I think this account suffers from two problems. First, the intuition that guilt or apologies may lessen an agent’s blameworthiness is not, it seems to me, best explained by that fact that these emotions and actions are evidence for a changed personality. To illustrate, suppose my flaw that made me blameworthy in the first place just disappears by accident. In that case, it seems to me that the agent would still be blameworthy. Or suppose I take a pill to remove the flaw, not in order to become a better person, but – knowing Khoury and Matheson’s theory – in order to avoid fitting blame in the future. In this case, my decision to take the blame is
obviously indicative of some kind of flaw. But the pill works and now I don’t have the flaw anymore. Given that resentment on this theory is indexed to the present, it follows that the agent is no longer blameworthy. This seems to be the wrong result.

Second, it also seems that the theory can sometimes make it too hard to avoid blame. Suppose you acted wrongly because of some flaw. Afterwards you feel guilt, sincerely try to reform, and ask for forgiveness. However, the flaw remains. It does not seem obvious that you are still blameworthy (at least to the same extent) for that very action.

5. A Mixed Account

Both of the accounts of resentment considered so far assumes that resentment has a simple representational content. But perhaps resentment is more complex? On Pamela Hieronymi’s (2001) view, resentment is constituted by several judgments. Hieronymi’s account is particularly relevant for our purposes because it is developed with the aim of explaining how forgiveness could be “articulate”. Forgiveness, she argues, requires an overcoming of resentment but not every kind of overcoming would count as forgiveness. The fact that resentment can be painful, counterproductive or damaging to one’s relationships constitutes perfectly good reasons to let go of resentment. But letting go of blame for these reasons is not forgiveness. Hieronymi wants to develop an account of forgiveness on which it can, to use the terminology I have been employing in this paper, be fitting to stop resenting. The challenge is to explain how it could be fitting to overcome blame without excusing the wrongdoer. To be blameworthy is to be an appropriate target of blame. David Shoemaker puts the challenge well: “how might one appropriately withdraw blaming emotions fitting for someone’s culpable moral transgression without also withdrawing the judgment making them fitting, namely that that person culpably committed a moral transgression” (forthcoming: 6). Hieronymi’s solution is to develop an account of blame on which resentment is constituted by several different judgments. In order for me to resent a wrongdoer, she argues, I must make the following three judgments: 1) the action was wrong; 2) the wrongdoer was blameworthy for the action; 3) I, as the one wronged, ought not to be wronged (2001:530). These judgments are necessary but not sufficient for resentment. Importantly, Hieronymi also adds a fourth judgment: 4) the past wrong makes the false claim that it is acceptable that I am treated this way, and this claim, if left unaddressed, poses a threat to me. Resentment, on Hieronymi’s view, is fundamentally a protest against the past action that persists as a threat (2001:546). These judgments allow Hieronymi to provide a solution to the puzzle of forgiveness. When a wrongdoer sincerely apologizes and repudiates his wrong, the fourth judgment is no longer true: the past wrong no longer makes a threatening
claim. This means that the resentment can fittingly disappear. Moreover, this can happen without excusing the wrongdoer, since the first three claims remain true.

Hieronymi’s project in that paper is to develop an account of forgiveness and not to explain how blameworthiness might change over time. However, we might use some elements from her account of resentment in order to address this problem. Hieronymi is offering a *mixed* account of resentment, where some of the constitutive judgments of blame are indexed to the past, whereas one is indexed to the present. Judgments 1-3 will always remain fitting. The judgment that the wrongdoing constitutes a threatening claim, on the other hand, might change depending on the wrongdoer’s repudiation of his action.

Some tweaking is necessary for this account to work as a theory of blameworthiness. The third thought seems superfluous for our purposes. The second thought will not work when employed in the **Blameworthiness as fitting resentment**-framework. If a constitutive thought of resentment is that the agent is blameworthy, the account will be circular and non-informative. So for the purposes of this paper, let us substitute it with a judgment about the agent’s quality of will. We then get the following mixed account:

**Mixed account:** Resentment represents the wrongdoer as having a) acted wrongly; b) with a bad quality of will; and as making c) a threatening claim.

This account would have several benefits compared to **Resentment indexed to the past**. Like Matheson and Khoury’s account, it can explain how blameworthiness can change over time: once the threatening claim is sincerely repudiated, the threat is no longer present and resentment is no longer fitting, so the agent is no longer blameworthy. On this view, blame is fundamentally a protest against the threatening claim. Once this claim is taken back, there is nothing left to protest against. Similarly, it can account for the intuition that guilt expressed to the victim through atonement and apologies can influence an agent’s blameworthiness. These actions can all be seen as ways of repudiating the threat. Moreover, this view avoids some of the problems I raised for **Resentment indexed to the present**. On the mixed account, it is not sufficient to change one’s quality of will or to remove a flaw. It matters how one does it. If one’s psychology changed dramatically, either by accident or in an intentional attempt avoid blame in the future, without a repudiation of one’s past wrongdoing the agent will still be blameworthy, since the threatening claim will remain unaddressed.

However, the mixed account faces problems of its own. First, even though expressed guilt is one way of repudiating the wrongful claim implicit in one’s action, it seems possible to
renounce one’s action in ways that do not involve any painful emotions whatsoever. The wrongdoer might simply apologize to the victim while calmly explaining that he now understands that his action was wrong and promise never to anything like that again. This seems sufficient to retract the threat. Of course, the victim might have more confidence in the wrongdoer’s sincerity if he also expresses guilt, contrition or remorse for his action. But this painful emotion only plays an epistemic role for the mixed account. This leaves the possibility of a fully blameworthy agent who ceases be blameworthy without ever being pained by what he has done simply because his action no longer constitutes a threat. This strikes me as implausible.

Moreover, even in cases where the wrongful threat is repudiated through sincerely felt and expressed guilt and apologies, it is not clear that it is sufficient to render blame inappropriate. It seems that agents who have sincerely repudiated their past wrongs might nevertheless remain blameworthy. Consider two cases:

**Cheating.** Tom is married to Mary. He cheats on her. The morning after, Tom experiences a pained acknowledgement of what he has done. He feels guilty and apologizes to Mary. Nevertheless, upon hearing Tom’s expression of guilt, Mary blames Tom. She will continue to blame him for a long time.

**Assault.** John assaults Jerry. Jerry is traumatized physically and mentally by the ordeal. After being apprehended by police, John comes to feel guilty for what he done; he painfully acknowledges what he has done. During the trial, he expresses his sincere guilt to Jerry and repudiates his action. Nevertheless, Jerry continues to blame John for assaulting him.

As Hannah Tierney (2019) points out, blaming wrongdoers who already experiences and expresses their guilt, is in many cases perfectly appropriate. This, Tierney, argues, seems to create a problem for Hieronymi’s view. In the cases above, it is not clear in which sense Tom’s and John’s actions persist as threats. If resentment is fundamentally a protest against an unaddressed threat, how could it be fitting to blame wrongdoers for wrongs which do no longer count as threat on Hieronymi’s account? (Tierney 2019: 12). At this point one may appeal to the other constitutive judgments of resentment mentioned by Hieronymi. Although the judgment about the threatening claim is no longer true, it is still the case that what Tom’s and John’s actions are wrong, that they acted with insufficiently good will and that their victims ought not be treated as they were. However, the mixed view would need to choose between two
options. If the truth of the threatening claim is necessary to render the agent blameworthy, it seems that Tom and John are no longer blameworthy. If, on the other hand, the truth of these other thoughts involved in resentment is sufficient to render the agent blameworthy, even without the threatening claim, the view collapses into Resentment indexed to the past. Tom and John will be blameworthy forever. But it seems plausible that if they keep on experiencing guilt and amend their wrongdoings, they will at some point stop being blameworthy. At the very least, they will be less blameworthy than they once were.

David Shoemaker (forthcoming) develops a view that is in some ways similar to the mixed account sketched above. According to him, what he takes to be the paradigmatic blaming emotion – agential anger – represents actions or attitudes as slights. The anger is directed towards the wrongdoer’s failure to take the victim sufficiently seriously and to properly acknowledge him. The action tendency of agential anger is to communicate the anger to the wrongdoer, thereby “delivering a request that the slighter emphatically acknowledge what he did” (Shoemaker forthcoming: 30). Shoemaker argues that the wrongdoer’s lack of sufficient acknowledgment creates a normative disequilibrium between the victim and the wrongdoer. Whereas for Hieronymi blame is fundamentally a protest against a persisting threat, Shoemaker understands blame as an angry request for acknowledgement.

This account of the nature of blame also helps explain how blame can cease to be fitting. When the wrongdoer has emphatically acknowledged the victim, the normative equilibrium is restored. This happens, according to Shoemaker when the wrongdoer feels remorse. Shoemaker understands remorse as a “painful emotional response to my recognition of having caused you a irremediable loss of value, a response which constitutively involves my being moved to reflect on (over and over) what I did from your perspective” (Shoemaker forthcoming: 11). For a wrongdoer to feel remorse he must take up the victim’s perspective and be open to feeling some approximation how the victim felt as result of the wrongdoing. Shoemaker thus understands remorse as a painful emphatic acknowledgement.

Hieronymi takes blame to represent an ongoing threat. This threat disappears when the wrongdoer repudiates his action. As a result, blame is no longer fitting. Shoemaker, on the other hand, takes blame to represent an ongoing slight: a failure of acknowledgement. On his account, the “offender’s emphatic acknowledgement is sufficient to make it the case that the slight is no more” [Shoemaker’s italics]. As a result, blame is no longer fitting. It remains true, of course,

14 Shoemaker’s methodological approach in this paper is to start out with an account of what makes anger go away. Just as Hieronymi, Shoemaker develops an account of how blame can stop being fitting as a way of explaining the puzzle of forgiveness.
that there was a normative disequilibrium – a slight – so the victim can still view the wrongdoer as someone who merited anger. But because the victim has gotten what he requested from the wrongdoer – emphatic acknowledgement – the victim no longer has a reason of fit to blame the wrongdoer. Hence, the wrongdoer is no longer blameworthy.

Shoemaker’s account is an improvement on the mixed account. Whereas a repudiation need not involve any painful emotion, the emphatic acknowledgement involved in remorse is painful. So Shoemaker’s account is not vulnerable to the objection that a mere repudiation of one’s action is insufficient to stop being blameworthy. It might seem, however, that Shoemaker is vulnerable to the counterexamples concerning serious wrongdoing discussed above. In *Cheating* and *Assault* it would seem perfectly appropriate for Mary and Jerry to continue to blame Tom and John even if they had experienced and expressed remorse for their wrongdoing. Here, however, Shoemaker might object that the kind of remorse he has in mind involves *emphatic acknowledgement*: the wrongdoers need to fully appreciate what it is like to be pained or damaged by the wrongdoing. For serious wrongdoing, it may take time to fully appreciate this. Shoemaker might thus argue that the reason why Mary and Jerry still have reasons (of fit) to blame Tom and John is that these wrongdoers have not yet fully appreciated the pain they caused.

Nevertheless, I believe that Shoemaker’s account suffers from serious problems. The crucial element in Shoemaker’s story is the wrongdoer’s emotional perceptual stance rather than the duration and intensity of his painful emotion. If I wronged you, I must take up your emotional perspective and come to imagine how it must have felt for you when I treated you badly. I will then have experienced a simulacrum of your pain. Once that happens, Shoemaker argues, “it looks appropriate for you to abandon hard feelings in favor of forgiveness just as soon as you have witnessed (and believed) my own sincere emotional devastation in light of what I did. It is, after all, obvious that I clearly and truly get what I did, which looks to be enough to disarm your blaming appraisal of me.” (Shoemaker forthcoming 16-17, my italics).

However, one might appropriately continue to blame a wrongdoer even after it becomes obvious that they clearly and truly get what they did. This seems clear in cases of serious wrongdoing. The fact that I have taken up your emotional perspective and that I am pained by the loss I caused you is one crucial step in a process towards blame becoming inappropriate, but it is not its end point. Once I have truly acknowledged the pain I caused you, it will be appropriate for me to apologize, try to make up for my wrongdoing, and ask for forgiveness.

---

15 See Shoemaker’s contribution to this volume, footnote 30.
These practices make sense in light of the fact that victim’s blame continues to be appropriate. But on Shoemaker’s account the blame would cease to be fitting as soon as the wrongdoer truly get what they did. Moreover, Shoemaker’s account will also struggle to make sense of the appropriateness of *self-blame*. If it is no longer fitting to blame the wrongdoer, he is no longer blameworthy. If the wrongdoer is no longer blameworthy, it would also, presumably, no longer be fitting for him to blame himself. But the fact that a wrongdoer has truly acknowledged the pain he has caused his victim does not mean that it is appropriate for him to (immediately) stop blaming himself. That acknowledgement seems to be the *beginning* of a process of appropriate self-blame. Once this acknowledgement happens, the wrongdoer will typically begin to feel guilty, and it seems appropriate to do so, at least for some time.

### 6. Blameworthiness as Deserved Guilt

I have argued that accounts of blameworthiness that understand blameworthiness in terms of fitting resentment or agential anger will face difficulties in explaining how blameworthiness may change over time. Intuitively, there are certain things that a wrongdoer can do or feel that will diminish her blameworthiness. She can experience guilt, apologize or make amends. Versions of Blameworthiness as *fitting resentment* struggle to account for this, either because resentment is indexed to the past, or because it is indexed to the present. What matters to a wrongdoer’s blameworthiness *now* is not merely her mental state at the time of the action, nor her mental state at the present time, but also the *moral process* she has undergone since the wrongdoing in the past. Even accounts, like Hieronymi’s and Shoemaker’s, that aim to capture these features struggle to get the details right. Fitting resentment views seem ill suited to capture this drama of atonement.

In this section, I will present an alternative theory. Accounts of blameworthiness tend to start with other-directed blame: Agents are blameworthy for *x* to the extent that it is appropriate that others blame them for *x*. I suggest that we instead build our theory of blameworthiness around *blaming oneself*. To blame oneself is to experience guilt (Wallace 1994; Clarke 2013). Guilt is characterized by its action tendencies, its representational content and its affect. The action tendency of guilt is to express the guilt and attempting to repair the relationship that has been damaged by the wrongdoing (Baumeister et al. 1994; Haidt 2003; 16)

16 This section summarizes arguments presented more fully in Carlsson (2017) and (2019)

17 More specifically, to blame oneself in the accountability sense is to experience guilt. I believe that to blame oneself in the attributability sense is to experience shame. See Carlsson (2019). For alternative accounts of self-blame, see McKenna’s, Pereboom’s, and Shoemaker’s contribution to this volume.
Lazarus 1991; Nichols 2007). I do not want to attempt a full account of the representational content of guilt. This strikes me as an extremely difficult empirical question. Answering it requires introspection, and there is no guarantee that people will find the same thing when they look inside. I think the best we can do is to give a rough gloss. In an earlier paper I suggested (following Graham 2014) that the representational content of guilt might simply be that one acted with insufficiently good will.\(^{18}\) It now seems to me that one can feel guilty for cases in which one does not have an insufficiently good will, for example, for unwitting omissions, or for missing a crucial penalty shot. It seems to me now that the most promising gloss on what guilt represents might be that one has violated a legitimate normative expectation (Wallace 1994), or perhaps simply that one has engaged in either a wrongdoing or in a personal betrayal (D’Arms and Jacobson this volume).

Importantly, guilt is also characterized by its painful affect. When one experiences guilt, one is pained by the thought that one has violated a legitimate expectation. The thought does not necessitate the unpleasant affect. It is possible to believe that one has violated a legitimate expectation without being pained by this thought. But in that case it would not be guilt. In this sense, guilt is constituted by its painful affect. This is important and sets guilt apart from other-directed blame. Unlike resentment and indignation, guilt is necessarily experienced by the wrongdoer: the suffering involved in guilt is a necessary part of what it is to blame oneself. This suffering comes in different degrees. It can vary from a mild discomfort to a prolonged state of agony (Clarke 2013: 155). But if the emotional state does not involve suffering at all, it is not guilt.

An agent is blameworthy, I suggest, to the extent that it is appropriate for her to feel guilt for what she has done. What is the relevant notion of appropriateness? I want to argue that the relevant notion is desert. Just as we can ask whether pain involved in punishment, sanction or harsh treatment is deserved, we can also ask whether the painfulness of guilt is deserved.

Desert differs from fittingness in several respects.\(^{19}\) First, desert, in contrast to fittingness, is a distinctly moral notion, which entails non-instrumental goodness (McKenna 2012).\(^{20}\) Although some fitting emotional responses, such as grief or gratitude, may be non-

\(^{18}\) See Carlsson (2017).

\(^{19}\) The sense of desert I’m concerned with here is what Feinberg (1970) calls “personal desert.” It differs from the sense in which we sometimes say that an artwork deserves admiration or a problem deserves consideration. It seems that we can capture the latter sense of desert by using the word “merit.” See Howard (2018) for some examples of how merit and personal desert differ.

\(^{20}\) For an account of desert that neither entails non-instrumental goodness, nor provide pro tanto reasons to bring about what is deserved, see Nelkin 2019.
instrumentally good, non-instrumental goodness is not entailed by fittingness. It can be fitting to feel envy or disgust or to be amused by a racist joke, but it is not non-instrumentally good to be envious, disgusted or amused by racist jokes. Moreover, whether an agent deserves something harmful is a consideration of justice (Feinberg 1970; Clarke 2016). It can be fitting to feel regret or grief, but it is not just. Finally, desert is uncontroversially a robustly normative notion. The normative status of fittingness in the alethic sense, according to which the fittingness of an emotion just consists in the correctness of its representation is less clear cut (Tappolet 2011). In general, it is not obvious that we have reasons to have true representations. It is even less obvious that the fact that an emotion would be an accurate representation of some evaluative property would be a reason to experience that emotion. It seems doubtful that we have any reason to feel schadenfreude, although schadenfreude might be fitting. By contrast, the fact that an agent is deserving of something, provides pro tanto reasons to make it the case that the agent gets what he deserves. These features of desert help explain why it is prima facie permissible to impose deserved harms on people who deserve it. It is prima facie permissible to make a wrongdoer experience the guilt he deserves, for example by blaming him. But it is not prima facie permissible to impose painful emotions such as grief, embarrassment or regret. The reason is that these emotions can be fitting, but they are not deserved (Pereboom 2017; Carlsson 2017).

By substituting resentment with guilt, and fittingness with desert, we get the following account of blameworthiness:

**Blameworthiness as deserved guilt (DG):** An agent S is blameworthy for x if and only if, and in virtue of the fact that S deserves to feel guilty for x (Carlsson 2017)

An agent is blameworthy, I propose, because she deserves to feel guilt. Deserving guilt is both a necessary and sufficient condition on being blameworthy. DG provides an explanation of how blameworthiness relates both to emotions and to harm. Like Blameworthiness as fitting resentment, it understands blame as a matter of emotional reactions and not as an overt form of treatment. This allows for the possibility of private blame. But unlike Blameworthiness as fitting resentment, it also makes sense of the thought that being blameworthy involves deserved harms or suffering. However, the suffering is of a special kind. When you feel guilt,

---

21 Although there might be a standing reason against holding false beliefs. For discussion, see Sharadin (2015).
22 For accounts that understands blame as an overt treatment, see Wallace (1994) and McKenna (2012)
you suffer in recognizing what you have done. I believe that this captures the specific kind of suffering that we take blameworthy agents to deserve. It is not suffering in general, but rather the specific pain of acknowledging that you have engaged in wrongdoing or violated a legitimate expectation. An agent may suffer as a result of being blamed, without recognizing his own fault, simply because he finds criticism unpleasant or because it affects his social standing. However, what he really deserves is the pain of guilt; the suffering involved in recognizing what he has done.

Guilt must thus be fitting in the alethic sense in order to be deserved. It is only when what guilt represents is correct that the agent will be deserving of the specific pain involved in guilt. But although the fittingness of guilt is a necessary condition on blameworthiness, it is not sufficient. If an agent acted wrongly or violated a legitimate normative expectation guilt will be fitting, but this does not mean that the pain of guilt is deserved. If the agent had an excuse, it will neither be noninstrumentally good, nor just that the agent is pained by the recognition that he acted wrongly.

Another strength of DG is that it can give a straightforward explanation of excuses and exemptions. A common way to explain excuses and exemptions is by appealing to the harm of blame. When these hostile emotions are expressed, they can be harmful. Because of this prospect of pain and hostile treatment, certain control conditions must be met for blame to be justified (Wallace 1994; Nelkin 2011, Rosen 2015). The problem with this approach is that other - directed blame need not be expressed at all, and even when it is expressed it need not be harmful (Nelkin 2013). Self-blame, understood as guilt, by contrast, is necessarily painful. DG provides a forceful and straightforward explanation for why control is a necessary requirement on moral blameworthiness without identifying blame with any form of overt treatment. To be blameworthy is to deserve to feel guilt. To feel guilt is to suffer. But no one deserves to suffer for what they cannot control. So no one is blameworthy for they cannot control.

Similarly, when people act wrongly without having the capacity to recognize moral reasons, or under duress, or while being coerced, they are not blameworthy because they do not deserve to experience the pain of guilt. The benefits of this way of accounting for excuses and exemptions can be seen by comparing it to Blameworthiness as fitting resentment. For such accounts to explain the variety of excuses and exemptions, the conditions of blameworthiness would need to be a part of, or explained by, the representational content of resentment. The content of resentment could not simply be that the agent acted with an insufficient quality of will, nor could it be that she

---

23 See D’Arms and Jacobson (this volume) as well as Portmore (this volume) for more on this point.
engaged in wrongdoing or violated a legitimate expectation, for it is possible for those representations to be true without the agent being blameworthy. A child might act with an insufficient quality of will and someone might be ignorant about the legitimate expectation he violated. A conjunction of such conditions would do better. However, to capture all of the commonly accepted conditions of blameworthiness, such as voluntariness, knowledge and control in the representational content of resentment while at the same time maintaining psychological plausibility is very difficult. The content would become implausibly complex and run the danger of violating what Gideon Rosen (2015) has called the naivety constraint: the representational content of guilt must be framed in terms that everyone capable of guilt understands. DG avoids the problem by explaining exemptions and excuses not by the content of the blaming emotion, but rather by appealing to moral facts about when the pain of guilt is deserved.

DG also makes sense of a common observation about the nature of blame. Many philosophers have recently argued that the communicative aim of blame is to generate guilt or remorse in the wrongdoer. According to Miranda Fricker (2016: 173) “[I]n Communicative Blame the speech act is geared specifically to bring us to feel the proper pang of remorse, where remorse is understood as a cognitively charged moral emotion – a moral perception that delivers a pained understanding of the moral wrong we have done.” Macnamara argues “Specifically, emotional uptake of the representational content of resentment or indignation by the wrongdoer amounts to guilt.” (2015: 559). Other philosophers suggest that guilt is the retributive aim of blame. As Gideon Rosen notes, “The wrongdoer who responds to outward blame with a sincere and cheerful promise to do better next time but without a hint of guilt or remorse palpably frustrates a desire implicit in resentment.” (Rosen 2015: 83). Similarly, Susan Wolf (2011) argues that the point of angry blame is to get the target of blame “…to experience the painful feelings of guilt and remorse” (2011: 338).

If it is an essential part of the nature of blame that it aims at generating a painful emotional state in its recipient, this raises a question of how blame can be justified. DG provides a natural justification for this kind of blaming. When wrongdoers deserve to feel guilty, it will be prima facie permissible for others to induce guilt in the wrongdoer. Moreover, the blameworthy wrongdoer cannot reasonably object to the fact that blame generates guilt in him, if feeling guilty is what he deserves.

---

24 See also Mason (2019).
26 However, he could object if blaming lead to other kinds of harms, or to disproportionate guilt.
7. Deserved Guilt and Blameworthiness over Time

DG has the following implication: If an agent does not deserve to feel guilty, she is not blameworthy. This means that if an agent who did deserve to feel guilt, for some reason no longer deserves to feel guilty, she is no longer blameworthy. How could one stop deserving guilt? One way would be simply by having experienced a certain amount of feeling guilty. To be blameworthy, I suggested, is to deserve to feel guilty. However, this claim must be modified. As Clarke (2013) has noted, the pain of guilt comes in different degrees and it can last for a longer or shorter time. A blameworthy agent deserves to feel guilty to the right degree and for the right amount of time. Moreover, as Portmore (2019a) has argued, experiencing a certain amount of guilt for a particular wrongdoing will influence how much guilt one deserves to experience for that wrongdoing in the future. “After all, one deserves to suffer only so much self-reproach for any given wrongdoing” (2019: 62-63). DG thus provides a clear explanation for how an agent’s blameworthiness can change over time. A blameworthy agent deserves to feel a certain amount of guilt. When she has experienced the right amount of guilt, she has gotten what she deserves. It is no longer just and non-instrumentally good that she suffers the pain of guilt. Since being blameworthy on DG is to deserve feeling guilt, this means that the agent is no longer blameworthy. Moreover, the fact that she already has experienced some but not all of the guilt she deserves, makes her less blameworthy than she otherwise would have been.

DG is well equipped to solve the problems I raised for Blameworthiness as fitting resentment. First, in contrast to Resentment indexed to the past, it can explain why experienced guilt matters to one’s degree of blameworthiness. Recall the two agents whose wrongdoings were similar in all relevant respects. One of them experiences guilt afterwards, the other does not. DG can explain why the former seems to be less blameworthy than the latter. He has gotten more of what he deserved.

Second, in contrast to Resentment indexed to the present, DG can explain how one can cease to be blameworthy without the disappearance of a flaw being either sufficient or necessary: agents can remain blameworthy even though they no longer have the flaw. A person who simply loses one’s flaw by accident or take a pill in order to remove the flaw so that they can avoid appropriate blame in the future will remain blameworthy because he still deserves to feel guilty. On the other hand, an agent may cease being blameworthy even though he still has the same flaw so long as he has experienced the right amount of guilt.
Third, in contrast to The mixed account, it can explain how agents can remain being blameworthy even though they have experienced guilt and repudiated their wrongdoing so that it does not persist as a threat. Tom and John are less blameworthy than they would have been without their guilt. But they are nevertheless still blameworthy since they still deserve to feel guilty. Their victims would thus have a pro tanto justification for continued blame. Similarly, in contrast to Shoemaker’s account, it can also explain why both other-directed and self-directed blame can remain appropriate even after the wrongdoer has come to fully acknowledge the pain he caused his victim. Emphatic acknowledgement does not settle the issue of whether the wrongdoer still deserves to feel guilty.

DG claims that an agent ceases to be blameworthy as soon as she has experienced the right amount of guilt. Is this plausible? DG is compatible with the claim that some agents who have committed grievous wrongs will always remain blameworthy. It might very well be the case that such agents will continue to deserve guilt for their entire life, and that no amount of experienced guilt will equal the guilt they deserve to experience. However, it seems that, at least for many kinds of wrongdoing, there are limits to the amount of guilt a wrongdoer deserves. When a wrongdoer has experienced the amount of guilt she deserves, she is no longer blameworthy. 27

This raises three important questions for my account. First, does this mean that blame is never justified when agent already experiences the guilt they deserve? Hannah Tierney (2019) has emphasized that we often continue to blame wrongdoers who feel and express guilt to their victims. She argues that this raises a serious worry for my view. DG provides a pro tanto justification of blame because blameworthy agents deserve to feel guilty. When people already feel the right amount of guilt, this justification will no longer be applicable. Yet, Tierney argues, this kind of blame is often perfectly appropriate. She suggests that blame has two different characteristic functions. One is to make the wrongdoer feel the appropriate guilt. But another,

27 According to Portmore (2019a; 2019b; 2019c; forthcoming), guilt is constituted by the thought that one deserves the unpleasantness of guilt. So, on Portmore’s account, an agent is only blameworthy if guilt is fitting, and guilt is only fitting if the wrongdoer deserves the unpleasantness of guilt. It is an elegant package. Since both of our views take an agent’s blameworthiness to depend on whether she deserves to feel guilty, Portmore could explain how blameworthiness can change over time in exactly the same way as I have outlined in section 7. Why not embrace Portmore paper, but I think that we have good reasons to doubt that guilt is in fact constituted by a thought about deserved unpleasantness. First, it seems highly unusual that the content of any emotion would be constituted by a representation concerning its own justification. It seems more natural that we often have independent thoughts about whether our emotional reactions are justified. Second, it is crucial to Portmore’s account that guilt is constituted by a thought about deserved, rather than fitting unpleasantness. But it in order to distinguish fittingness from desert Portmore needs to appeal to notions such as non-instrumental goodness. If we accept the naivety constraint, the representational content must be framed in terms everybody capable of experiencing the emotions understand. I don’t believe that this is the case when guilt is constituted by desert and desert in understood in terms of non-instrumental goodness and justice.
equally important, function is to communicate to the wider community as well as to the wrongdoer that the victim possesses self-respect. In expressing blame, the victim, according to Tierney, communicates that she should not be treated this way, and that she deserves moral attention, care or concern in light of the perpetrator’s behaviour. I believe that Tierney is right that blame often serves this function. However, I also believe that DG can allow for this. Blame can be justified in many ways. If an agent has already suffered the guilt he deserves, he is no longer blameworthy. This provides a *pro tanto* reason not to generate any more guilt in him. However, the importance of communicating self-respect to the wrongdoer and to the wider community provides a *pro tanto* reason to continue blaming. These considerations need to be weighed against each other. Moreover, I don’t think that expressing the claim that the victim deserves moral attention, care or concern in light of the wrongdoer’s behaviour requires that the wrongdoer is blameworthy. It is for example contested whether people are blameworthy for their unwitting omissions. Suppose I forgot my wife’s birthday. It may be the case that I don’t deserve to feel guilty for this. Nevertheless, it seems quite clear to me that my wife ought not to be treated this way, and that she deserves moral attention, care etc. in light of my omission. This aspect of blame will often be justified even when the agent is not blameworthy.

Another worry is that explanation I have given is simply too focused on the wrongdoer’s emotions. As I emphasized when criticizing versions of *Blameworthiness as fitting resentment*, it also matters what the wrongdoer does after his wrongdoing. Feeling guilty can lessen a wrongdoer’s blameworthiness, but so, intuitively, can apologizing, compensating, and various sorts making amends. It would be problematic if these sorts of actions matter to a wrongdoer’s level of blameworthiness only to the extent that they served as evidence for her feelings of guilt. But DG does not have this implication. According to DG an agent is blameworthy to the extent she deserves to feel guilt. Experiencing guilt is *one* reason why one might deserve less guilt, but it is not the only one. Intuitively, the fact that one has apologized, atoned, or made amends will also matter to how much guilt one deserves to feel. What a blameworthy agent deserves is typically understood as a direct result of the wrongness of her action, her mental states at the time of action, her opportunity to avoid what she did, and the difficulty of avoiding doing what she did. These features relate to the time of her action. However, when we investigate what an agent deserves now, it also seems relevant to ask what she has done in the interval between the wrongful action and the present. Plausibly, when one acts wrongfully, with knowledge and control, one will not only deserve to feel guilty, but also incur certain duties towards one’s victim. These can come in various forms, but at the very least it seems that such wrongdoers often will have a duty to apologize, compensate and make
amends. It seems plausible that an agent will continue to deserve to feel guilty for his wrongful action until such duties are fulfilled.

A similar story might be told about forgiveness. Certain views on forgiveness, which I find attractive, argue that in committing a blameworthy wrongdoing a wrongdoer incurs a debt to the victim from which the victim might release the wrongdoer by forgiving him (Nelkin 2013b). One might worry that if my account were correct, forgiveness would play no role in releasing the victim from this debt. But this does not follow from my account. At least for certain grievous wrongs, it may be that agents will continue to deserve guilt, not only until they have experienced a certain amount of guilt, and fulfilled the duties incurred by their wrongdoing, but also until they are released by their debts to their victims by being forgiven.

The final worry concerns fittingness. I have argued that we should understand blameworthiness in terms of deserved guilt. Whether a wrongdoer will continue to deserve to feel guilty can change. But, as mentioned in section 6, for guilt to be deserved, it must be fitting in the alethic sense, it must be a correct representation. I also argued that guilt’s representational content is indexed to the past. It represents the agent has having violated a legitimate normative expectation of having engaged either in wrongdoing or personal betrayal. Since it is indexed to the past this representation will be true forever. This means that guilt might be fitting although the agent no longer deserves to feel guilty, and thus, no longer is blameworthy. This, however, is not a problem for my view. As noted in section 6, the fittingness of guilt is necessary but not sufficient for blameworthiness. The possibility of fitting but non-deserved guilt might perhaps have been problematic if fittingness were analysed in terms of reason or value, or if it was understood as a generic notion of appropriateness. But this is not the sense of fittingness that I have been discussing in the paper. On the alethic sense of fittingness, fittingness is simply a matter of correct representations. And the representation that I engaged in wrongdoing or violated a legitimate normative expectation is compatible with the fact that I no longer deserve to feel the pain of guilt for the wrongdoing, as well as with the fact that I’m no longer blameworthy for it.

8. Conclusion

In this paper I have argued that accounts that understand blameworthiness in terms of fitting resentment will struggle to make sense of how blameworthiness can change over time. It matters to an agent’s blameworthiness whether she has experienced guilt, apologized and made amends, but these views have difficulties in explaining why this is the case. The solution to the problem, I argued, is to adopt a different account of blameworthiness. An agent is blameworthy
to the extent that she deserves to feel guilty. When she no longer deserves to feel guilty, she is no longer blameworthy.28

References


Callard, A. (2017). The reason to be angry forever. The moral psychology of anger, 123-137.


28 Thanks to the participants of the workshop on self-blame at the University of Oslo in September 2019, including Randy Clarke, Justin D’Arms, Christel Fricke, Dan Jacobson, Coleen Macnamara, Michelle Mason, Michael McKenna, Dana Nelkin, Derk Pereboom, Doug Portmore, Piers Rawling, and Krista Thomason. Special thanks to Randy Clarke and Justin D’Arms for very helpful discussions of my paper. For excellent written comments on my paper, I’m very grateful to Leonhard Menges and Daniel Telech.
Hieronymi, P. 2001. “Articulating an Uncompromising Forgiveness”, *Philosophy and Phenomenological Research*


