

Shame and Attributability

Andreas Brekke Carlsson, University of Oslo

Penultimate draft, for Oxford Studies in Agency and Responsibility, vol. 6.

1. Introduction

It is increasingly common to distinguish between two kinds of moral responsibility: *responsibility as attributability* and *responsibility as accountability* (Watson 2004).¹ These kinds of responsibility are taken to correspond to different forms of blame. To be responsible in the attributability sense means that a certain kind of evaluation, what Gary Watson calls aretaic appraisal, is appropriate. This evaluation concerns how well or poorly an agent's actions, omissions or attitudes reflect on her character, on what she cares about or stands for: she might be kind, cowardly, brave, or selfish. Aretaic appraisals are meant to reflect "one's purposes, ends, choices, concerns, cares, attachments and commitments" (Watson 2004: 287). Negative aretaic appraisal will often concern immoral conduct, but Watson emphasizes that one can be blameworthy in the attributability sense even for behaviour that does not fall squarely into this category. All that is needed in order to be a proper object for negative aretaic appraisal is that the conduct is lacking in a way that reflects poorly on one's character:

If someone betrays her ideals by choosing a dull but secure occupation in favor of a riskier but potentially more enriching one, or endangers something of deep importance to her life for trivial ends (by sleeping too little and drinking too much before important performances, for example) then she has acted badly- cowardly, self-indulgently, at least unwisely. (Watson 2004: 266-276)

To be blameworthy in the *accountability* sense means that some more robust kind of blame than mere aretaic appraisal would be licensed. There are different ways of spelling out what this extra feature of blame consists in. But it is often taken to be the emotional opprobrium of a certain class of negative

¹ Watson (2004), Nelkin (2011), (2015), Fischer and Tognazzini (2011); Shoemaker (2015); Darwall (2016) all accept this distinction, although they draw it in slightly different ways.

reactive attitudes - resentment, indignation and guilt - or the harm in adverse treatment and sanctions. Accountability concerns moral wrongdoing and actions, as opposed to attributability which primarily concerns character and actions only insofar they reveal something about one's character.

A fundamental difference between accountability and attributability is the kind of control they require for moral blameworthiness. Accountability is typically taken to have rather strict control conditions; attributability is not. Control is a central issue for our practices of holding people accountable, but not, it seems, for the justification of aretaic appraisals.

These different control conditions are one important reason why the distinction between accountability and attributability plays a useful and important role in making sense of our responsibility practices. The distinction allows us to explain the ambivalence we feel towards people who seem to be responsible in one respect but not in another. It is controversial whether we are accountability responsible for our emotional reactions, lack of attention, and moral ignorance. In all these cases it may seem as if we lack control. However, it seems plausible that we can be attributability responsible for such states and omissions, as they often reveal something about our aims, values and characters. The distinction can also make sense of our ambivalence towards agents with psychological deficiencies, such as psychopaths (Watson 2011), people with autism, mild intellectual disability, poor formative circumstances and dementia (Shoemaker 2015). The distinction provides resources to philosophers who are skeptical of whether *anyone* can satisfy the control conditions of accountability. It allows them to explain why their skepticism (concerning one kind of responsibility) does not mean that we have to give up *all* of our responsibility practices. Moreover, the notion of attributability helps illuminate the plausible idea that we can also be responsible for actions that do not fall squarely into the moral domain.

Why do these two kinds of responsibility have different control conditions? The most straightforward explanation goes as follows: accountability involves targeting the wrongdoer with the reactive attitudes - resentment or indignation. When these hostile emotions are expressed, they can be viewed as sanctions or adverse treatment (Wallace 1994; Watson 2004; Rosen 2002; Nelkin 2011).

Because of this prospect of pain and adverse treatment, accountability raises questions of desert. The wrongdoer must meet certain control criteria in order for accountability blame to be deserved.²

Similar concerns do not, it is often argued, arise for attributability. The kind of blame involved in attributability is taken to be milder than the one involved in accountability. Watson describes a case of a colleague who fails to return a book. This conduct might be called shoddy; we see the conduct “as ‘inferior goods’, as poor exercise of human evaluative capacities, as characteristic of someone who cares little about the standards of excellence in human affairs” (Watson 2004: 265). By calling this conduct shoddy one *is* blaming the colleague, Watson maintains. But this kind of blame - let us call it *attributability blame* - does not presuppose that the agent “deserves adverse treatment or ‘negative attitudes’ in response to their faulty conduct” (Watson 2004: 266).

This standard explanation of accountability’s and attributability’s different control conditions needs to be revised. Being blamed by others in the accountability sense is not a sanction and need not be painful to the wrongdoer. However, to blame oneself - to feel guilty - is necessarily painful. I will argue that it is the painfulness of guilt that provides the most promising way of justifying accountability’s control conditions. This explanation prompts the question of what it is to blame oneself in the attributability sense.

I will argue that the best candidate for self-blame in the attributability sense is *shame*. This raises a puzzle. *Both* guilt and shame are intrinsically painful. Hence, appealing merely to the different forms of blame involved in accountability and attributability cannot explain their different control conditions. I will then propose a solution to this puzzle. Accountability and attributability are not merely distinguished by two different forms of self-blame; these forms of blame are also governed by different norms of appropriateness. An agent S is accountability blameworthy for X only if S *deserves* to feel guilty for X; an agent S is attributability blameworthy for X only if it is *fitting* that S feels shame for X. The painfulness of these emotions, I will argue, plays different roles, depending on

² Nelkin (2015) and Watson (2011) give somewhat different explanations of the different control criteria, based on the kind of demands associated with in accountability blame but not with attributability blame.

whether we ask if they are fitting or deserved. Control is relevant for the question of whether a painful emotion is deserved, but not for the question about its fittingness.

The paper is structured as follows. Section 2 presents the painfulness of self-blame as the basis for an argument for accountability's control conditions. In section 3 and 4 I argue that to blame oneself in the attributability sense is to feel shame. Section 5 argues that shame is appropriate when it is fitting. Section 6 explores the differences between fittingness and desert. In Section 7, I argue that accountability is governed by desert, whereas attributability is governed by fittingness. Section 8 concludes.

2. Accountability and Control

Accountability blame, as we have seen, is understood in terms of the reactive attitudes resentment and indignation.³ This has led to the following, commonly endorsed argument for control.⁴

The interpersonal argument for accountability control

- (1) An agent S is blameworthy for X in the accountability sense only if S deserves to be targeted with the reactive attitudes resentment or indignation because of X.
- (2) To be targeted with the reactive attitudes resentment or indignation is to suffer.
- (3) S deserves to suffer for X only if X was under S's control.

Therefore

- (4) S is blameworthy for X in the accountability sense only if X was under S's control.

³ This section draws on material from Carlsson (2017).

⁴ For versions of this argument in the literature see Wallace (1994); Watson (1996); Nelkin (2011) and Pereboom (2014). Some of the versions use fairness instead of desert, but my objections could be applied equally well to that version of the argument.

Is premise (2) true? That depends on how we understand the notion of blame. To “target” someone with resentment or indignation could be understood either as *experiencing* or *expressing* these emotions. Blame can be private or outwardly expressed. If blame is identified with the reactive attitudes, it seems natural to understand blameworthiness in terms of the appropriateness of *experiencing* these emotions. After all, there seem to be numerous characteristic instances of blame that are not expressed. We blame people from afar and we blame the dead (Nelkin ms). One might argue that private blame is the more fundamental notion, since it is possible to experience resentment and indignation without expressing these sentiments. The possibility of private blame suggests that the expression of reactive attitudes, in contrast to experiencing the attitudes, is not a necessary part of what blame is. But if private blame is used in our conception of blameworthiness it is not obvious why control should be a necessary condition. Merely experiencing the reactive attitudes does not amount to sanctioning the wrongdoer. Feeling certain emotions is not a form of adverse treatment (Graham 2014: 391). This gives us reason to be skeptical of premise (2). If blame fundamentally is to experience an emotion, it is unclear how this relates to the suffering of the blamed agent.

Watson and R. J. Wallace emphasize that reactive attitudes come with a *disposition* to sanctioning behaviour. Watson claims that “blaming attitudes involve a readiness to adverse treatment” (2004: 275). Wallace argues that blame “involves a disposition to engage in a variety of sanctioning activities” (1994: 94). This might very well be true as an empirical fact, but there is no necessary connection between blame and sanctions. Even though having the blame emotions involves a disposition to treat the wrongdoer in a certain way, this disposition need not be manifested (Nelkin 2013: 124; Graham 2014: 391). Moreover, even when the blame emotions are expressed, they will not always be experienced as sanctions. Sometimes we do not care about the disapproval of others. We can also imagine cases in which the wrongdoer might appreciate being the object of expressed blame. Consider a group of bullies who challenge each other to find the most effective way of humiliating their victim. When the victim expresses his resentment towards one of them after a particularly

ingenious humiliation, the bully might not experience any harm at all, but rather pride as well as recognition from his peers.⁵ Premise (2), it seems, is false.

Moreover, if premise (2) is false we will have reasons to doubt premise (1) as well. Premise (1) states that agents are blameworthy in the accountability sense only if they *deserve* to be targeted with blame. However, if blame is not essentially harmful to the agent being blamed, it is unclear why the relevant notion of appropriateness should be a moral notion such as desert. Desert, in the context of responsibility, concerns ways in which agents could be harmed or benefitted (Feinberg 1970; Nelkin 2016). If blame is not essentially harmful it is unclear how it could be deserved.⁶ The fact that blame need not be expressed is one of the main reasons for why many philosophers have recently argued that the relevant notion of appropriateness is fittingness and not desert. Graham (2014) claims that because blame as such is not a form of adverse treatment at all, we have no reasons to assume that the relevant notion of appropriateness is a moral notion like fairness or desert. Similarly, David Shoemaker (2015 221-223) draws a sharp distinction between the question of the appropriateness of the blaming emotion and the appropriateness of its expression. The former question is matter of *fittingness*: whether the emotion correctly appraises its object. The latter is a matter of *ethics*: whether the harsh treatment is fair or deserved. Because expressed blame can be harmful to its recipients, it is plausibly governed by moral norms. But this does not, according to Shoemaker, mean that blameworthiness *itself* is governed by desert: blame as such is not harmful and it is the appropriateness of blame, not its expression, that determines whether an agent is blameworthy.⁷

⁵ See Nelkin (2013: 124) for a similar example.

⁶ When blame is not expressed the blamee will not suffer as a result. But perhaps one could argue that it is always *harmful* for an agent to be blamed, even though he is unaware of being blamed? I do not think so. This kind of unacknowledged blame may perhaps *sometimes* be harmful, for example if the blamee cares about how the blamer feels about him. But even if we grant that there can be cases like this, silent blame is often not harmful. If the blamee does not care about how the blamer feels or thinks about him, it is hard to see how he could be harmed by unexpressed emotions of blame. Thanks to Daniel Telech and Gideon Rosen for discussion.

⁷ Both Graham (2014) and Shoemaker (2015) take *experienced* blame to be the fundamental notion. One way to preserve the connection between desert and the harms of blame is to argue that it is expressed blame that is basic, and that being blameworthy is a matter of deserving to be the target of expressed blame (McKenna 2012; Pereboom 2013). As I argued above, however, taking expressed blame is the fundamental notion, however, is problematic. For further criticism of understanding accountability in terms of deserved, outwardly expressed blame, see Carlsson (2017).

I have pointed to two interconnected problems for the interpersonal argument for accountability control: blame is not essentially harmful and we therefore have reasons to doubt that blame is governed by desert. There is, however, another way of arguing for the control condition on responsibility as accountability. It is a crucial, but often overlooked, desideratum of any adequate theory of blameworthiness to incorporate self-blame in its definition: An agent is blameworthy not only to the extent that it is appropriate that she is blamed, but also to the extent that it is appropriate that she blames herself.⁸ There is no necessary connection between other-directed blame and suffering. The resentment and indignation may not be expressed, or the wrongdoer might not care. But there is a necessary connection between blaming oneself and suffering. To blame oneself in the accountability sense is to feel guilt. Feeling guilty is intrinsically painful (Clarke 2013: 155). In contrast to resentment or indignation, 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. This provides the material for another, *intrapersonal*, argument for why control is required for accountability blameworthiness:

The intrapersonal argument for accountability control:

- (1) An agent S is accountability blameworthy for X only if S deserves to feel guilt for X.
- (2) To feel guilt is to suffer.
- (3) S deserves to suffer for X only if X was under S's control.

Therefore,

- (4) S is accountability blameworthy for X only if X was under S's control.

⁸ In Carlsson (2017) I argue that being blameworthy (in the accountability sense) most fundamentally is to deserve guilt. In this paper I only argue for the claim that deserving guilt is a necessary condition on blameworthiness in the accountability sense. See also Clarke 2013; 2016.

My claim is that this argument provides a more straightforward way of justifying the control condition on accountability. It may be somewhat surprising to argue for the control condition via self - blame. Normally the control condition is explained by reference to the harmfulness of blaming *other* people, not the harmfulness of blaming oneself. But the shift is not as radical as it may seem. After all, other-directed blame and self-blame are closely connected. First, it is uncontroversial that blameworthiness does not only depend on whether it is appropriate to feel resentment or indignation, but also of whether it is appropriate to feel guilt. Moreover, if it is appropriate for me to blame you for X, it would also be appropriate for you to blame yourself for X. Hence, we should not only be concerned about the harms of blaming other people, but also about the harm of self-blame that is entailed by our other-directed blame.⁹

One might also object that this way of explaining the control conditions of accountability fails to acknowledge an important asymmetry between the way I am allowed to treat myself and the way I am allowed treat others. We may do things that cause ourselves pain, even though we are not allowed to cause pain to others.¹⁰ Given this asymmetry one might argue as follows: Since guilt is a form of blame, perhaps I might be justified in blaming myself, even in cases where it would not be justified to blame others? If that is correct, focusing on self-blame might be taken to provide *weaker* control conditions than focusing on other-directed blame. Note, however, that the intrapersonal argument for control relies on a specific form of justification, namely desert. We can think of many ways in which we may be *justified* to blame ourselves even if we would not be justified in blaming others. It may for instance be commendable to hold oneself to a higher standard. We are also often uncertain about whether a wrongdoer displayed ill will or whether he knew that what he was doing was wrong. There will typically be less epistemic uncertainty concerning one's own mental states. In these cases I might be justified in feeling guilty, but not blaming others or causing them to feel guilt. However, these

⁹ It is commonly observed that expressed blame aims at guilt. If the agent being blamed does not feel in any way bad about his action, both the communicative and the retributive purpose of blame seems to be thwarted. When blaming other people, we typically want them to *painfully recognize* what they have done (Clarke 2013; Shoemaker 2015; Rosen 2015; Macnamara 2015; Fricker 2016; Mason 2019).

¹⁰ I thank an anonymous reviewer for raising this issue

considerations do not bear on the question of whether I *deserve* to feel guilty. The asymmetry between what I am allowed to do to myself and what I am allowed to do to others is relevant to the question of whether agents ought to feel guilty, but it is not relevant to the question of whether they are blameworthy.¹¹

The lesson I want to draw from this section is this: It is blaming *oneself* - by experiencing guilt - that provides the most direct explanation for the control conditions on responsibility as accountability. If we want to uphold that there are weaker control conditions on attributability, we need to give an account of what self-blame is in the attributability sense in order to make sure that a parallel argument cannot be made for that kind of responsibility.

3. Attributability and Self-blame

What is it to blame oneself in the attributability sense? So far this question has received little attention.¹² Let me start by considering recent discussions of *other-directed* attributability blame. Attributability blame can be understood in two different ways, either as emotional responses or as judgments. Given that we understand accountability blame in terms of the reactive attitudes resentment, indignation and guilt one natural idea is understand attributability blame in terms of *other* reactive attitudes (Eshleman 2004; Shoemaker 2015). Certain negative reactive attitudes such as disdain or contempt are clearly instances of aretaic appraisals; they are

¹¹ Duggan (2018) argues that the conditions of deserved other-directed blame are stricter than the conditions of deserved guilt. He claims for example that agents can deserve guilt but not blame for suberogatory actions. However, he arrives that conclusion by weakening the notion of guilt and strengthening the notion of blame. Blame, according to Duggan, is characterized by a desire to alienate the wrongdoer from the moral community, and for the blamer to seek repentance from the blamed. This seems to be an overly robust notion of blame. Duggan also understands guilt not as a form of blame, but rather as the “sorrow in virtue of morally disapproving of oneself”. I find this characterization unconvincing. Sorrow and guilt seem to be quite distinct emotions.

¹² Eshleman (2004) is a notable exception. He takes self-blame in the attributability sense to be shame, a view I will endorse and develop in the next section. Shoemaker (2015: 26) mentions shame as one candidate for a self-directed responsibility response for attributability, but also suggests that shame may be fitting for accountability and answerability “as it can take as its object qualities of character, judgment and regard at different occasions.”

negative responses to an agent's character.¹³

Watson, by contrast, takes attributability blame to be an aretaic *judgment*. To blame someone in the attributability sense is to judge that the agent has failed with respect to some standard of excellence and that this failure reflects what she cares about or what she stands for. Watson suggests that this judgement can be made dispassionately: this kind of blame does not require that one is emotionally exercised in any way. Thus, one natural thought would be to interpret self-blame in the same way. To blame oneself in the attributability sense would be to make a judgment that one's own action has failed some standard of excellence, and that this action reflects one's own character, cares or aims. In this section I will discuss and reject this proposal.

Consider some of the common objections to a cognitive account of other-directed blame. First, if blame can be a dispassionate judgment, it will be difficult to explain the plausible distinction between blaming someone and judging them to be blameworthy. These two states can come apart. It seems possible to judge someone to be blameworthy without actually blaming them. I might judge that Napoleon was blameworthy for invading Russia, but it is quite unclear in what sense I actually blame him. If a good friend of mine did something wrong, cheated on his taxes, say, I might very well judge him to be blameworthy, without actually blaming him. If this is right, a cognitive judgment is not sufficient for blame.¹⁴

Another common objection against cognitivist accounts of blame is that it is implausible to think of blame "as though in blaming we were mainly moral clerks, recording moral faults [...] from a detached and austerely 'objective' standpoint" (Watson 2004: 226–27). Blame, it seems, requires an emotional involvement that goes beyond mere registering non - excused moral fault. These arguments are typically raised against cognitivist versions of accountability blame. But I do not see any

¹³ For different accounts of aretaic reactive attitudes see Eshleman (2004); Bell (2013); Shoemaker 2015 and Darwall (2016)

¹⁴ Nor does it seem necessary. It seems possible to blame someone without believing them to be blameworthy. See Pickard (2013) for other-directed blame, and Wallace (1994) and Gibbard (1990) for self-directed blame.

principled reasons for why we should not want to maintain the distinction between blaming and judging blameworthy also for responsibility as attributability.

I believe this point becomes particularly vivid once we move from other-directed blame to self-blame. There is a difference between registering or acknowledging that one is blameworthy and actively blaming oneself. If I judge in a cool and dispassionate way that my action was bad (and reflected badly on my character), but did not experience any negative emotion whatsoever, something about blame seems to be lacking. Consider the following case:

The rival: A colleague and rival is being humiliated by a number of crushing comments during a philosophy talk. As you observe him blushing and stuttering, a warm feeling of satisfaction wells up inside you. You are not in control of this emotion, so let us assume that accountability blame is not at issue. On reflection, it becomes clear to you that your reaction does reveal a character trait of yours. You have noticed several other examples of your small-mindedness. You also take your emotional reaction to be a failure to live up to the standard of excellence to which you hold yourself. Suppose, however, that you do not feel bad because of your reaction.

There are two natural responses to this scenario. First, one might be skeptical of whether you actually believe that your response was wrong, or reflected badly on your character. Second, let us suppose that your beliefs are genuine. If so, it seems that you believe that you are blameworthy and it would be appropriate for others to make the same aretaic appraisal of your character. However, if you do not feel in any way bad about your reaction, it does not seem as if you are *blaming* yourself. This suggests the following conclusion: Any kind of self-directed blame, attributability blame included, requires at least some negative affect. A pure cognitive account for self-blame in the attributability sense is not plausible.

4. Attributability and Shame

How we should then understand self-directed blame in the attributability sense? In what follows I argue that we should understand it as *shame*. Shame along with guilt, embarrassment and pride

belongs to the class of self-reflective emotions “that are evoked by self-reflection and self-evaluation” (Tangney and Stuewig 2007: 347). In the psychological literature there have been several attempts to distinguish shame from guilt. One approach draws on the distinction between private and public. On this view, shame arises from public exposure and is an essentially social emotion. Whereas one can feel guilty in private, based on one’s own negative evaluation of an act, shame requires an audience. However, even though *embarrassment* requires the presence of others, there seem to be little evidence that this also is true for shame. Tangney et al. (1994) found that shame and guilt are equally likely to be experienced in the presence of others and that solitary experiences of shame were about as common as solitary guilt experiences.

Another approach focuses on the events that elicit shame and guilt. But it seems that most types of events that can elicit guilt can also elicit shame. One can feel guilty or ashamed for shoplifting, or for lying, or for breaking a promise. The object of the emotion, however, will be different. Whereas guilt typically concerns some specific behaviour, shame emotions involve a negative evaluation of one’s global self. This is the most empirically robust and commonly accepted way of distinguishing shame and guilt. (Lewis 1972; Tangney and Stuewig 2007; Deonna et al 2012). One feels shame not for what one does, but rather for whom one is. If I feel guilty for telling a lie, it will be because I recognize that I did something wrong. If I feel ashamed for telling the lie, it will be because this reflects badly on whom I have revealed myself to be (Tangney and Stuewig 2007; Mason 2010). Shame’s more comprehensive negative assessment of the self is reflected in its characteristic phenomenology, often reported as a feeling of worthlessness and incapacity, feeling small, shrunken or exposed (Tangney and Stuewig 2007; Deonna et al. 2012), as well as its characteristic action tendencies: withdrawal or escape from the shame inducing situation (as opposed to confessing, amending or apologizing for the wrongdoing in the case of guilt).

Shame has several features which make it the most plausible contender for self-blame in the attributability sense. Just as attributability concerns actions only insofar as it reveals the agent’s cares and commitments, shame is elicited by actions insofar as they are taken by the agent to reveal something about whom she is. Whereas we typically feel guilty for things we could have prevented, shame does not seem to require a sense of control (Prinz and Nichols 2010). We can feel ashamed of

unbidden thoughts or our lack of a sense of humour. Guilt is often taken to concern the transgressions of norms; I typically feel guilty when I believe that I did something morally wrong. Guilt and accountability are both concerned with the deontic: right and wrong, duty, obligation and permissibility (Darwall 2006; 2016). Shame, on the other hand is typically experienced as the failure to exemplify some particular value pertaining to oneself (Deonna et al 2012). When one feels ashamed for a stupid remark, for the creation of a lousy song, or for failing to speak up to a bullying boss, one is pained not for acting wrongly, but for a failure to live up to certain values or ideals. Stephen Darwall (2016:17) remarks that “[a]ttributions of virtue and vice concern how *estimable* someone is; they call on attitudes of esteem and disesteem.” Shame is a way of feeling disesteem for oneself. It is commonly taken to be an emotion of self-assessment that is appropriate as a response to one’s failure of character (Rawls 1999: 388; Taylor 1985: 64; Mason 2010: 418; Deonna et. al., 2012: 177). These characterizations of shame fit nicely with several important features of responsibility as attributability: we are attributability responsible for our character and for what we care about. Shame, I will argue, is appropriate when our standards, characters and cares are flawed.

In order to test my view that self-blame in the attributability sense is best captured as experiencing shame, let us consider behaviour for which we can be responsible in the attributability sense, but for which is it controversial whether we are accountable: emotional reactions, unwitting omissions, and actions from moral ignorance. Consider *The Rival* from above. You are attributability responsible for your warm feelings of satisfaction at your rival’s humiliation. Given that you are not in control of your emotions, it is debatable whether you deserve to feel guilty. But it seems perfectly appropriate to feel ashamed of your emotions. Consider two more examples:

Birthday: Angela forgot a close friend’s birthday. A few days after the fact, she realized that this important date had come and gone without her so much as sending a card or giving the friend a call. Angela was mortified. What kind of a friend could forget such a thing? (Adapted from Smith 2005: 236)

Upbringing: Abigail, who is raised in a family or community that is deeply racist, say, or religiously intolerant. It would not be at all surprising for Abigail to develop evaluative tendencies and corresponding attitudes in line with those she sees operative in her family or surrounding community. As an adult, her attitudes may

continue to reflect the vicious evaluative judgments thus formed in her childhood. (Smith 2005: 267)

Since Angela is not in direct control of her unwitting omissions, it is debatable whether she is accountable for them. However, it seems perfectly appropriate for Angela to feel ashamed of her omission if it reflects her lack of concern for her friend. Similarly, given her upbringing it is debatable whether Abigail is accountable for her attitudes and whether she deserves the pain of guilt for what she did. Nevertheless, it seems appropriate for Abigail to feel ashamed of her attitudes.

To blame oneself in the attributability sense, I have argued, is to feel shame. This gives us the following necessary condition on responsibility as attributability:

Attributability: An agent S is blameworthy for X in the attributability sense only if it is appropriate for S to feel shame for X.

Attributability and accountability are thus distinguished by two different forms of self-blame: to blame oneself in the accountability sense is to feel guilt; to blame oneself in the attributability sense is to feel shame. But now it becomes unclear how we can explain why accountability has stricter control conditions than attributability. Above I argued that the best explanation for why accountability requires control starts with the fact that guilt is intrinsically painful. However, shame *necessarily involves suffering*, just as guilt does. Indeed, Tangney and Dearing (2007: 349) report that feelings of shame are typically experienced as being *more* painful than guilt. It is not merely one's behaviour that is at stake, but one's core self. This pain will of course come in different degrees. The pain of shame can vary from discomfort to a prolonged state of agony. But if the emotional state does not involve suffering at all, it is not shame. In this respect, shame is similar to guilt. This provides the material for a parallel, intrapersonal, argument for why control is required also for *attributability* blameworthiness:

The intrapersonal argument for attributability control

(1) An agent S is attributability blameworthy for X only if S deserves to feel shame for X.

(2) To feel shame is to suffer.

(3) S deserves to suffer for X only if X was under S's control.

Therefore,

(4) S is attributability blameworthy for X only if X was under S's control.

Something has clearly gone wrong here. If this argument is sound, *both* accountability and attributability require control. I have argued that self-blame, both in the accountability and in the attributability sense, entails suffering. If accountability and attributability have different control conditions, the explanation cannot merely be a difference in the harm of blame. In the remainder of this paper I will attempt to explain the different control conditions by appealing to two different norms of appropriateness of emotions. Accountability, I will argue, is governed by desert, whereas attributability is governed by fittingness. Once we draw this distinction, we can see why the intrapersonal argument for *attributability* control fails, whereas the intrapersonal argument for *accountability* control succeeds.

5. Fitting Shame

In this section I will argue that an agent is blameworthy in the attributability sense only if it is fitting that she feels shame. Fittingness is a distinct normative property that applies to attitudes, such as emotions, desires, beliefs and intentions. The property is understood in different ways.¹⁵ I will assume, as is increasingly common in the literature on responsibility, that the fittingness of emotions is a

¹⁵ For a good overview of different ways of understanding fittingness, see McHugh and Way (2016) and Howard 2018.

matter of *correct representations* (D'Arms and Jacobson 2000; Graham 2014; Rosen 2015; Shoemaker 2015).¹⁶ Emotions represent their object as having certain properties.¹⁷ Fear represents its object as being dangerous or threatening. Envy, roughly, portrays one's rival as having something that one lacks, and casts this circumstance in a negative light. (D'Arms and Jacobson 2000: 66). Regret represents some action of yours as a mistake. Emotions are fitting when these representations are correct or accurate. Fear of tigers is fitting; fear of kittens is not. It is not fitting to envy one's rival for being wealthier than you when he is in fact approaching bankruptcy; and it is not fitting to regret something that was not a mistake.

What does shame represent? One possibility is to take the representational content of shame to be exclusively about one's failures of character. According to Gabrielle Taylor (1985: 64), shame represents the agent as "not the sort of person she believed, assumed, or hoped she was or anyway should be." On Michelle Mason's (2010: 417-418) view, "to experience shame is to experience oneself as diminished in merited esteem on the grounds that one has violated a legitimate ideal of the self."¹⁸ The above views would fit very well with the account of shame as the relevant kind of self-blame in the attributability sense: *Shame is fitting when the agent has failed to live up to an aretaic standard*. These views provide a thin, but explanatory characterization. A striking feature of shame is that different people experience different things as shameful (Gibbard 1990:137). This characterization explains different experiences of shamefulness by people's different ideals of the self. It also allows for the fact that people can be mistaken about what is shameful. When Huckleberry Finn finds himself psychologically unable to turn in Jim, the slave he has befriended and helped to escape, he feels

¹⁶ However, both Graham and Rosen speak of correctness or truth instead of fittingness. Rosen explicitly distinguish fittingness (as a sui generis concept) from correctness.

¹⁷ Accepting this idea does not commit one to cognitivism about emotions. The representational content need not be judgments or beliefs. Emotions may be recalcitrant. I may be afraid of ducks, while at the same time believing that ducks are not dangerous. My fear will nevertheless represent the ducks as dangerous: it will seem to me, or strike me that they are dangerous, even though I believe they are not, just as it may seem to me that the lines in the Müller-Lyer illusion differ in length, even though I believe they are equally long (D'Arms and Jacobson 2000: 67; Rosen 2015: 72).

¹⁸ Deonna et. al (2012: 102) take shame to represent "a trait or an action of ours that we take to exemplify the polar opposite of a self-relevant value as indicating our incapacity to exemplify [a]self-relevant value even to a minimal degree."

ashamed of himself. His shame is understandable given what he took to be a legitimate ideal of a person, yet it is not *shameful* because Huckleberry is mistaken about that ideal. Correspondingly, Abigail who was raised in a racist society might not be ashamed of her attitudes, but it would nevertheless be fitting if she were.¹⁹ On this view, it would also be fitting to feel ashamed for one's emotions in *The Rival*, or for one's unwitting omissions in *Birthday*, because these examples exemplify failures to live up to aretaic standard.

However, we also experience shame for features that do not reveal anything about our character. People are sometimes ashamed of their looks, their physical incapacities, their family or their country. This raises a problem for a view that takes the representational content of shame to be exclusively about character. First, it is hard to see how shame for one's family or country can be construed as being about the person one wants to be, or what one takes to be legitimate ideal of the self. As a result, it seems that both Taylor and Mason would need to deny that such experiences really are instances of shame. This is problematic. Whereas it seems plausible to deny that such experiences are fitting, it is much harder to deny that they are (perhaps unfitting) instances of shame. In order to accommodate the wide range of cases in which people experience shame, we could instead adopt a thinner representational content: What shame represents is simply that *something reflects badly on the agent*. This would allow for shame about one's appearances, physical incapacities, family or country to count as genuine instances of shame. The question of whether these kinds of shame are fitting would then depend on whether such non-agential features in fact reflect badly on the agent.

This wider notion of shame's representational content is compatible with the view I am arguing for in this paper. My claim is that fitting shame is a *necessary* condition on being blameworthy in the attributability sense. Failures of character do reflect badly on agents, and would therefore make shame fitting. If the wider notion of shame is correct, there might be instances of fitting non-agential shame (say for one's physical incapacities) that would not make one blameworthy

¹⁹ Consider the reproach, "You should feel ashamed of yourself!" One plausible interpretation would be to say that aim of the reproach is to direct the person's attention to an unrecognized flaw—i.e., that their apparent absence of shame is taken as evidence that they do not see their flaw. I thank an anonymous Reader for pointing this out to me

in the attributability sense. Given the plausibility of a wider notion of shame we could think of the agential shame proposed by Mason and Taylor as a *cognitively sharpened emotion*. D'Arms and Jacobson (2003) define a cognitively sharpened emotion as a subclass of an emotion, specified by the thought these instances of the emotion share. Resentment, on their view, is a subclass of anger, unified by the thought that one was wronged. Similarly, agential shame might be a subclass of shame, unified by the thought that this reflects poorly *on my character*. It is the fittingness of this kind of shame, I argue, that is a necessary condition of blameworthiness in the attributability sense.²⁰

It seems plausible that fittingness is the notion of appropriateness that governs attributability (Shoemaker 2015; Darwall 2016). Whether an agent is responsible in the attributability sense depends on whether her action *in fact* revealed a failure of character, or whether the faulty character trait *in fact* is attributable to her. Aretaic appraisals of ourselves as lazy, cowardly or mean are appropriate when their objects actually manifest these character traits. This is a matter of correct representations. If we understand self-blame in the attributability sense as shame, what matters for attributability is whether the representational content of this emotion is correct, i.e. whether it is fitting. This gives us the following necessary condition on attributability:

Attributability: an agent S is attributability blameworthy for X only if it is *fitting* that S feels shame for X.

We can now explain what went wrong in the intrapersonal argument for attributability control. Once we substitute fittingness for desert, we cannot argue from suffering to control. Recall premise 3 from that argument: S *deserves* to suffer for X only if X was under S's control. This is a plausible claim. But a parallel claim about *fitting* suffering would not be plausible. Grief, regret, and embarrassment are all, like shame, intrinsically painful. Yet these emotions do not require a control condition in order

²⁰ Appealing to cognitively sharpened emotions might be somewhat circular if one attempts to give an analysis of attributability, since the focus on character is used to specify the emotion. But I am not attempting such an analysis. I am merely arguing for one necessary condition on blameworthiness in the attributability sense.

for them to be appropriate. This is because these emotions are most plausibly governed by fittingness, and not desert.

The painfulness of an emotion is *one* kind of reason not to feel that emotion. But it does not seem to have any bearing on whether it would be *fitting* to feel it. The fact that shame is painful does not have any bearing on whether the action was shameful. In this sense, the painfulness of shame is the *wrong kind of reason* (D'Arms and Jacobson 2000). What matters for the fittingness of agential shame is only whether the agent in fact failed to live up to an aretaic standard. And this, as I argued above, does not require a control condition.

6. Fittingness and Desert

Above I argued that attributability is governed by fittingness. In the next section I will argue that accountability is governed by desert. The purpose of this section is to explain the difference between these two kinds of appropriateness. Desert and fittingness are distinct normative properties. It can be fitting to feel grief when a loved one dies, envy at your rival's success, or amusement at a funny joke, but these reactions are not *deserved*.²¹ However, it is not easy to explain exactly what the difference between fittingness and desert consists in.

As Joel Feinberg (1970) observes, we sometimes speak of artworks deserving admiration, or problems deserving consideration.²² I will follow Feinberg in limiting my discussion to personal desert.²³ The desert relation is typically taken to be a three-place relation between a person, a desert basis (the things for which the person is deserving) and what is deserved (Feinberg 1970). For example, I might deserve to feel guilty in virtue of having acted knowingly, freely and wrongly.

²¹ "Fittingness" is sometimes used interchangeably with "merit" (King 2012; Shoemaker 2015). But perhaps the relation picked out by "merit" is identical, or at least close enough, to the one picked out by desert? I believe that merit and desert are distinct notions. Howard (2018) gives the following examples of how "merit" and "desert" comes apart: a proposition might merit, but never deserve belief. Someone might merit admiration in virtue of his good looks, but it would be strange to think that he deserves admiration in virtue of this quality.

²² This kind of usage seems closer to the examples of merit I mentioned in the previous footnote.

²³ As does Nelkin (2016) and Clarke (2016). As Nelkin points out, this notion also seems to capture what Pereboom has in mind when he talks about "basic desert."

Derk Pereboom (2017: 123) characterizes desert as follows:

The desert at issue here is basic in the sense that the agent, to be morally responsible, would deserve to be blamed or praised just because she has performed the action, given sensitivity to its moral status; and not, for example, by virtue of consequentialist or contractualist considerations.²⁴

I think this is correct, but it does not suffice to distinguish fittingness from desert. Pereboom's characterization is mainly negative: blame is deserved simply by virtue of the wrongdoer's action and not by virtue of any other moral justifications. But the same would be true of fitting blame, so this quote would be compatible with identifying desert with fittingness. If blame is fitting, it is fitting just because the agent has performed the action, and not by virtue of consequentialist or contractualist considerations.

Michael McKenna (2012) adds a plausible positive feature to Pereboom's negative characterization. Although desert is a basic concept, which cannot be analysed in terms of or reduced to other concepts, we can say something about what is *entailed* by desert. Desert, according to McKenna, entails noninstrumental goodness.²⁵ If an agent deserves some harm, then it will be noninstrumentally good that the agent suffers that harm.²⁶ This seems to be a more promising way of distinguishing desert from fittingness. It may be fitting to be amused by a racist joke, to be envious of one's rival, or to despair at a disaster. But it does not seem noninstrumentally good to have these emotions. Yet, it seems plausible that fitting emotions can *sometimes* be noninstrumentally good. To see why that is the case, consider Randolph Clarke's (2013) explanation of why it is noninstrumentally good that blameworthy wrongdoers feel *guilty*. First, feeling guilty is to acknowledge one's own moral

²⁴ See also Pereboom (2014: 2) and Scanlon (2013: 101) for similar formulations.

²⁵ I take noninstrumental value to cover both intrinsic value and extrinsic value. Something can be noninstrumentally valuable even if its value is partly extrinsic. Appreciation of a beautiful artwork for example can be noninstrumentally valuable, even if its value depends on the value of the artwork.

²⁶ See also McMahan (2009); Kagan (2012); Clarke (2016); Carlsson (2017) and Portmore (forthcoming).

fault. This involves recognizing a morally important fact about oneself and there is a value in this recognition. Second, the painful affect that is constitutive of feeling guilt expresses a valuable moral concern. However, the very same reasons for why experiencing deserved guilt is noninstrumentally valuable could also plausibly be marshalled for why experiencing *fitting agential shame* is noninstrumentally valuable. To feel fitting agential shame is to acknowledge that one has failed to live up to an aretaic standard. There is arguably noninstrumental value in this recognition. To experience the painful affect of shame expresses a valuable concern for one's character. Shamelessness, the failure to feel fitting agential shame, is plausibly disvaluable.²⁷

A similar argument could be made for other intrinsically painful emotions such as grief. Grief at the death of a loved one is very painful. Yet, it is also plausibly noninstrumentally good. If we compare two worlds, one in which people grieve their dead parents, and one in which they do not, the first does seem better, quite independently of any instrumental value attached to the grief.

Grief and shame are fitting reactions to deaths and failures of character. I have also suggested that they can be noninstrumentally valuable. Yet they are not *deserved*. At this point I would like to highlight two crucial differences between fittingness and desert. First: desert, but not fittingness, *entails* noninstrumental goodness. The relation between fittingness and noninstrumental value is contingent. For desert, I suggest, the connection is necessary. It is noninstrumentally good that people get what they deserve. If they deserve to feel guilt for a wrongdoing, it will be noninstrumentally good that they feel guilty.²⁸ Grief when a loved one dies, on the other hand, is both fitting and

²⁷ In fact, there seem to be a number of fitting responses, which are noninstrumentally good without being deserved. Consider Hurka's (2000) value-based theory of virtue. Hurka starts out with states of affairs that are either valuable or disvaluable. Pleasure, knowledge and achievement, according to Hurka, are noninstrumentally good whereas pain, false beliefs and failure in the pursuit of achievement are noninstrumentally bad. He then goes on to argue that the attitudes we have towards these valuable or disvaluable states of affairs can themselves be noninstrumentally good or bad. On Hurka's view, positive attitudes towards valuable states of affairs are valuable, negative attitudes towards disvaluable states of affairs are valuable, positive attitudes to disvaluable states of affairs are disvaluable, and negative attitudes to valuable states of affairs are disvaluable. Being happy about other people's happiness is valuable, whereas taking pleasure in other peoples suffering is disvaluable. It is valuable to feel sad when a loved one dies, but disvaluable to resent the achievement of others.

²⁸ This does not mean that people who deserve guilt will have a *decisive* reason to feel guilty. The noninstrumental goodness of desert only provides *pro tanto* reasons to feel guilty (Clarke 2016) and these reasons can be outweighed. Consider someone who, we stipulate, deserves

noninstrumentally good but the noninstrumental value of this emotion is *not explained* by its fittingness.²⁹ The fittingness of emotion consists in the correctness of its representations. But correctness in itself is normatively neutral.³⁰ Despair, rage or envy might be correct ways of representing the world, but nothing noninstrumentally valuable seems to *follow* from these fitting states. In the case of fitting grief, the value of these emotions are not explained by the mere correctness of the representation but rather by the fact that it is valuable to be pained by something that is disvaluable. This source of value is orthogonal to the general structure of fittingness: there can be fitting responses that are not valuable in this sense. Suppose that your rival has made some significant achievement. In this case it seems like a *positive* attitude towards the achievement is plausibly valuable for its own sake, whereas a negative attitude would be disvaluable for its own sake. Yet, envy towards your rival's achievement - definitely a negative attitude - can nevertheless be fitting.

The second difference between fittingness and desert is that desert, but not fittingness is plausibly seen as a *consideration of justice*. This point is made by Joel Feinberg (1970) and Randolph Clarke (2016). It may be fitting to feel grief or regret but it is not *just*.³¹ I have argued that such emotions may sometimes be noninstrumentally good: a world in which we feel grief over our dead parents or regret over our serious mistakes is arguably better than a world in which we do not. Yet the world would not be more just.

to feel guilty, but who has also suffered a lot of hardships. I would maintain that it is noninstrumentally good that he deserves guilt, but this does not mean that he ought all things considered to feel guilt. Thanks to Robert Hartman and David Shoemaker for raising this issue.

²⁹ See Portmore (forthcoming) ch. 2, page 9 for a similar claim.

³⁰ One might argue that correctness is noninstrumentally good, if it is noninstrumentally good to have true beliefs. But whereas there is value in having certain kinds of true beliefs, it is less plausible that this is true in general. So fittingness does entail noninstrumental goodness. If we compare two worlds, one in which I have the true belief that my bowl of cereals contain 234 grain and one in which I lack this belief, it is by no means obvious that the former world is noninstrumentally better.

³¹ However, in some cases fittingness might be relevant to *distributive* justice. Suppose you admire someone who is not admirable instead of someone who is. This might be viewed as unjust in a distributive sense. Thanks to Anthony Duff and David Shoemaker for raising this issue.

7. Deserved Guilt, Fitting Shame and Control

I have suggested the following necessary condition on blameworthiness in the accountability sense:

Accountability: an agent S is accountability blameworthy for X only if S *deserves* to feel guilt for X.

In this section I will argue that accountability is governed by desert. Above I argued that attributability is governed by fittingness. Why think that accountability is governed by a different notion of appropriateness? Accountability is traditionally taken to concern the appropriateness *of holding* agents responsible (Watson 2004; Wallace 1994; Nelkin 2011; McKenna 2012). This requires answering distinctively moral questions concerning the justification of the harm related to *expressions* of the reactive attitudes. When the reactive attitudes are expressed, they will take the form of reproach, scolding, remonstrance, etc. These kinds of harsh treatment can be harmful and are plausibly governed by desert.

It might be less obvious why desert should govern accountability once we think of blame as *experiencing* reactive attitudes, rather than expressing them. As noted in section 2, both Graham (2014) and Shoemaker (2015) argue that because blame in the accountability sense is fundamentally an emotional reaction, and those are not as such harmful, it should be governed by fittingness rather than desert. Neither of them considers the harmfulness of guilt, but they could perhaps argue as I have done for the fittingness of shame: guilt is necessarily painful, but this pain is not relevant to whether guilt is appropriate or not. I think this would be a mistake. The pain of guilt, I will argue, is *crucial* for the appropriateness of accountability, in a way that is different from the appropriateness of attributability. We can use this difference to explain the different control conditions of these kinds of responsibility.

Emotions consist of both representational contents and of affects. They can be fitting in virtue of their representational content. After we have confirmed that an emotion represented the world correctly, we can still ask the question of whether the painfulness of this emotion is *deserved*. This is

not the case for fittingness: once the world is represented in the right way any further reason for or against having the emotion seems like the wrong kind of reason. Guilt also has a representational content, and this content can be correct or incorrect. Yet, independently of how we specify the representational content of guilt, we can still ask the further question of whether the painful affect of guilt is deserved. Suppose for example that the representational content of guilt were that the agent acted with ill (or insufficiently good) will (Graham 2014). Even if he did act with ill will, and the feeling is fitting, we can *still* ask the question of whether the agent *deserves* to be pained by the recognition of what he done, i.e. whether it is just and noninstrumentally good that he is pained by his guilt. The fact that guilt is painful is irrelevant to its fittingness but it is not irrelevant to whether guilt is *deserved or not*.³²

One way to illustrate this point is to consider the duration of deserved guilt. Blameworthy agents deserve to suffer the pain of guilt at the right time and to the right degree (Clarke 2013: 155). When I deserve to feel guilty, I deserve to be pained by certain amount of guilt. When I have suffered the right amount of guilt, I do no longer deserve to feel pain, because I have suffered enough: I have gotten what I deserved. The fact that I have suffered the right amount of pain of guilt, explains why I no longer deserve to feel guilt.³³ This is not the case for fittingness. Consider grief. The fact that I suffered a certain amount of grief does not make my grief any less fitting. What grief represents is not affected by my suffering, and fittingness is a matter of correct representations. The same goes for shame. The fact that I have already suffered the pain of shame, does not make shame less fitting.

³² Consider the following objection. I demand that you watch a documentary about the Holocaust because you continually underplay its importance in world history. You come away with a powerful sense of it, you acknowledge the horror and in so doing you feel pain: empathic pain, say. The fact that the acknowledgement is painful is irrelevant to its fittingness but it also appears irrelevant to whether that experience was deserved or not. We can understand this example in two different ways. If what you deserve is the pain of acknowledging the horror because of your moral insensitivity, it seems that both the degree and duration of the pain will matter to whether you deserve it. However, if what you deserve is simply the acknowledgement itself, the pain that it causes you will not matter, but this is because what you deserve is not the pain. Guilt is different in this respect because it is, in contrast to acknowledgement, partly constituted by painfulness. I thank David Shoemaker for raising this objection.

³³ This paragraph draws on Portmore (forthcoming: ch 2). He uses the point about the duration of guilt to argue for the claim that guilt is partly constituted by the thought that I deserve the unpleasantness of guilt.

The question of whether an agent deserves to feel guilty is not merely a question of correspondence or fit between what an emotion represents and the world but rather *a specific moral question*; namely whether the pain of guilt is just and noninstrumentally good in virtue of the agent's activities, circumstances and capacities.³⁴ By making this question central to whether agents are blameworthy we preserve the normative force of accountability, something it would be hard to account for if we were to understand blameworthiness merely in terms of correct emotional representations. As noted above, correctness itself, is normatively neutral. The normative relation between blameworthy wrongdoing and feeling guilty is *stronger* than that between a (racist) joke and amusement, or a rival's success and envy.

In answering the question of whether an agent deserves to feel guilt one can also apply all the issues traditionally taken to be relevant to responsibility as accountability. We can ask whether the wrongful action or omission was the product of the agent's history, whether the agent had a reasonable opportunity to avoid it, whether his action or omission was the product of bad luck, etc. We can also apply general moral principles, as the one I used in my intrapersonal argument for control: Nobody deserves to suffer for things they cannot control. If we treat the appropriateness of guilt as an issue of deserved suffering, we can explain why questions typically considered to be relevant for accountability are still relevant *even though* blame is essentially an emotion and not a sanction.

8. Concluding Remarks

I have argued for the following necessary requirements on accountability and attributability:

³⁴ What if desert were a part of the representational content of guilt? Douglas Portmore (forthcoming: ch. 2) argues that "To feel guilt, one must have the thought that one deserves to feel bad for what one has done—or, at least, one must experience one's feeling guilty about what one has done as in some respect noninstrumentally good". In that case guilt would only be fitting if it is deserved. This would be an elegant way of preserving the idea that accountability requires deserved guilt. However, there are several different competing accounts of the representational content of guilt in the literature (Taylor 1985; Greenspan 1995; Clarke 2016; Graham 2014). I would not like my account would be hostage to the precise cognitive content of guilt, which I find it hard to pin down with any confidence. Even if desert is not part of the representational content, I would still insist that whether an agent is responsible in the accountability sense depends on whether she deserves the pain of that emotion.

Accountability: an agent S is accountability blameworthy for X only if S *deserves* to feel guilt for X.

Attributability: an agent S is attributability blameworthy for X only if it is *fitting* that S feels shame for X.

My claim is that accountability is governed by desert whereas attributability is governed by fittingness. Because guilt is self-blame in the accountability sense, and shame is self-blame in the attributability sense, deserved guilt and fitting agential shame are the relevant notions for these two kinds of responsibility. This does not mean that guilt cannot be fitting or agential shame cannot be deserved. As an emotion, guilt can be fitting. But I have argued that fitting guilt is not *sufficient* for accountability. It is also conceivable that agential shame can be deserved. But deserved shame is not necessary for attributability. Moreover, given that we only deserve pain for things over which we have control, *deserved* shame, in contrast to fitting shame, would require that we have control over that which shame takes as its object, like character. It is possible that we in some instances can be in control of our character development. For such cases, however, we will typically also deserve guilt for some action or omission that led to the bad character.

Whether a person is responsible in the accountability sense depends on whether blame is deserved, i.e. whether it is just and noninstrumentally good that they are blamed. Whether a person is responsible in the attributability sense depends on a different question: on whether blame is fitting, i.e. whether our blaming emotions accurately represent them, or ourselves, as revealing some fault of character or vice. In the case of accountability we are concerned with the goodness and justice of harm or suffering. In the case of attributability we are concerned with the correctness of evaluations.

Why does accountability and attributability have different control conditions, given that both senses of responsibility necessarily involve suffering? The answer, I have argued, is that the painfulness of the self-blame plays two very different roles depending on whether we use desert or

fittingness as our notion of appropriateness. When we ask whether guilt is *deserved*, we are posing a moral question concerning deserved suffering. Whether the agent was in control of his behaviour is relevant to this question. Consider again the protagonists of *The Rival*, *Birthday* and *Upbringing*. Are they *accountable* for their emotional reactions, unwitting omissions and moral ignorance? I have argued that we should answer this question by considering whether they deserve the pain of guilt. Given the plausible moral principle that no one deserves to suffer for what they cannot control, one might argue that they do not deserve the pain of guilt. If so, they are not blameworthy in the accountability sense if they were not in control of their behaviour. The point here is not to settle this issue, but rather to illustrate that the painfulness of their guilt is relevant to the issue of whether they are accountable.

Are they blameworthy in the *attributability sense*? I have argued that we should answer *this* question by considering whether it would be fitting for them to feel ashamed. This is not a question about deserved suffering. Rather, it is a question concerning the correctness of the representational content of agential shame. The fact that shame is intrinsically painful is not relevant to *this* question. The control conditions of accountability and attributability differ; not because attributability blame is less painful than accountability blame, but rather because the painfulness of blame plays different roles depending on whether we ask whether blame is fitting or deserved.

Acknowledgements: For very helpful comments and discussion, I wish to thank Gunnar Björnsson, Randolph Clarke, Wessel van Dommelen, Anthony Duff, Austin Duggan, Anna Drożdżowicz, Jacob Elster, Jan Gertken, Robert Hartman, Sofia Jeppsson, Benjamin Kiesewetter, Kristin Michelson, Per Erik-Milam, Dana Nelkin, Douglas Portmore, Knut Olav Skarsaune, David Shoemaker, Caj Strandberg, Daniel Telech, Aness Webster, two Anonymous Readers, audiences at the University of Oslo, The Einstein Group Early Career Workshop in Political and Moral Philosophy, and the New Orleans Workshop on Agency and Responsibility.

References

- Bell, Malacester. 2013. *Hard feelings: The moral psychology of contempt*. Oxford University Press
- Carlsson, Andreas Brekke. 2017. "Blameworthiness as Deserved Guilt". *The Journal of Ethics* 21 (1): 89-115
- Clarke, Randolph. 2013. "Some Theses on Desert". *Philosophical Explorations*, 16 (2): 153-64.
- Clarke, Randolph. 2016. "Moral Responsibility, Guilt and Retributivism". *The Journal of Ethics* 20 (1): 121-137
- D'Arms, Justin and Daniel Jacobson. 2000. "The Moralistic Fallacy" In *Philosophy and Phenomenological Research*, 61, (1): 66 - 90.
- D'Arms, Justin and Jacobson, Daniel. 2014. "Wrong Kinds of Reason and the Opacity of Normative Force". In *Oxford Studies in Metaethics* vol. 9, Oxford University Press.
- Darwall, Stephen. 2006. *The Second Person Standpoint*. Oxford: Oxford University Press
- Darwall, Stephen. 2016. Taking Account of Character and Being an Accountable Person. In *Oxford Studies in Normative Ethics*, vol. 6, Oxford University Press, pp 12-36.
- Deonna, Julien, Raffaele Rodogno, Fabrice Teroni. 2012. *In Defense of Shame*. Oxford: Oxford University Press
- Eshleman, Andrew. 2014. Responsibility for Character. *Philosophical Topics* 32 (1/2): 65-94
- Feinberg, Joel. 1970. *Doing and Deserving*. Princeton N.J.: Princeton University Press
- Fischer, John and Neal A. Tognazzini. (2011) "The physiognomy of responsibility." *Philosophy and Phenomenological Research* 82 (2): 381-417.
- Greenspan, Patrica. 1995. *Practical Guilt*. Oxford: Oxford University Press

- Graham, Peter. 2014. "A Sketch of a Theory of Blameworthiness". *Philosophy and Phenomenological Research* 88 (2): 388-409.
- Hieronymi, Pamela. (2004). The force and fairness of blame. *Philosophical Perspectives*, 18(1): 115-148.
- Howard, Christopher M. 2018 "Fittingness". *Philosophical Compass*
- Kagan, Shelley. *The Geometry of Desert*. Oxford: Oxford University Press.
- Lewis, Michael. 1992. *Shame: The Exposed Self*. New York: The Free Press.
- King, Matt. 2012. "Moral Responsibility and Merit". *Journal of Ethics and Social Philosophy* 6 (2): 1-17.
- Mason, Elinor. 2019. *Ways to be Blameworthy*. Oxford: Oxford University Press
- Mason, Michelle. 2010. "On Shamelessness." *Philosophical Papers* 39 (3): 401-425
- McHugh, Conor, and Jonathan Way. 2016. "Fittingness first". *Ethics* 126 (3): 575-606.
- McKenna, Michael. 2012. *Conversation and Responsibility*. Oxford: Oxford University Press.
- McMahan, Jeff. 2009. *Killing in War*. Oxford University Press
- Nelkin, Dana. 2011. *Making Sense of Freedom and Responsibility*. Oxford: Oxford University Press
- Nelkin, Dana. 2013. "Desert, Fairness and Resentment". *Philosophical Explorations* 16 (2): 1-16.
- Nelkin, Dana. 2015. "Psychopaths, Incurable Racists and the Two Faces of Responsibility". *Ethics* 125 (2): 357-390.
- Nelkin, Dana. 2016. "Accountability and Desert". *Journal of Ethics* 20 (1-3): 173-189.
- Nelkin, Dana. Manuscript. *Blame and Desert*.

- Pereboom, Derk. 2013. "Free Will Skepticism, Blame, and Obligation". In D. J. Coates and N. A. Tognazzini, (eds.), *Blame: Its Nature and Norms*, New York: Oxford University Press, pp. 89–206
- Pereboom, Derk. 2014. *Free Will, Agency, and Meaning in Life*. Oxford: Oxford University Press.
- Pereboom, Derk. 2017. "Responsibility, Regret, and Protest". In *Oxford Studies in Agency and Responsibility*, vol. 4., pp 121-140
- Pickard, Hanna. (2013). "Irrational blame". *Analysis*, 73(4), 613-626.
- Portmore, Douglas. Forthcoming. *Opting for the Best*. Oxford University Press
- Prinz, Jesse. J. and Shaun Nichols. 2010. "Moral emotions". In *The Moral Psychology Handbook*. Oxford University Press, pp. 111-146.
- Rosen, Gideon. 2002. "Culpability and Ignorance". *Proceedings of the Aristotelian Society* 103: 61-84.
- Rosen, Gideon. 2015. "The Alethic Conception of Moral Responsibility". In Clarke, Randolph, Michael McKenna and Angela. Smith (eds.), *The Nature of Moral Responsibility. New Essays*, Oxford: Oxford University Press, pp. 65-89.
- Shoemaker, D. 2015. *Responsibility from the Margins*. Oxford: Oxford University Press.
- Smith, Angela. 2005. "Responsibility for Attitudes: Activity and Passivity in Mental Life". *Ethics* 115: 236-271.
- Tangney, J. P. and R. L. Dearing. 2003. *Shame and Guilt*. The Guilford Press
- Tangney, June Price, Jeff Stuewig, and Debra J. Mashek. 2007. "Moral emotions and moral behavior." *Annu. Rev. Psychol.* 58 (2007): 345-372.
- Taylor, Gabrielle. 1985. *Shame, Pride and Guilt*. Oxford University Press.
- Velleman, David. 2001. "The Genesis of Shame." *Philosophy and Public Affairs* 30 (1) 27-52.
- Wallace, R. Jay. 1994. *Responsibility and the Moral Sentiments*. Harvard University Press.

Watson, Gary. 2004. "Two Faces of Responsibility." In *Agency and Answerability* New York: Oxford University Press, pp. 260-289.

Watson, G. 2011. The Trouble with Psychopaths. In Wallace R., J., R. Kumar & S., R. Freeman (eds.) *Reasons and Recognition: Essays on the Philosophy of T. M. Scanlon*. Oxford University Press, pp. 307-331.