

A Conditional Defense of Shame and Shame Punishment

Erick Jose Ramirez

Abstract: In this paper I argue that, if we properly understand the nature of shame, it is sometimes justifiable to shame others in the context of a pluralistic multicultural society. I begin by assessing the accounts of shame provided by Cheshire Calhoun (2004) and Julien Deonna, Raffaele Rodogno and Fabrice Teroni (2012). I argue that both views have problems. I defend a theory of shame and embarrassment that connects both emotions to 'whole-self' properties. Shame and embarrassment, I claim, are products of the same underlying emotion. I distinguish between moralized and non-moralized shame in order to show when, and how, moral and non-moral shame may be justly deployed. Shame is appropriate, I argue, if and only if it targets malleable moral or non-moral normative imperfections of a person's 'whole-self.' Shame is unjustifiable when it targets durable aspects of a person's 'whole-self.' I conclude by distinguishing shame punishments from guilt punishments and show that my account can explain why it is wrong to shame individuals on account of their race, sex, gender, or body while permitting us to sometimes levy shame and shame punishment against others, even those otherwise immune to moral reasons.

Keywords: Calhoun, embarrassment, gender, mental illness, race, responsibility, shame.

Shame and embarrassment are almost always unpleasant to feel. They belong to a family of 'self-conscious' emotions that include pride and guilt and which are closely associated with self-esteem and self-worth. Additionally, because of the social nature of shame and embarrassment, it can be difficult for us to fully extinguish these feelings. Furthermore, because shame and, to a lesser extent, embarrassment appear to focus on who we are (as opposed to what we have done), some philosophers have argued that these emotions do not have a place in a liberal pluralistic society (Nussbaum 2004). In this paper I make appeal to empirical work on shame and embarrassment to argue that shame and shaming can be justly deployed in pluralistic multicultural societies. Beyond the informal use of shame, I offer a conditional defense of state-issued shaming punishment on the basis of my account of shame.

First, I argue that empirical work on shame and embarrassment should lead us to see both as different ways of expressing the same underlying emotional process. This view diverges from recent philosophical analyses of shame and embarrassment that draw a sharp boundary between these emotions. I argue that we have good reason for unifying these feelings and rejecting other accounts. The distinguishing features of shame and embarrassment will be

shown to depend upon how the person feeling the emotion conceptualizes the situation she finds herself in. Whether the emotion is called shame or embarrassment expresses something about the relationship between the agent and how she perceives the judgments that others make of what I will call her 'whole-self' properties.

I then put this account of shame and embarrassment into practice. I offer a conditional defense of shaming and shame punishment in the context of a pluralistic society. Although recent defenses of shame have focused on a distinctly moralized conception of the emotion (Calhoun 2004), I argue that shame can be appropriate not only against moralized character flaws but also against non-moral normative failings. Though I believe that shame has these uses, we must be careful, I suggest, to avoid shaming individuals for aspects of their selves that are either not easily malleable or about which reasonable evaluative disagreement is possible.

Shame: Recent Accounts

Shame has many critics. Some have argued that shame is always inappropriate because it is fundamentally degrading (Nussbaum 2004). Others worry that a society in which shaming is sanctioned will be subject to the tyranny of the moral majority or in which shame-based violence is used to violate the rights and dignity of citizens (Hall 2013). Even if we grant that the politics of shame requires reformation (and it does!), I will work from the assumption that shame, and indeed all our negative emotions, can have value (D'Arms and Jacobson 2000). Reformation, not abolition, is what is needed (Arneson 2007).

Julien Deonna, Raffaele Rodogno, and Fabrice Teroni (2012), hereafter DRT, have recently proposed a new account, and defense, of shame. On DRT's view, shame is defined in terms of the judgments that an individual makes about her values, her actions, and her character. On their view, shame is subjectively determined in a relatively extreme way by these judgments. Although social factors may play a *causal* role on their account of shame, the thoughts, judgments, or beliefs of others are neither necessary nor sufficient for the emotion. Additionally, although DRT attempt to distinguish shame from an emotion they call 'humiliation,' this account is impoverished. They do not explore the relationships between shame and anger (which they blend together in their account of humiliation) that would help to better distinguish embarrassment from humiliation. In these respects, their account of shame fails to capture important social and normative aspects of shame.

Second, Cheshire Calhoun (2004) has produced what may be the best known defense of shame in the recent literature. Unlike DRT's (2012) account, Calhoun rejects purely subjective explanations of the determinants of shame. Calhoun sees shame as an emotion that emerges when we are judged by others with whom we are co-participants in a normative framework. Shame, on her view, *requires* the existence of others (and their judgments) in order to function.

According to Calhoun, “[m]oral criticism has practical weight when we see it as issuing from those who are to be taken seriously because they are co-participants with us in some shared social practice of morality” (Calhoun 2004, 139). Because we are co-participants, and because co-participation is always aimed at shared activities (moral, academic, or other practice) co-participants must develop and exist within the context of shared social norms. Thus, on Calhoun’s account, even if we reject a co-participant’s judgment about us, we remain shamed on her account because these judgments retain pragmatic social significance for us: “[s]haming criticisms work by impressing upon the person that she has disappointed not just one individual’s expectations but what some ‘we’ expected from her” (Calhoun 2004, 140). One difficult question arising from such an analysis is who our co-participants are. Many of our shared normative institutions include hundreds of millions of others (in the case of national and global institutions). Notwithstanding this question, Calhoun’s view has several advantages over DRT’s account.

First, Calhoun’s view takes seriously the social elements of shame and shaming. DRT go to great lengths to deny that shame has this essentially social dimension. Calhoun is right, I argue, to reject this claim. The judgments of real or imagined others are essential for shame and shaming to take place. Without reference to the judgments of real or imagined others, DRT’s account of shame turns shame into a form of guilt over one’s character traits.¹

Calhoun’s account also has another, less appreciated, advantage over DRT’s more recent theory: it takes seriously the fact that shame can be experienced even when it goes against *one’s considered judgment or values*. Because Calhoun’s view of shame is grounded on the judgments of our co-participants, it is clear how those judgments could differ from our own. In other words, Calhoun’s account can explain why an immigrant may feel ashamed about his legal status even if he consciously judges that there is nothing to be ashamed of. It can explain why a person may feel ashamed about their sexual identity and preferences even if they think that they should reject the values that govern this shame. This kind of shame is neither an aberration nor is it rare. That DRT do not allow for its existence is a significant drawback of their view. Worse, because DRT make shame an emotion tied to an individual’s consciously held judgments and values they would need to claim that these individuals consciously believe

¹ Guilt is traditionally defined as an emotion that is subjectively determined (i.e., only the guilty party can trigger guilt feelings). These feelings arise from judging that something we have done was morally wrong. In this sense, guilt is focused on the moral wrongness of our actions and only our judgment can trigger the feeling. That others believe I acted wrongly has no bearing on whether I feel guilty so long as I think what I have done is permissible. Shame, on the other hand, focuses not on the wrongness of our actions but instead on more global judgments about our whole selves. We feel guilt for what we have done but we are ashamed of ourselves.

that they ought to feel ashamed in order to explain this shame. This is an unpalatable account of these forms of shame.

Although Calhoun's view has these advantages, I argue that it too should be rejected. I do this for two reasons. First, Calhoun's account of shame, like DRT's, makes it difficult to differentiate shame from embarrassment. Although both Calhoun and DRT distinguish shame from guilt, Calhoun does not clearly distinguish shame from embarrassment. This is a mistake for reasons that will become clear later. Although DRT distinguish shame from humiliation, their account of humiliation fails to capture the essence of embarrassment (and, indeed, cannot explain why we are sometimes embarrassed, but not humiliated, by positive attention).²

Additionally, both Calhoun and DRT neglect relevant empirical data on implicit bias, dual systems heuristics, and recalcitrant emotion that can help us make better sense of shame. Calhoun and DRT offer opposed, but polarized, theories of the role of subjectivity when it comes to shame. DRT assert that shame is entirely dependent on an individual agent's subjective values and judgments. It is impossible for others to shame me, on their account, unless I judge that I have failed to live up to values that I myself endorse (more accurately, when I judge that I have not only failed to live up to my values but have instead acted in a way that I judge to be antithetical to *my* values). This view places too much responsibility on an agent for her own shame (and provides an unpalatable account of shame as it relates to gender, race, and sexual identity).

Calhoun, on the other hand, moves too far in the opposite direction. On her view we are subjects of shame whenever anyone with whom we are co-participants in a normative enterprise judges us negatively. Given the large scale of contemporary normative enterprises, Calhoun's view allows for situations that appear to elicit not shame but anger instead. It is at this point that better theories of shame, embarrassment, and anger can help us to make sense of this landscape and better explain when shame and shaming could be justifiable.

Shame, Embarrassment, Humiliation, and Anger

Although many theories of shame and embarrassment can be found in the psychological and philosophical literature (Keltner and Buswell 1997; Maibom 2010; Tagney and Miller 1996; Taylor 1985; Williams 1986/2006), I focus on the particularly promising account of shame and embarrassment that emerges from research conducted by John Sabini and his colleagues. On Sabini's view (Sabini

² Interestingly, DRT approvingly cite Sabini's work in the context of discussing embarrassment (2012, 117). This is odd given Sabini's rejection of a bright line between shame and embarrassment, something DRT go to great lengths to argue *for*. It's not clear why DRT think it possible to accept Sabini's work on embarrassment without undermining their major thesis about the nature of shame.

and Silver 1997; Sabini, Garvey, and Hall 2001) shame and embarrassment are two faces of the same underlying affective process. They are not distinct emotions but instead represent different ways of framing the same emotion. These different frames carry contrasting expressive messages about the feeler's appraisals (conscious or unconscious) about the situation that elicited the feeling.

Why should we think that shame and embarrassment represent different faces of the same emotion? On the one hand, we lack good evidence for uniquely distinguishing them. Shame and embarrassment cannot be distinguished via reference to unique patterns of bodily changes nor can they be distinguished by looking at patterns of neurological activation (Sabini et al. 2000; Sabini, Garvey, and Hall 2001). Additionally, Sabini and his colleagues argue that shame and embarrassment cannot be distinguished from one another in terms of their severity (shame is not necessarily a more severe or intense version of embarrassment), nor can the two feelings be distinguished by their characteristic objects (both shame and embarrassment target what are called 'whole-self' properties). Lastly, the failure to find distinct neurological or physiological markers for shame and embarrassment is not a unique phenomenon. Most emotions lack these features (Barrett 2006; Gendron, Roberson, and Barrett 2015).

Lacking mind-independent unifying characteristics, Sabini and colleagues argue that shame and embarrassment are products of the same underlying affective process. This process is triggered whenever real or imagined others are in a position to judge one or more of our 'whole-self' properties. Whole-self properties are features of a person's identity that may manifest in action but which are present even when not immediately acted upon: character traits, racial traits, gender and sexual identities, and the like represent some of our whole-self properties. According to Sabini:

When someone makes the appraisal that something has happened that might be taken as evidence that his or herself has been discredited, an emotional state is triggered. That painful, inhibiting state, one that leads one to want to become small and hide, might be called "State A." If the person is later asked (or for some other reason chooses) to describe that state, if the person wishes to imply that he or she sees the revealed flaw as real, then he or she will call it shame. However, if he or she wishes not to license the inference that he or she believes a real flaw of the self was revealed, then he or she will call it embarrassment. (Sabini et al. 2000)

Any scenario where one or more of our whole-self properties is judged by real or imagined others is therefore an occasion for what Sabini calls 'State A.' Conceptualization, on this view, plays a key role in the differentiation of shame from embarrassment. We call the state shame as an indication that the negative evaluation of the whole-self property is true. We call the state embarrassment as a way of acknowledging that, although the situation *appears* to license the judgments that others are making of us, we don't really think we have the whole-self property in question.

For example, accidentally bumping into someone seems like a cause for embarrassment. This is because, in calling the emotion embarrassment, I acknowledge that it appears that I am clumsy or uncaring about personal spaces (both are relevant whole-self properties that may be up for judgment). However, I also, in saying that I am embarrassed, indicate that I think I actually *do* care about the personal spaces of others and that the fact that I bumped into you shouldn't lead you to think that I lack respect for your space. This way of thinking about embarrassment also helps us make sense of the seemingly puzzling instances of feeling embarrassed by *positive* judgments of our whole-self properties.

Feeling embarrassed by a compliment, in this instance, expresses the same message as embarrassment in other contexts: "although I grant that it may seem like I'm x, I don't really think I am x." DRT's conception of humiliation does not capture this phenomenon. On their view, humiliation necessarily requires a form of unjust public attack on our person. Embarrassment does not require such an injustice. Sabin's conception of shame and embarrassment, however, does make room for a form of anger that derives from 'State A' experiences that captures DRT's sense of humiliation as well.

On the view I am proposing, shame is a way of acknowledging an acceptance (conscious or not) of the real or imagined judgments of others.³ Embarrassment is a way of acknowledging that, while a situation appears to justify an inference about our whole-self, we reject the judgment the situation appears to license. This leaves a space open for situations in which real or imagined others judge a whole-self property of ours and where we believe that those judgments are not at all licensed by the situation. For example, suppose a colleague calls me lazy despite significant evidence to the contrary. In such cases, Sabin claims, the emotion I am likely to feel is neither shame nor embarrassment (nor humiliation) but rather, I will feel anger. This account functions better both descriptively and normatively. Anger, not humiliation, is the right response to baseless judgments about our character.

As I have argued, Calhoun's account, while possessing advantages over DRT's view, does not match up to the picture of shame that emerges from the empirical literature. Calhoun is motivated to defend her account of shame, in part, out of a wider rejection of traditional conceptions of autonomy that fail to appreciate how autonomy can be socially constructed. This appreciation of the social dimension thus leads her to view shame not only as socially constructed in a shallow sense (i.e., that shame requires a social world and that this social world defines the sorts of exchanges that make shame possible for us) but in a much deeper sense. Shame is constructed, on her view, without input on the part of the

³ Unconscious appraisals (discussed below), allow appraisal theories like Sabin's to make room for the effects of internalized norms not present in consciousness.

agent who is an object of shame (her judgments are neither necessary nor sufficient for shaming).

On Calhoun's view, the judgment of co-participants ought to trigger shame even when we believe that they lack any grounding. This seems wrong on both empirical and normative fronts. First, it does not appear that subjects in situations like these actually do tend to feel ashamed (Sabini, Garvey, and Hall 2001). Second, shame would be the wrong response to a groundless judgment about our whole-self properties. It is understandable why some might think that I am uncaring if I arrive late (through no fault of my own) to the department meeting, and hence why such a situation may be embarrassing. However, given no evidence at all for such a judgment (I have always been to meetings on time, go above and beyond with my teaching and service, etc.), being called uncaring in such circumstances, even by those with whom I am co-participants in a normative framework, calls for anger, not shame.

Acceptance of this view of embarrassment and shame need not imply, as Calhoun seems to claim, that we lose track of the essentially social aspects of these emotions. Both shame and embarrassment make sense only in the context of social spaces. For example, suppose that you are in the grocery store and I accidentally knock over a display of paper rolls which now begin to cascade onto the floor. This was an accident, of course, though the (real or imagined) gaze of others demands a response from me.

In a series of now classic studies, Semin and Manstead (1982) showed subjects a series of video clips of exactly the sort of scenario described. They argued that in a case like this "embarrassment displays can be construed as non-verbal apologies for social transgressions" (Semin and Manstead 1982, 369). Semin and Manstead furthermore found that individuals made less severe judgments about someone who knocked over a store display who then appeared embarrassed relative to those who were not embarrassed (1982). Sabini's research on shame and embarrassment can help us make sense of how all of this can be the case. Embarrassment signals a recognition, on the part of the embarrassed agent, that the situation appears to license a judgment about one or more of her whole-self properties. Furthermore, it signals that the embarrassed agent *rejects* this judgment about her properties. In doing this, she joins her audience in condemning the whole-self property that is *apparently* on display while, simultaneously, signalling that she does not actually possess the property in question. The social dimension is therefore an essential feature of any explanation of 'State-A' emotions.

Sabini's view provides a better account of the relationship between shame, embarrassment, and anger than prior accounts. However, it also has the resources to better explain the phenomena of 'recalcitrant' emotions. An emotion is recalcitrant when we continue to feel it despite our judgment that the feeling is unwarranted (D'Arms and Jacobson 2003). Survivor's guilt, for example, is often felt by individuals even though they (rightly) believe that they have done

nothing wrong. All emotions have the ability to manifest as recalcitrant. Because DRT's view of shame requires that agents consciously judge that their actions flout their intentionally held values in order for shame to manifest, it carries the consequence that shame cannot be recalcitrant. Given the widely experienced nature of emotional recalcitrance, this is a significant cost of DRT's view.

Although Calhoun's view allows for a kind of recalcitrance, I argue that Sabini's view has the resources to provide a better explanation. Because Calhoun's view of shame allows for the judgements of others, all on their own, to shame an individual, it is possible for an individual to feel ashamed despite her own opposition. As previously noted, many of these situations would seem to elicit anger and not shame. Additionally, by ignoring the shamed agent, Calhoun provides an impoverished account of how recalcitrance can arise in us.

Sabini's theory of shame and embarrassment is an appraisal theory. In other words, subjects must conceptualize (i.e., appraise) their situation in the right way in order to feel certain emotions. These conceptualizations, importantly, need *not* be conscious. Many appraisals happen at the subdoxastic level. Dual process theories of judgment and decision making (Stanovich and West 2000; Kahneman 2003) for example, help to explain how subjects can feel ashamed by aspects of their whole-selves as a result of internalized implicit biases that run counter to their explicitly chosen conscious values.

Tension between implicit, internalized, norms and explicit, consciously held values better explains the phenomena of recalcitrant shame. Internalizing anti-immigrant norms, for example, may explain why someone may feel ashamed of being an undocumented immigrant even though they believe that open immigration policies are just. This explanation maintains contact with the individual feeling shame, unlike Calhoun's view, while keeping the judgments of real or imagined others an essential feature of the emotion. I now turn my attention to the regulation of what I will call 'moralized' and 'non-moralized' shame in the context of a pluralistic multicultural society.

The Regulation of Moral and Non-Moral Shame

I have argued that we should understand shame as an emotion that expresses something about the explicitly held or implicitly internalized norms of a speaker when one or more of her whole-self properties are judged by others (real or imagined). I now wish to distinguish between two different senses of shame that I will refer to as moralized and non-moralized shame.

The objects of shame (and embarrassment) are whole-person properties. These properties can manifest themselves in our actions but are not, themselves, easily reducible to an action. Although character traits are paradigmatic whole-person properties (i.e, one can be honest, manipulative, brave, selfish, cruel and so on), whole-person properties include features of a whole-self that are not, arguably, character/virtue based. For example one may be a good swimmer, a bad comedian, a mediocre artist, or an attractive person. Additionally, features of

a whole-self include properties relating to racial, sexual, and gender identity along with a person's genetic, cultural, and ethnic heritage. This allows us to provide a unifying explanation of shame that connects shame felt about whole-self properties that vary across many dimensions (character properties, identity properties, embodied properties, and historical properties).

There are important distinctions to note about shame that relate to the kind of whole-self property involved. When a person is shamed over a whole-self property that represents, in the eyes of the shamer, a moral failing, we can refer to this form of shame as moralized shame. Moralized shame targets a feature of a person's whole-self in an attempt to use a subject's shame to enact their own conceptions of moral progress. Moralized shame, in this sense, is defined largely in terms of the intentions of the shaming party. This is in contrast to traditional accounts of guilt which are largely focused on an individual judging that he has acted wrongfully. As a result, it is possible for a committed racist to use race-shame in a moralized way. Although this form of moralized shaming is wrong, I will argue that at least some forms of moralized shame can be justifiably deployed in a pluralistic society. My sense of moral, in this application, will therefore be limited to moral norms that must be shared in order for pluralistic communities to flourish.

Non-moralized shame can be defined as shame which targets a whole-self property for a non-moral failure. Non-moral shame is still normative, and this normativity will play an important role in my defense of non-moral shame, but this normativity is not essentially moral. For example, one may be ashamed of being a bad teacher in a way that does not implicate a clear moral failing.

We can also mark distinctions grounded on features of the whole-self property being targeted by the shamer. As previously noted, some whole-self properties are more straightforwardly related to an agent's actions than others. A person can only manifest the whole-self property of 'being a liar' or 'being a narcissist' or 'being courageous' if they have a history of acting in ways that support the existence of such properties.⁴

Other whole-self properties are less clearly connected to an agent's actions. For example, being a member of a certain racial group is explicable largely in terms of the social and historical context that an individual is located in and are less dependent on an individual's actions, beliefs, and judgments than other whole-self properties that may adhere to us.⁵ Whole-self properties connected to

⁴ In marking this distinction, I step into somewhat contentious territory about the nature of character traits (Sreenivasan 2013; Doris 2002). Because I focus here on the ethics of shaming, it is essential, to my account, that shame is only permissible, in the rare cases where it is permissible, only on account of whole-self properties that manifest in behavior.

⁵ The construction of race is, of course, much more complicated than I have laid out here. Additionally, insofar as race has performative aspects, individual behavior can be a relevant (but not determining) feature of one's racial identity. For example, Mallon (2004) and Appiah

action are therefore more malleable than those whole-self properties not so easily connected. This distinction is important to keep in mind when discussing the permissibility of shaming, especially when enacting shame punishments.

While it may be difficult to shift from being a liar to someone who is honest, it is even more difficult to shift more durable whole-self properties (those relating to race and sex are especially durable in this sense). Other properties lie somewhere in between these extremes. Bodily whole-self properties, for example, are often subject to shame. Though our body shape is loosely under our control, the malleability of bodies (especially with regard to weight and appearance) is less susceptible to individual action than other more straightforwardly character-based whole-self properties. I argue below that this fact makes body shaming unjustifiable.

With these distinctions in mind, I now propose (and defend) both moralized and non-moralized shame and, in rare circumstances, even the enactment of shame punishments. Moralized and non-moralized shame, in the context of a pluralistic society, can be used to effect positive change. Shame is permissible, on my view, when it meets the following criteria:

1. Shame does not target a deeply durable whole-self property: durable properties are those non-character based whole-self properties that are difficult, if not impossible, for an individual to alter.
2. Shame does not target features of the whole-self over which reasonable pluralistic disagreement exists: a disagreement is reasonable if it is possible for parties to the disagreement to live alongside one another while maintaining their disagreement.

The first criterion helps us to see which whole-self properties are proper targets for shaming. The second criterion situates shaming in the context of pluralism and multiculturalism. One of the many concerns that opponents of shame punishment raise is the danger of a tyranny of a moral majority (Nussbaum 2008; Hall 2013). Shame can be dangerous if used to force a universalist normative ideal on a diverse citizenry. We can help to mitigate the problem of a moral mob by appealing to Rawls' notion of a public reason in the context of multicultural and pluralistic nations.⁶ My defense of shame and

(1996) go to great lengths to demonstrate the various ways in which institutional and social forces can shape racial categories and racial identity.

⁶ My account of reasonableness is indebted to Rawls' (1996) account of the reasonable citizen. In doing so, I do not intend to commit myself fully to Rawlsian political ideology. For example, though I appeal to Rawls' notion of a public reason to define a reasonable disagreement, I use it only to discuss the permissibility of shame and, in an even more limited sense, shame punishment. I take seriously concerns others have raised that a state that adopts Rawls' principle of public reasons may operate on anti-democratic norms (Benhabib, 2002). I do not aim to restrict the content of speech in a pluralistic society. My argument is intended to justify the shaming of individuals who hold views that are unreasonable in the thin sense of reasonable implied by the notion of a public reason.

shaming is conditional: only whole-self properties over which reasonable disagreement is impossible are proper targets for shame.

Examples will be helpful to clarify the criteria. Reasonableness is admittedly difficult to define. For my purposes, a commitment to pluralism requires the acknowledgement that others may disagree with me about matters of value. So long as it is possible for individuals to peaceably live alongside one another while maintaining their disagreements, the disagreement is a reasonable one. Persons of different religious groups, for example, hold deeply inconsistent metaphysical views and/or have conflicting beliefs about value. However, it is possible for most parties of religious disagreements to live alongside one another provided neither's beliefs command the conversion or elimination of non-believers. This type of disagreement, while deep and longstanding, is reasonable. Religious shaming, in these situations, would be inappropriate as the change it would effect is inconsistent with value pluralism.

Contrast this with the conflict that holds between a committed racist of the sort who believes that members of other races (as he construes them) are sub-human. In such a case, it is impossible for members of different races to live alongside such a person without radically constraining the liberty of at least one of the parties. The disagreement between the racist and his neighbors is unreasonable. Racism, in this sense, is defined by a set of beliefs. Even if the process of changing these beliefs is difficult, beliefs are malleable. The racist can therefore be rightfully shamed in such a scenario and *ought* to be ashamed of himself for his beliefs. The shame, in this circumstance, would ideally function to change his whole-self property of 'being a racist' to something that fits more comfortably in a pluralistic society.⁷ The same would apply to non-racist political ideologies that would qualify as unreasonable in the thin sense I intend here (i.e., those that restrict voting rights along gender or class lines). These would also be justifiably open to shame and shaming punishments.

Consider another example: aesthetic whole-self properties. Properties like 'being beautiful', 'being elegant' and the like may be *durable* and thus blocked by my first criterion from shaming. However, even if they are not durable, that is, even if individuals can exercise some control over their aesthetic whole-self properties, I argue that they still form an inappropriate basis for shaming. Aesthetic properties will *always* be subject to reasonable disagreement even where culturally dominant models of beauty exist. This is because it is always

⁷ Success, in this instance, does not mean that we fully convert the racist or remove all of his noxious beliefs. All that is needed is to shame the racist into a more reasonable position about the humanity of his neighbors so that he may peaceably live alongside them. Though moralized shamers may seek to enact even greater change in the racist, this form of shaming would run afoul of my second criterion. The moral majority does not have a right to shame others into accepting its values so long as evaluative disagreements are reasonable. It's not clear whether any forms of racism would be reasonable in the public reasons sense though at least some offensive ways of being would remain.

possible for individuals with conflicting aesthetic ideologies to live alongside one another. The second criterion, grounded on the notion of reasonable disagreement, blocks the imposition of aesthetic mob rule just as much as it blocks the rule of the moral majority. The only exceptions to these generalizations are aesthetic ideologies that combine with one or more moral failings to threaten the existence of shared public spaces.⁸

Shame Punishments

The function of shame is to enact a positive change in whole-self properties that, for moral or non-moral reasons, are inconsistent with pluralistic multicultural societies. Shame is poised to enact these changes in ways that other reactive attitudes, like guilt, are not. Specifically, this is because guilt is typically thought to focus on an agent's *actions*. That is, an agent must judge that they have *done something wrong* in order to feel guilty. Reparative behaviors related to guilt therefore focus on repairing the negative consequences of a bad action. Guilt does not, importantly, require that agents consider their whole-self properties from which those bad actions issue (Williams 1989/2006). Reparative action related to shame, therefore, is focused on *changing* the whole-self properties that an agent is ashamed of. It is not enough to apologize for lying in order to rid oneself of the shame of being a liar. One must *cease to be* a liar. This is the special value of shame. When properly deployed, shaming can enact positive change in both the shamed person and, as a consequence, provides a benefit to his society as well (Arneson, 2007).

Shame punishments represent punishments whose main aim is to encourage a specific form of self-reflection in the hopes of enacting positive change in the target's whole-self; importantly, shame punishments extend beyond traditional guilt-based incarceration and often include dialogue between the subject of shame punishment and the individuals and/or community enacting the punishment (Book 1999). Successful shame punishments change a person's perspective so that he comes to reject the targeted whole-self properties and is driven to change them in order to relieve himself of shame. Shame punishments can vary though they all share the goal of publicizing that the shamed person has the whole-self property for which he is being shamed. For example, some shame punishments require that an individual stand in front of a courthouse holding a sign identifying their crime or to place a similar sign near their home (Book, 1999).

Before defending shame punishment, I wish to say several things about shame and its connection to shame punishment. First, if my earlier argument has

⁸ In cases like these, shame is appropriate only because the relevant whole-self property poses a threat to public spaces and not, purely because it is the subject of aesthetic disagreement. Examples are difficult to come by though aesthetized versions of Feinberg's (1985) bus cases may help elucidate the relevant sense of 'threat to public spaces' that I have in mind.

force, then shame and shaming are sometimes justifiable in the context of a multicultural and pluralistic society. To say that shame is justified in this context is to say that a person *may* be shamed in a narrow range of cases (for malleable whole-self properties that are not properties over which reasonable disagreement is possible). However, it does not tell us who is justified in shaming the individual. I argue that the *state* is best poised to shame individuals for such properties. Why is this the case? If shame is left at the purely informal level, concerns about mob rule are far more likely to arise. In a minimally just pluralistic society, the rule of law would prevent states from enshrining into law legislation that would run afoul of my proposed criteria for shaming. As a result, state-sanctioned shaming (i.e., shame punishments) is justified if shaming is justified.

We must be cautious in the application of such punishment. Historical shame punishments almost all run afoul of my proposed criteria. Many historical shame punishments target whole-self properties that are subject to reasonable disagreement. Adultery, apostasy, and bigamy, for example, are norms over which reasonable disagreement is possible. Shame punishments, historical or contemporary, that target these properties are unjustifiable. Secondly, historical shame punishments often faltered because they made it impossible to enact positive change in the shamed individual by transforming a malleable whole-self property into a durable one. For example, although 'being a thief' is a whole-self property that it may be appropriate to shame someone for, branding an individual makes it impossible to change the property and hence is an inappropriate form of shaming.⁹ Lastly, we must be cautious in the forms of shaming that we allow. In what follows, I specify the particular constraints that should limit the form of shame punishments.

The infliction of shame, in the limited circumstances I outline here, is meant to cause positive change in the individual being shamed. Shame, as I defend it, targets malleable whole-self properties that render it impossible for an individual to live peaceably alongside his neighbors. Because the removal of shame requires the removal of whole-self property targeted, shaming can cause positive change to the individual. Questions about the efficacy of shaming in terms of enacting positive changes in a person are partly empirical and in that domain the answer appears to be that shaming *can* effect positive change in a person if used correctly. For example, shaming appears to reduce recidivism more than traditional guilt-based incarceration (Tangney, Stuewig, and Martinez 2014). Shame, moral or non-moral, must be aimed very specifically only to those whole-self properties that we wish to change.

⁹ In the United States, many states require that any individual convicted of a sex offense list themselves in a 'sex offender registry' for their lifetime and to update this registry whenever they change their home address. This form of shaming may also transform a malleable property ('being a sex offender') into a durable one and hence may also be an inappropriate form of shaming.

Shame must not be used to merely reject a whole-agent as a member of society. Call such shame disintegrative shame. Such disintegrative shaming, insofar as it targets an entire person, unfairly targets durable whole-person properties along with those that are subject to reasonable disagreement. Disintegrative shaming is ineffective at producing positive character change. For example, some have argued that “when individuals are shamed so remorselessly and unforgivingly that they become outcasts...it becomes more rewarding [for them] to associate with others who are perceived in some limited or total way as also at odds with mainstream standards” (Braithwaite 1989, 67). For this reason, I favor appealing to what is known as *reintegrative* shaming in the context of shame punishment (Braithwaite 2002; Morris 2002).

Reintegrative shaming aims to respect the individual while attempting to get him to see one of his particular whole-self properties as negative. Insofar as shame has a purpose, its purpose should be reintegrative. However, I have argued that shaming and shame punishments can be used in several ways. The following examples should clarify the extent to which shame, properly attenuated, has a place in a pluralistic society.

Take the following three cases: Alzheimer’s disease, antisocial personality disorder, and gender dysphoria. Mental illness is a notoriously difficult and multifaceted construction with many critics (Maddux 2001; Wakefield 1992). The 400+ list of mental illnesses and disorders classified in the *Diagnostic and Statistical Manual of Mental Disorders* (APA 2013) notoriously bear little relationship with one another in a way that allows for an easy unification of the concept of mental illness (Wakefield 1992). Given the variable nature of particular mental illnesses, individuals with some mental illnesses may be open to shame while others, due to the nature of their illnesses, would not be. This is because the DSM identifies conditions that include whole-self properties over which individuals can exercise significant control along with conditions over which individuals have little to no control.

Alzheimer’s disease is a complex disorder arising largely from factors beyond an agent’s ability to modify and over which agents are able to exercise little present control. Although “being a person with Alzheimer’s disease” is a whole-self property, and hence could be something that a person is ashamed of, I argue that it is not a proper object of personal or social shaming. This is true even though it is understandable why some may feel shamed by the effects and progression of the disease. ‘Being a person with Alzheimer’s disease’ is too durable a property for social shaming. We may additionally argue that taking neurodiversity seriously should lead us to view the value of ‘being a person with Alzheimer’s disease’ as the subject of reasonable disagreement.¹⁰

¹⁰ Alexandra Perry and C.D. Herrera (2013) speak broadly on the issue of the commitment to pluralism embedded within the context of neurodiversity movements.

On the other hand, consider an individual diagnosed with antisocial personality disorder (ASPD). A person with this condition would be an apt candidate for shaming based on my proposed criteria. ASPD is diagnosed largely via reference to a person's harmful antisocial behavior and their emotional character (APA 2013, 476). Both behavior and affectivity are things over which agents are able to exercise at least some control, they are malleable.¹¹

Non-moral shame and shame punishment are especially significant for individuals diagnosed with ASPD (and especially the subset of those with ASPD who are also psychopaths under the PCL-R test). This is because some have argued that those with ASPD may not understand moral reasons and thus are constitutively unable to feel guilty about what they do. Insofar as understanding moral wrongness is thought to license retributive punishment, this line of reasoning would suggest that those with ASPD cannot be held responsible for what they do (Levy 2007). Importantly, though individuals with ASPD may not be susceptible to guilt, evidence suggests that they are capable of feeling shame (Morrison & Gilbert 2001).

Insofar as those with ASPD are amenable to non-moral forms of shame, concerns about holding these individuals responsible can be largely circumvented (Ramirez 2013; Ramirez *forthcoming*; Talbert 2008). This is a very real advantage of including non-moral shaming in a theory of shame. So long as a person can be ashamed of a noxious whole-self property, their inability to feel guilty about it (or to understand that what they are doing is *morally* wrong) does not excuse them from accountability. Furthermore, the motivating effects of shame and shame punishment can be used to enact positive behavioral improvement in such individuals in ways that have proven more successful than traditional therapeutic approaches (Tangney, Stuewig, and Martinez 2014). Some have argued that ASPD, along with all 'cluster-B personality disorders,' are better construed as moral disorders given that they are characterized primarily in terms of moral failings (Charland 2007). Because ASPD is malleable and because ASPD is largely defined in terms of antisocial behavioral traits that make it difficult, if not impossible, for someone with ASPD to live peaceably with others, it is permissible, assuming my criteria are met, to target these individuals for shame and shame punishment.

Lastly, I wish to consider the diagnosis currently described as gender dysphoria (GD) in the DSM (APA 2013, 451). Gender dysphoria is described as a persistent and pervasive dissatisfaction or unhappiness (dysphoria) caused by a "marked incongruence between one's experienced/expressed gender and primary and/or secondary sex characteristics" (APA 2013, 452). Gender dysphoria replaced a disorder known as 'Gender Identity Disorder' (GID) in the previous incarnations of the DSM (APA 2000, 581). Gender identity (either under

¹¹ Some may have impulse control disorders that transform affectivity into an especially durable trait. Such individuals may be exempt from shaming on that basis.

Erick Jose Ramirez

the GD or GID labels) should not, on my account of shame, be something about which individuals should be made to feel shamed about. Gender identity is a durable property. Additionally, social constructions of sex and gender are topics about which reasonable disagreement is possible. Much as with racialized shame, we can make sense of how individuals who internalize norms that denigrate their identity can come to feel a pervasive sense of shame about this whole-self property. The proper response, in the context of value pluralism, is to prevent the shaming of people over their gender identities and to encourage a shift in perspective so that the shamed individual comes to reject the norms that ground it.

Conclusion

Shame is a powerful reactive attitude that is often associated with the worst aspects of the social world. Shame is unpleasant to feel and, when recalcitrant, can make us doubt our own value and even our identity. It is important, for these reasons, that we better understand the conditions that trigger shame and how this emotion is related to humiliation and anger.

I argue that it is best to think of shame and embarrassment as two faces of the same underlying emotion. Once better understood, I argue that shame, used in the right way, can be an important instrument for positive change in the context of pluralistic liberal societies. Provided that a society targets malleable and unreasonable whole-self properties, shame can be used to reintegrate individuals into society in ways that have been too-long neglected.

Shame is a dangerous tool, however, when used incorrectly and caution must always be exercised to ensure that shame is used in reintegrative not disintegrative ways. Though the dangers of shame and shame punishment have led some to argue in favor of social prohibitions against shame, this conclusion is too strong. Shame, suitably, rehabilitated, can be an important aspect of social progress and social justice.

References

- American Psychiatric Association. 2013. *Diagnostic and Statistical Manual of Mental Disorders*. 5th ed. Washington, DC: American Psychiatric Association.
- . 2000. *Diagnostic and Statistical Manual of Mental Disorders: DSM-IV-TR*. Washington, DC: American Psychiatric Association.
- Appiah, Kwame Anthony. 1996. "Race, Culture, Identity: Misunderstood Connections." In *Color Conscious: The Political Morality of Race*, K. Anthony Appiah and Amy Gutmann, 30-105. Princeton, NJ: Princeton University Press.
- Arneson, Richard J. 2007. "Shame, Stigma, and Disgust in the Decent Society." *The Journal of Ethics* 11(1): 31-63.

- Barrett, Lisa Feldman. 2006. "Are Emotions Natural Kinds?" *Perspectives on Psychological Science* 1: 28-58.
- Benhabib, Seyla. 2002. *The Claims of Culture: Equality and Diversity in the Global Era*. Princeton: Princeton University Press.
- Braithwaite, John. 1989. *Crime, Shame and Reintegration*. Cambridge: Cambridge University Press.
- . 2002. "Setting Standards for Restorative Justice." *British Journal of Criminology* 42: 563-577.
- Book, Aaron S. 1999. "Shame on You: An Analysis of Modern Shame Punishment as an Alternative to Incarceration." *William and Mary Law Review* 40: 653-686.
- Calhoun, Cheshire. 2004. "An Apology for Moral Shame." *The Journal of Political Philosophy* 12(2): 127-146.
- Charland, Louis. 2007. "Character: Moral Treatment and the Personality Disorders." In *The Philosophy of Psychiatry: A Companion*, edited by Jennifer Radden, 64-77. Oxford UP: New York.
- D'Arms, Justin and Jacobson, Daniel. 2000. "The Moralistic Fallacy." *Philosophy and Phenomenological Research* 61(1): 65-90.
- . 2003. "The Significance of Recalcitrant Emotion." In *Philosophy and the Emotions*, edited by Anthony Hatzimoysis, 127-146. Cambridge: Cambridge University Press.
- Deonna, Julien, Rodogno Rafaele and Teroni, Fabrice. 2012. *In Defense of Shame: The Faces of an Emotion*. New York: Oxford University Press.
- Doris, John. 2002. *Lack of Character: Personality and Moral Behavior*. Cambridge: Cambridge University Press.
- Feinberg, Joel. 1985. *Offense to Others*. New York: Oxford University Press.
- Gendron, Maria, Debie Roberson, and Lisa Feldman Barrett. 2015. "Cultural Variation in Emotion Perception is Real: A Response to Sauter et al." *Psychological Science* 26: 357-359.
- Hall, Casey. 2013. *Shame, Humiliation, and Punishment in the Liberal Society*. (Unpublished doctoral dissertation). University of California Irvine, CA, USA.
- Kahneman, Daniel. 2003. "Maps of Bounded Rationality: Psychology for Behavioral Economics." *The American economic review* 93(5): 1449-1475.
- Keltner, Dacher, and Brenda N. Buswell. 1997. "Embarrassment: Its Distinct Form and Appeasement Functions." *Psychological Bulletin* 122: 250-270.
- Levy, Neil. 2007. "The Responsibility of the Psychopath Revisited." *Philosophy, Psychiatry, & Psychology* 14(2): 129-138.
- Maddux James. 2001. "Stopping the Madness." In *The Handbook of Positive Psychology*, edited by Charles R. Snyder & Shane J. Lopez, 13-25. Oxford: Oxford University Press.
- Maibom, Heidi. 2010. "The Descent of Shame." *Philosophy and Phenomenological Research* 80(3): 566-594.

Erick Jose Ramirez

- Mallon, Ron. 2004. "Passing, Traveling and Reality: Social Constructionism and the Metaphysics of Race." *Noûs* 38(4): 644–673.
- Morrison, David and Paul Gilbert. 2001. "Social Rank, Shame and Anger in Primary and Secondary Psychopaths." *The Journal of Forensic Psychiatry* 12(2): 330-356.
- Morris, Allison. 2002. "Critiquing the Critics: A Brief Response to Critics of Restorative Justice." *British Journal of Criminology* (42)3: 596-615.
- Nussbaum, Martha. 2004. *Hiding from Humanity: Disgust, Shame, and the Law*. Princeton: Princeton University Press.
- . 2008. "Hiding From Humanity: Replies to Charlton, Haldane, Archard, and Brooks." *Journal of Applied Philosophy* 25(4): 335-349.
- Perry, Alexandra and Chris D. Herrera. 2013. *Ethics and Neurodiversity*. Newcastle upon Tyne: Cambridge Scholars Publishing.
- Semin, Gun R. and Anthony S.R. Manstead. 1982. "The Social Implications of Embarrassment Displays and Restitution Behavior." *European Journal of Social Psychology* 12: 367-377.
- Ramirez, Erick. 2013. "Psychopathy, Moral Reasons, and Responsibility." In *Ethics and Neurodiversity*, edited by Alexandra Perry and Chris D. Herrera, 217-237. Newcastle upon Tyne: Cambridge Scholars.
- . *forthcoming*. "Psychopathy, Autism, and Basic Moral Emotions: Evidence for Sentimentalist Constructivism." In *Bloomsbury Companion to Psychiatry*, edited by Robyn Bluhm and Şerif Tekin. New York; London: Bloomsbury.
- Rawls, John. 1996. *Political Liberalism*. New York: Columbia University Press.
- Sabini, John, Brian Garvey, and Amanda Hall. 2001. "Shame and Embarrassment Revisited." *Personality and Social Psychology Bulletin* 27: 104-117.
- Sabini, John and Mauri Silver. 1997. "In Defense of Shame: Shame in the Context of Guilt and Embarrassment." *Journal for the Theory of Social Behavior* 27(1): 1-15.
- Sabini, John, Michael Siepmann, Julia Stein and Marcia Meyerowitz. 2000. "Who Is Embarrassed by What?" *Cognition and Emotion* 14(2): 213-240.
- Sreenivasan, Gopal. 2013. "The Situationist Critique of Virtue Ethics." In *The Cambridge Companion to Virtue Ethics*, edited by Daniel Russell, 290-314. Cambridge: Cambridge University Press.
- Stanovich, Keith, E. and Richard F. West. 2000. "Individual Differences in Reasoning: Implications for the Rationality Debate?" *Behavioral and Brain Sciences* 23(5): 645–65.
- Talbert, Matthew. 2008. "Blame and Responsiveness to Moral Reasons: Are Psychopaths Blameworthy?" *Pacific Philosophical Quarterly* 89(4): 516-535.
- Tangney, June and Rowland Miller. 1996. "Are Shame, Guilt, and Embarrassment Distinct Emotions?" *Journal of Personality & Social Psychology* 70: 1256-1269.

A Conditional Defense of Shame and Shame Punishment

- Tangney, June, Jeffrey Stuewig and Andreas G. Martinez. 2014. "Two Faces of Shame: the Roles of Shame and Guilt in Predicting Recidivism." *Psychological Science* 25(3): 799-805.
- Taylor, Gabriele. 1985. *Pride, Shame, and Guilt: Emotions of Self-Assessment*. Oxford: Oxford University Press.
- Wakefield, Jerome. 1992. "The Concept of Mental Disorder: On the Boundary between Biological Facts and Social Values." *American Psychologist* 47(3): 373-388.
- Williams, Bernard. (1989/2006). *Shame and Necessity*. Berkeley: University of California Press.