Online Public Shaming: Virtues and Vices

We are witnessing increasing use of the Internet, particular social media, to criticize (perceived) moral failings and misdemeanors. This phenomenon of so-called 'online public shaming' could provide a powerful tool for reinforcing valuable social norms. But it also threatens unwarranted and severe punishments meted out by online mobs. Drawing on the work of John Locke, as well as on recent discussions of social norms, we analyze the dangers associated with the informal enforcement of norms, but also highlight the promise of this practice. We then consider two crucial conditions that online public shaming must meet in order to be justifiable: proportionality and accountability. We argue that these requirements are in fact frequently violated, rendering most cases of online public shaming unjustified. While the use of online public shaming against others' vices has some apparent virtues, it is currently rarely justified, given its own vices.

1. Introduction

The phrase "public shaming" brings to mind stocks, pillories, and rotten tomatoes. But the phenomenon of public shaming has kept pace with technological progress. Public shaming has gone digital. The Internet affords unprecedented opportunities to criticize those we consider to have done wrong. Ill-advised tweets, photos, and Facebook posts regularly go viral and attract streams of criticism. Moreover, individuals regularly take to the Internet to denounce conduct they have observed offline, and appeal to others to join them in their condemnation. The list of (in)famous cases of online public shaming gets longer by the month – from the very early case of "dog poop girl"[1] to a more recent example in which a video of a customer racially abusing a post office employee in London was put on Facebook and attracted nearly a million views.[2]

Many of those who have been the target of this kind of online public shaming have faced severe consequences – suffering deep distress, having great detail about their private lives made public, and losing their jobs. In telling the story of Lindsey Stone, who was publicly shamed for a joke photo in which she appeared to be acting disrespectfully at a military cemetery, Jon Ronson reports that Stone "fell into depression, became an insomniac, and barely left home for a year" (2015a, 202).[3] The

---

[1] "Dog poop girl" was a young woman in South Korea who, in 2005, refused to clean up the mess after her dog had defecated on a subway train. See http://knowyour-meme.com/memes/dog-poo-girl; Solove (2007, 1-3).

[2] This was then picked up by the mainstream media and reported in many media outlets, no doubt causing millions more people to see the video. See BBC News (2018).

[3] Ronson (2015b) also tells Stone's story.

severity of the consequences suffered by those shamed online is one of the main themes of Ronson's book about the phenomenon, which vividly recounts several shaming stories.

However, we should not be too quick to decry all online public shaming. Thousands of people participate in this practice, believing that it can be a force for good. Many consider themselves to be upholding valuable standards of conduct, and rightly censuring those who fall short of those standards. This kind of informal sanctioning is essential to ensuring the continuation of these standards. Indeed, some might argue that it would surely be *wrong* to allow racist, sexist, dangerous, and disrespectful conduct to go unchallenged. Would this not be to connive with wrongdoing, and perhaps even make those who overlook these offences complicit? Online public shaming can be an effective way to censure wrongdoing, draw attention to social ills such as racism and sexism, and deter future wrongdoers. As Samrat Hanif, who put the video of the post office worker being racially abused online, said, posting the video will hopefully mean that "this lady" – and others, we can assume – "will never do this again" (quoted in BBC News 2018).

This justification for public shaming holds that it is an effective way in which to enforce valuable social norms – rules of behavior that ought to govern our conduct in relation to one another. Public shaming can highlight violations of these norms, reaffirm the importance of compliance, help in causing the norm violator to repent, and deter future violations.

In some cases, at least, these beneficial results do indeed occur. In May 2015, Shauna Hunt was interviewing two sports fans live on-air, when she was interrupted

by a man who said into her microphone "Fuck her right in the pussy".[4] With the camera still rolling, Hunt proceeded to question a group of men who had cheered on the heckler. Asked to explain why he found it so funny, Shawn Simoes responded: "It's fucking awesome! You're lucky there's not a fucking vibrator in your ear!" The footage was posted online, and went viral. After having been shamed in this way, Simoes sent an apology letter to the journalist, made a donation to a male-led group seeking to make men take more responsibility to stop violence against women, and voluntarily attended sensitivity training (see Huffington Post Canada 2015). Irrespective of whether this instance of public shaming was justified all things considered, it would be a mistake to overlook these positives.

As we have already intimated, however, things do not always work out so well. In December 2013, on her way from New York to Cape Town to visit family, Justine Sacco tweeted "Going to Africa. Hope I don't get AIDS. Just kidding. I'm white!"[5] She meant this as an ironic comment about the attitudes of middle-class white Americans towards AIDS, but many observers considered it racist.[6] Sacco faced a barrage of online criticism, much of it threatening and abusive, and was fired from her job – all while she was on a plane, and so unable to defend herself or explain her tweet.

Various commentators have warned of the excessive nature of the punishments imposed through online public shaming. Daniel Solove writes that "Internet shaming has a tendency to become overzealous" (2007, 95). Further, it "creates an

---

[4] Our account here is drawn from Gillis (2015).

[5] Sacco's story is told in Ronson (2015a, 63-77); Ronson (2015c).

[6] For Sacco's explanation, see Ronson (2015a, 69).

indelible blemish on a person's identity. Being shamed in cyberspace is akin to being marked for life… People acquire permanent digital baggage" (2007, 94). In the same vein, Kate Klonick notes that "the punishments are both extreme and endless and seem deeply uncalibrated to the transgressions" (2016, 1054). The stories told by Ronson clearly support these claims.

So, what are we to make of online public shaming, given that it sometimes does good, yet so often is overzealous and excessive? Our aim in this paper is to draw upon the resources of political philosophy to illuminate both our reasons to welcome public shaming, due to its role in enforcing valuable social norms, and the rather stringent conditions that we must meet in order for instances of public shaming to be justified.

Our starting point is perhaps a surprising one; it is with the writings of John Locke. While the kinds of public shaming with which we are concerned in this paper use new technologies, the idea that we might have worries regarding the enforcement of moral rules through informal sanctioning is not a new one. Indeed, it is a central insight within Locke's *Second Treatise of Government*, and our first task will be to show how his analysis helps us to understand online public shaming as a phenomenon.

Before turning to that, however, it is important to note two limits of our inquiry. First, we focus on the potential role of public shaming in upholding *morally authoritative* social norms – extant social norms with which individuals have a duty to comply, such that violating them is wrongful. We say more about the nature of social norms and moral authoritativeness below (see Section 4.1). But we do not address the question of what explains why particular norms are authoritative, or

indeed why social norms can be authoritative at all. These deeper questions in moral philosophy fall beyond the scope of our inquiry. Instead, we hope it is uncontroversial that the examples that we use do involve morally authoritative norms, such as norms against racism and sexism. Our focus on morally authoritative norms also means that we do not discuss complexities that arise when norms are contested, or when public shaming is used in an effort to establish new social norms. We recognize that this excludes an important range of cases, which are worthy of sustained attention. But we think it is crucial first to address the more straightforward cases of online public shaming directed at violations of morally authoritative social norms. After all, if public shaming were always (or almost always) impermissible, even in those cases, when the target of shaming has violated a duty to comply with a morally authoritative norm, then it is very unlikely ever to be justified in the more complicated cases.

A second limit of our inquiry is that we do not examine precisely what constitutes 'shame' or whether 'online public shaming' is the correct term to use for our phenomenon of interest. These are interesting issues, but we lack space to explore them here. Instead, we use the term 'online public shaming' in the way that has become common, i.e. to refer to the use of the Internet to draw attention to and criticize (perceived) wrongdoing. Our aim is to identify pro tanto reasons in favor of such shaming, and to consider some of the stringent conditions that it must meet in order to be justified. We argue that using online public shaming against others' vices has some apparent virtues, but that it is in fact currently rarely justified, given *its own vices*.

2. Lessons from Locke

Locke argues that in the pre-political "state of nature" every individual has the right to enforce the law of nature, which is the moral law that each of us ought to follow. While the state of nature is "a state of liberty... it is not a state of license" (Locke 1689, II.II.§6).[7] It is governed by the law of nature, according to which each individual is bound to preserve both herself and the rest of mankind. Those who violate the law of nature, for example by harming another "in his life, health, liberty, or possessions" (Locke 1689, II.II.§6), may be rightfully punished in order to preserve the safety and security of all. Moreover, "Every man hath a right to punish the offender, and be executioner of the law of nature" (Locke 1689, II.II.§8). Without rights of enforcement, the law of nature would be in vain. And, given that the state of nature is a "state of perfect equality" (Locke 1689, II.II.§7), if any individual has a right to punish offenders then all must have that right.

Sensibly, Locke insists that the offender's punishment should not be arbitrary or unlimited. It must be "proportionate to his transgression, which is so much as may serve for reparation and restraint" (Locke 1689, II.II.§8). In other words, punishment has specific purposes: making reparations for harm to others and deterring future violations. Punishment should not go beyond what is necessary to fulfill these purposes (Locke 1689, II.II.§12; see also §§10-11).

Locke recognizes that there is a problem here, however, and he imagines someone objecting to his view along the following lines:

---

[7] By this, we mean Book II, Chapter II, §6. We use the same notational format throughout.

I doubt not but it will be objected, that it is unreasonable for men to be judges in their own cases, that self-love will make men partial to themselves and their friends: and on the other side, that ill-nature, passion and revenge will carry them too far in punishing others; and hence nothing but confusion and disorder will follow (Locke 1689, II.II.§13).

Locke readily admits that this is one of the great "inconveniences" of life in the state of nature. Various aspects of human psychology, including our very "passion" for exacting justice, make it likely that some will be excessively punished. Equally, others will escape punishment, due to our "partiality" toward ourselves and "unconcernedness" for others (Locke 1689, II.IX.§125).

This problem is exacerbated by two further complications that Locke overlooks. First, the content of the law of nature is complex, such that even individuals who act in good faith and who do not exhibit partiality toward themselves and unconcernedness toward others will sometimes disagree about whether a transgression has taken place.[8] This is because the nature of an individual's conduct is sometimes unclear, such as in the case of Sacco, whose comments were intended as ironic. This point is especially forceful once we recognize the existence of what John Rawls calls the "burdens of judgment" – that is, the many "hazards involved in the correct (and conscientious) exercise of our powers of reason and judgment in the ordinary course of political life" (1996, 55-56). Second, rather than unevenness in punishment being fairly distributed, it is instead likely to reflect racist, sexist, and classist

---

[8] In this way, we are doubtful regarding Locke's claim that the law of nature is "plain and intelligible to all rational creatures" (Locke 1689, II.IX.§124). Locke admits that bias will lead to disagreement, but seems to overlook reasonable disagreement.

biases, and so be more targeted against members of marginalized groups. Those who are more powerful, and so enjoy greater capacities for retaliation, are less likely to be punished than those who are comparatively powerless.

All three of these features are evident in contemporary examples of online public shaming, and so germane to our inquiry. First, those who engage in this practice typically demonstrate "partiality" toward themselves and "unconcernedness" for others. Second, it is sometimes unclear whether the target has in fact violated the relevant social norm. Third, participants often exhibit racist, sexist, and classist biases.[9]

Locke offers the pressing problems with private individuals enforcing rules and norms as a central reason that individuals would choose to leave the state of nature and enter political society. The establishment of "civil government is the proper remedy for the inconveniences of the state of nature" (Locke 1689, II.II.§13). Crucially, government provides a "known and indifferent judge, with authority to determine all differences according to the established law" (Locke 1689, II.IX.§125). Individuals give up their natural right to punish transgressors, in favor of establishing a centralized authority that promulgates and enforces laws impartially. Private enforcement of the law of nature is replaced by governmental enforcement.

This solution is not available to us, however, since we are explicitly focused upon *informal sanctions* to enforce social norms, rather than on state sanctions. Social norms are rules that govern the attitudes and behavior of members of a group,

---

[9] For example, studies also reveal that online abuse is disproportionately directed at females and members of ethnic minorities. See Gardiner et. al. (2016).

specifying how members of that group ought (not) to act.[10] They specify a public standard to which members expect one another to adhere. Social norms concern a wider range of conduct than that regulated by law. They include rules of etiquette, manners, fair play, respect, and so on. The norms that we are most concerned with are ones that regulate conduct that ought not be subject to legal regulation. In some cases, this is for principled reasons, such as when using the coercive power of the state would be self-defeating. In other cases, this is for practical reasons, such as when the law is too blunt an instrument, and so it would be unduly costly to penalize violations of the norm. Informal sanctioning within civil society through public shaming is a vital enforcement mechanism with respect to these norms. As with the law of nature within Locke's state of nature, if individuals lack the right to enforce social norms, then we might fear that that they will cease to be effective regulators of behavior.

This is not to suggest that the state has *no* role in regulating public shaming. Laws against libel and hate speech can prevent its worst excesses. The "right to be forgotten" might provide a way that people can escape from past norm violations and live free from the fear of future sanctions.[11] We should also consider further innovative policy solutions, such as laws that provide employees with greater pro-

---

[10] For further discussion of the nature of social norms and their importance to well-functioning societies, see Brennan et. al. (2013); McTernan (2014).

[11] The European Court of Justice endorsed the 'right to be forgotten' in 2014. See Warman and Barratt (2014).

tection against being dismissed for actions outside of work that are deemed to damage their employer's reputation and profitability.[12] Crucially, however, all of these laws are ways to make informal sanctioning less severe or problematic, rather than the state itself taking over the role of sanctioning, in the way that it does within Locke's account. While the state can endeavor to prevent certain wrongful forms of informal sanctioning, it cannot itself *be* the sanctioner, precisely because we are considering the enforcement of informal social norms.

3. The Role of Social Norms and Public Shaming

We will return to Locke later in the paper. The upshot of the previous section, however, is that though Locke helps us to identify the dangers of informal sanctioning, his solution is not available in our context. This seems to leave us with two options. We could either abandon public shaming altogether or seek to improve it. The first of these options might seem attractive at first, since it means that we do not face the many problems associated with this kind of individualized enforcement. But this would also bring considerable costs. In particular, it risks undermining valuable social norms, which are essential for ensuring peaceful and cooperative coexistence. To see why this might be a problem, we must say more about the role of social norms and the goods that we can realize through their enforcement.

As Solove notes, "Norms bind societies together; they regulate everyday conduct; they foster civility. They are the oil that reduces the friction of human interaction... In short, norms are a central mechanism through which a society exercises

---

[12] We thank XX for this suggestion.

social control" (2007, 6). This kind of "social control" is crucial for the smooth functioning of all societies. While laws play an important role in this regard, they are not sufficient. We also need informal social norms, which create stable expectations and facilitate cooperation and concord. Analysis of sexism and racism helps to illustrate this point. Even though sexist and racist discrimination is almost universally legally prohibited, pervasive social norms that unjustly privilege the interests of white men continue to exist. It is widely accepted that overcoming racial and gender injustices requires us to develop and maintain better, egalitarian, social norms.[13]

External sanctions, such as praise, blame, commendation, and criticism, are vital for maintaining social norms. When norms are functioning well, members of a group take them to be authoritative or obligatory. They accept and internalize the norms, such that they consider themselves duty-bound to comply with them, and are likely to feel guilt or shame if they violate them. Nonetheless, the internal sanctions generated by these moral emotions are generally insufficient to motivate compliance with the norms, especially over a longer period of time. External sanctions are needed to fortify the moral capacities of individuals who cannot reliably depend on internal sanctions doing the job.[14] It is this combination of internal and external sanctions enforcing social norms that makes them effective as regulators of individuals' behavior. As Emily McTernan notes, empirical research shows that "social norms are powerful determinants of behaviour, and secure stable patterns of behaviour from the majority of those who internalise the norm" (2014, 95).

---

[13] For example, see Gheaus (2012); Gheaus (2018).

[14] For discussion of the idea of moral fortification, see Howard (2017).

Public shaming is one form of external sanction that can uphold and reinforce morally authoritative social norms. In doing so, it plays several valuable roles.[15] Most obviously, it makes the offender aware that she has violated a social norm. Ideally, this will lead her to recognize that she has acted wrongly, feel remorse, apologize, seek to make appropriate amends, and commit to complying with the norm in future. This is the "reparative" role that Locke identifies. The second role of punishment identified by Locke is "restraint", or deterrence. Public shaming can reduce future norm violations, by both the norm violator and others, since individuals are less likely to disobey social norms if they will be externally sanctioned for doing so. In this way, public shaming helps to protect potential victims against future violations. Further, it also demonstrates solidarity with those who are wronged by a norm violation, showing them that others are unwilling to let the wrongdoing go unnoticed and uncriticized. Finally, public shaming provides a way that we can express our endorsement of valuable social norms, thus strengthening our shared sense of commitment to those norms, and the values they promote or respect.

We can observe each of these points in the case of Shawn Simoes, who made sexist remarks to Shauna Hunt live on-air. First, the public shaming of Simoes helped to ensure that he was aware of the seriousness of his wrongdoing, for which he later apologized. Second, his public shaming provides a useful reminder of the fate that an individual may suffer by acting in this way. No doubt, this has deterred both Simoes and others from acting similarly in the future, and, in this way, it has helped to protect potential victims against similar kinds of sexist conduct. Finally, those who

---

[15] We disentangle these distinct roles more carefully and in greater depth in our [reference removed].

shamed Simoes have demonstrated their solidarity with Hunt, as well as strengthened their shared commitment to stamp out this kind of sexist behavior.

Part of what enables online shaming to play these various roles is its public nature. Clearly, shaming must be public if it is to strengthen our shared commitment to the norm and deter other prospective violators. Further, holding unacceptable behavior to public view and inviting others to affirm the wrongness of that behavior is perhaps more likely to cause the offender to feel remorse and seek to make amends, because they are likely to experience a greater degree of guilt and shame than would be the case otherwise. Private criticism can also serve some of these roles, of course, and is another important means of social norm enforcement. But the public nature of online shaming enables it to serve some roles more effectively, and to serve other roles that private criticism cannot.

The upshot of this analysis is that we have reasons to welcome public shaming, given the way in which it can enforce valuable social norms. Moreover, one might have high hopes for *online* public shaming in particular. More information is open to public view than ever before, and we can disseminate that information globally and instantaneously. This enables ordinary people to hold one another to account, uphold valuable social norms, and promote good conduct. The decentralized nature of online shaming means that it can reflect values shared by ordinary users of the Internet, rather than the views of powerful elites. Online public shaming might

appear to be an egalitarian, or even democratic, form of public accountability that can give voice to members of marginalized groups.[16]

This rosy picture of online public shaming does not seem to match present reality, however. Nothing that we have said eliminates the worries that Locke helped us to identify or that we saw in some of the examples above. Even if public shaming *can* play valuable roles, this does not mean that it always, or even often, *does*. We should still have grave concerns about the large number of recent cases where public shaming has been excessive, caused severe distress, and seemingly given people an excuse to let out their "ill-nature" (to use Locke's term) by insulting and threatening their targets. Even if public shaming, including online shaming, *could* be a valuable practice, experience suggests that it is in fact a dangerous one, just as Locke might have expected. Nonetheless, abandoning the practice altogether is not an unambiguously attractive option, for the reasons we have highlighted in this section.

These remarks point to the second option we mentioned: we might seek to improve our practices of online public shaming, by developing principles and guidelines to govern the practice and establish when, and in what forms, public shaming is and is not justifiable. In other words, there is a need for us to develop a moral framework that we can use to assess the justifiability of instances of online public shaming. The rest of this paper seeks to contribute to this endeavor, by discussing

---

[16] Here, we put aside worries relating to the fact that many websites use algorithms to structure the information that users receive, and so the ability to share information is not so decentralised as it may first seem. For discussion, see Hern (2016).

two principles that will be central to any such framework: proportionality and accountability.

4. Proportionality

We saw earlier that Locke held that the punishment of those who violate the law of nature must be proportionate to the aims of that punishment. The same applies with respect to the public shaming of those who violate social norms. An instance of public shaming is justifiable only if it is proportionate. This requires that its negative consequences are not excessive in comparison with its positive consequences. The positive consequences here are those that we identified in the previous section – causing remorse, deterring future violations, strengthening our commitment to and compliance with valuable social norms, and so on. The negative consequences include the psychological and reputational harms to the shamed individual, as well as any material costs that follow, such as the loss of her job. Klonick, Ronson, and Solove all argue that these negative consequences have been disproportionate to the positive consequences in many cases of online public shaming.

Determining whether this proportionality requirement is fulfilled in a particular case is a complex matter, which will depend on various contextually-specific factors. Nonetheless, it remains possible to make a number of more general points.

4.1. *Morally authoritative social norms*

Public shaming that enforces a social norm can be proportionate only if that norm is morally authoritative. If the relevant social norm is not morally authoritative, then

there is no value in its enforcement or reinforcement, so imposing burdens on people by shaming them is always disproportionate.[17] Social norms are morally authoritative when individuals are under a duty to comply with them. Some norms generate such duties by reflecting pre-existing moral principles. For example, we can justify norms against sexist and racist speech on the grounds that sexist and racist speech is itself wrongful. Of course, what counts as sexist or racist speech – for example, what epithets are sexist or racist – will often be determined by the particular social context. This is part of what makes Sacco's case so difficult.

Other norms relate to pre-existing moral principles in more complicated ways. For example, we might think that it is morally desirable, or even required, that there be some generally recognized way of deciding the order in which people will get onto buses, so that passengers do not come into conflict with one another. There is no moral requirement that this order be set by a queue that is formed on a first-come-first-served basis. We might imagine a society in which some other fair system is used instead. However, if the prevailing social norm is indeed queuing, then we should take it as morally authoritative, since it facilitates peaceful bus-boarding.[18]

---

[17] This conclusion parallels Jeff McMahan's claim that all unjust wars are disproportionate (2016).

[18] Moreover, this example also shows that not all social norms that are morally authoritative are morally required, or even morally optimal. For example, let's suppose we were to discover a new method of allocating seats on buses that was marginally fairer and more efficient than queuing. This would call into question whether the current social norm is morally optimal. But it would not – immediately, at least – call into question whether compliance with the existing social norm is morally required.

Various norms of politeness and respect operate in this way: it is morally desirable to have *some* norm regarding what behavior is polite or respectful in various situations, and compliance with whatever norm comes to prevail within a particular context is then morally required, providing that norm plays its role sufficiently well.

An implication of this discussion is that different norms can be present and morally authoritative in different contexts or within different communities.[19] For example, there might be different norms online and offline. This can create difficulties when online public shaming is used in response to conduct that occurred offline. In order for such shaming to be proportionate, the norm that was breached must have been morally authoritative in the context in which the targeted individual acted, rather than being a norm that is distinctive to online interaction. Proportionality requires the target of public shaming to have violated a duty.

4.2. *Narrow proportionality*

Some instances of public shaming are proportionate because they impose burdens upon norm violators that those individuals are *liable* to suffer, due to their violation of the norm. This is *narrow proportionality* (McMahan 2009, 20-24; Tadros 2011, 356-359).

A norm violator can be liable only if she has *culpably* violated the norm. This requires that she must have been able to gain knowledge of the relevant norm and that her conduct breached that norm in a way that she could have reasonably foreseen. Her violation of the norm can thus be considered to reflect an "ill will" toward

---

[19] Some of the complications that this creates are discussed in Valentini (forthcoming).

others. This culpability condition was clearly met in the case of Simoes, whose conduct was uncontroversially sexist, and the post office customer, whose conduct was uncontroversially racist. It is less clear that it was fulfilled in the case of Sacco, whose comments were intended to be ironic.

When thinking about public shaming's negative consequences, one factor that will affect the magnitude of the reputational and psychological harms imposed upon the norm violator is the size of the audience before whom she is shamed, and the number of people who then participate in the shaming. Though we have been talking about public shaming in a general way, it is important to recognize that *publicity* comes in degrees. There are big differences between an individual being criticized in front of a small group of friends or colleagues and her being chastised before the general public, through the media or Internet. One of the primary concerns with online public shaming is the fact that it can reach such a large audience, whereby thousands, or even millions, of people know about and condemn a norm violation. Many who have been the target of online public shaming describe a sense of the whole world being against them, and the intense shame, humiliation, and distress that this causes (see Ronson 2015a). These severe consequences are one of the reasons that online public shaming will often be narrowly disproportionate. Even if individuals are liable to suffer some negative consequences due to their norm violation, they will rarely be liable to suffer the kind of distress and humiliation that online public shaming regularly brings about.[20]

---

[20] There are important questions about how to understand narrow proportionality in cases that involve the systematic accumulation of smaller harms. We discuss this issue in our [reference removed].

Another factor that makes a significant difference to whether public shaming is narrowly proportionate is whether it is *reintegrative*.[21] Toni Massaro provides an account of shaming practices in pre-World War II Japan and colonial America, in which she argues that shaming was "not understood in these communities as an expression of contempt that permanently renounces or expels the shame[d] member" (1997, 682). Instead, it showed that norm violations will be criticized and punished, but with the aim of reconciling the violator and the community. Indeed, "the bond between the community and the shamed member is reinforced… rather than severed" (Massaro 1997, 682). Healthy forms of public shaming aim at, and make possible, the reintegration of the norm violator back into the community, rather than permanently stigmatizing them.

In the same vein, John Braithwaite draws a distinction between reintegrative and disintegrative shaming in the context of state-based punishment (2000). Reintegrative shaming seeks, and is receptive to, repentance on the part of the shamed. It functions as a means of moral education, seeking to bring the offender to an awareness of the wrongness of her conduct and to produce reformation. In contrast, disintegrative shaming stigmatizes and degrades the offender, posing a threat to her very identity and providing no means of reconciliation.

---

[21] In emphasising that reintegration makes a difference to whether public shaming is proportionate, we do not mean to deny that reintegration might also be important for other reasons or that it serves as an independent constraint of justifiable public shaming. Indeed, we argue elsewhere that it is an independent constraint, as well as a contributor to proportionality. See our [reference removed].

Online public shaming can alienate its targets from both online and offline communities.[22] It should thus make possible reintegration into whichever communities are relevant – that is, into any community in which the individual has lost their good standing. This also involves reintegration both with the shamers themselves and with the wider audience in whose eyes the individual's standing has been lowered.[23]

Shaming will almost always be narrowly disproportionate when it is not reintegrative. Yet online public shaming struggles in this respect, since in order to be reintegrative it must be carried out in a way that makes clear that it does not constitute a permanent rejection of the individual. There are limits to how effectively this can be done using media such as Twitter, which severely limit the space for explanation and are not well-suited to nuance. Online public shaming also often involves abuse and threats, which further undermine reintegration, as well as being morally reprehensible in themselves.

Beyond this, the very fact that so many people participate in the criticism and express their disapproval makes it harder for online shaming to be reintegrative.

---

[22] Some online public shaming affects its target's standing only in the online community, even when the shamed conduct occurs offline. But often online shaming has clear offline effects as well, as we have seen.

[23] There is a further question about whether public shaming should facilitate the reconciliation of a norm violator and the *victim(s)* of his act. To answer this question, we would need to know more about the morality of forgiveness and, in particular, whether there are duties to forgive. We do not commit to a position on this here, beyond tentatively suggesting that online public shaming should not make such reconciliation much more difficult or costly.

Explanations, defenses, and apologies are often caricatured or go unnoticed among the torrent of denunciation. Overcoming these barriers to reintegration is essential if online public shaming is to be more often justifiable.

4.3. *Wide proportionality*

All of our comments thus far have focused on narrow proportionality, in which the burdens fall on those who are liable. However, public shaming also imposes negative consequences on those who are not liable. Reputational and psychological harms often fall on third parties, such as the norm violator's associates (her friends, family, employers etc.). Moreover, the negative consequences that fall on the norm violator can often exceed that to which she is liable, as we have highlighted.

Nonetheless, it could be that the overall good consequences of an instance of public shaming greatly exceed the negative consequences that are imposed upon those who are not liable. This might justify the public shaming. In considering whether this is the case, we are examining *wide proportionality*.

Locke did not recognize the need for a separate standard of wide proportionality. However, it is an important component of a complete account of proportionality, since sanctions almost always affect individuals who are not liable. It is inevitable that the family and friends of a sanctioned individual will suffer at least some harm. This is true with respect to the state's enforcement of laws (Tadros 2011, 356-359), it would be true with respect to the enforcement of natural law within Locke's state of nature, and it is also true with respect to public shaming.[24] We should not

---

[24] It is also true with respect to military actions during war, which is the context in which much contemporary discussion of proportionality occurs. See McMahan (2009, 20-24).

conclude that sanctions can never be justifiable when they harm third parties in this way, since this would render practically all enforcement of laws and social norms impermissible. Instead, we should recognize a distinct standard of wide proportionality, which compares the overall positive consequences of imposing sanctions to the negative consequences imposed upon those who are not liable.

The standard for wide proportionality is much more demanding than that for narrow proportionality. This is because negative consequences are much more difficult to justify when they fall on those who are not liable. The positive consequences of public shaming must be *much* more valuable than the negative consequences in order for them to be widely proportionate.

This is not to say that this standard is impossible to meet. In at least some cases of online public shaming, the harms to third parties may be fairly small, and the benefits with respect to the enforcement and upholding of valuable social norms great. Similarly, these benefits might sometimes be sufficient to justify imposing burdens upon the norm violator that exceed those she is liable to suffer. We suspect that such cases will be fairly rare, but perhaps this is true in the case of Shawn Simoes.

There is a further important worry regarding the wide proportionality of online public shaming, however, which concerns its systemic effects. Even if some individual instances of online public shaming initially appear proportionate, the regular practice of public shaming might have a 'chilling effect'. The idea here is that the fear of being publicly shamed will lead many to retreat from online interactions, or to refrain from posting anything that could possibly be seen as controversial. The online environment would thus become far less active and discursive, in a way that

harms everyone. This kind of worry has been expressed by Ronson. He quotes from a friend who said that he would not dare post many of his jokes and observations online anymore. The friend told Ronson: "I suddenly feel with social media like I'm tiptoeing around an unpredictable, angry, unbalanced parent who might strike out at any moment" (2015a, 268). The actor and comedian Stephen Fry has also expressed this worry. In February 2016, Fry quit Twitter, after he faced online criticism for a joke he told while presenting the British Academy of Film and Television Awards (BBC News 2016). In a blog post explaining his decision, Fry complained that Twitter had become "a stalking ground for the sanctimoniously self-righteous who love to second-guess, to leap to conclusions and be offended – worse, to be offended on behalf of others they do not even know" (Fry 2016).

These concerns are especially problematic when the chilling effect occurs through the silencing of members of marginalized groups (rather than of upper-class, white men, such as Fry). This is for two reasons. First, it is because it is members of these groups that have the most to contribute in terms of making our online environment more discursive. Second, it is because the participation of members of marginalized groups is vital to any defense of online public shaming that appeals to its credentials as an egalitarian, or even democratic, form of public accountability.

The idea that the over-zealous enforcement of social norms might have this kind of chilling effect, and ultimately limit individuality and free expression, is not new. Indeed, it is one of the central claims in John Stuart Mill's *On Liberty* (1859, especially chapter III). Mill argues that individual freedom is threatened not only by the encroachment of the state, but also by civil society's informal sanctions. He writes that

> It is not by wearing down into uniformity all that is individual in them-
>
> selves, but by cultivating it and calling it forth, within the limits imposed
>
> by the rights and interests of others, that human beings become a noble
>
> and beautiful object of contemplation (Mill 1859, Chapter III, §9).

For Mill, the development of individuality was being stifled by undue levels of societal regulation, including through public shaming, in 19th century Britain. Ronson and Fry both claim that these dynamics also pose a grave threat in the contemporary online world.

These concerns arise mainly because there is currently such a high volume of online public shaming that is disproportionate. If shaming occurred only when it was proportionate, then Mill's and Ronson's worries about its systemic effects would have less force. That is, their warnings have force largely due to the fact that there is currently so much *unjustified* public shaming. But that is indeed the present reality, and in that context it might well be that even instances of online public shaming that initially appear proportionate are contributing to an overall practice that has extremely damaging systemic effects, and is thus disproportionate. If this is right, then online public shaming will be justifiable much less often than we might otherwise suppose.

4.4. *Implications*

Much contemporary online public shaming is disproportionate, in both the narrow and wide senses. What can be done about this? In our view, the most important response should be an increase in *awareness* regarding the risks of disproportionality,

and thus of public shaming being unjustified. Mill wrote *On Liberty* in order to entreat his readers to recognize the proper limits on society's control over the individual. Similarly, we hope that an increased awareness of the risks of disproportionality will lead people to be more hesitant to engage in online public shaming. Importantly, this hesitancy should extend even to cases where there clearly has been a culpable breach of a morally authoritative social norm. Even in those cases, it might well be that adding one's voice (or re-tweet) to the criticism of the norm violator will contribute to her and/or others suffering disproportionate harms.

In other words, we need new social norms regarding online activity: social norms that encourage us to be slow to conclude that an individual is culpable, to show restraint in criticism, to refrain from seeking disproportionate punishment, and to be willing to overlook past infractions, so as to allow people to be restored into the online community in good standing. While our concern in this essay is with the disproportionate enforcement of social norms, it is social norms that also provide the most promising solution.

Greater awareness alone is unlikely to be sufficient, however. We also need to consider whether certain negative consequences are almost always disproportionate, and thus ought to be ruled out completely – and perhaps legally regulated. One such consequence is the norm violator losing her job. Both Sacco and Simoes were fired, though the latter was re-hired after an arbitration purpose. While it is perhaps understandable that firms want to distance themselves from employees who have been shamed, this will almost always impose disproportionate burdens

on the norm violator and their dependents.[25] For this reason, we should consider tighter legal regulation in this domain, granting employees stronger protection against being fired.

5. Accountability

Another concept that is central to justified public shaming is *accountability* – specifically, the accountability of shamers. It is highly desirable that those who shame others can themselves be held to account, taking responsibility for their criticism and its effects, and being open to responses, discussion, and correction. Certainly, the target of shaming should have an effective right of reply, and ideally other individuals will do too.

5.1. *Locke, again*

Accountability can help to mitigate the problems that Locke identified, since it means that norm enforcers must be prepared to justify their actions. They must explain why the norm violator ought to be criticized, by showing that a valuable social norm has in fact been violated. They also must listen to other points of view, and consider the possibility that they might be mistaken. This helps to combat the "partiality" and "passion" that Locke considered central problems with informal sanctioning. Further, it acts as a safeguard against inaccurate and unwarranted criticism,

---

[25] Of course, there are notable exceptions to this, such as when an individual violates an important professional norm. This may include cases in which a doctor violates confidentiality norms or an academic has sexual relations with one of his students.

since it both deters people from engaging in such criticism and makes it easier for errors to be corrected.

Indeed, accountability is part of Locke's solution. Locke argues that the inconveniences of the state of nature would lead people to establish a public authority that lays down clear laws and then enforces those laws in a transparent way. Laws are enforced by impartial judges, those accused of crimes are able to offer a defense, and those who are convicted are able to appeal. Moreover, Locke argues that if the public authority ceases to govern in a fair way then the people can hold it to account by revolting against it, and replacing the unjust government.[26]

The formalized accountability that features in Locke's solution is not available in our context, since we are focusing on the justifiability of informal sanctioning within civil society. Nonetheless, the idea that those who sanction others must themselves be able to be held to account is an important one. This is a feature that is often missing from contemporary online public shaming. Indeed, one of the striking features of Sacco's case is that her tweet went viral whilst she was on a plane, accompanied by the hashtag #HasJustineLandedYet (Ronson 2015c). She had been the number one worldwide trend on Twitter, and lost her job, before her plane landed, and thus before she was able to offer any kind of explanation or defense.

As with proportionality, increasing accountability within online public shaming is centrally a matter of developing new social norms. We should seek to create a culture in which those who engage in public shaming make themselves accountable to others. In other words, they should recognize and respect others' right of reply,

---

[26] For discussion, see Simmons (1993, ch. 5).

and especially that of the target of their criticism. Those who accuse others of violating social norms should be willing to listen to the other side of the story and consider whether their criticisms might be misplaced.

5.2. *Anonymity*

Arguably, one way to increase accountability online is to prohibit, or otherwise prevent, anonymity. Anonymity makes it more difficult for the shamed, and others, to exercise their right of reply, since anonymous shamers shield themselves from criticism by concealing their identities. Anonymous shamers are also likely to be emboldened in a way that can lead to excessive criticism. Indeed, there is empirical evidence showing that those who write under the cloak of anonymity are more likely to be uncivil. Arthur Santana studied the comments on online newspaper articles, and found that anonymous comments were significantly more likely to include personal attacks, threats, abusive or hateful language, epithets, and racist sentiments. As Santana notes, his study supports the conclusions of recent psychological research, according to which "anonymity can foster a sense of impunity, loss of self-awareness and a likelihood of acting upon normally inhibited impulses in a way that is markedly inconsistent with a person's offline self" (2014, 23). Anonymity runs counter to the kind of accountability-conducive social norms that are needed in order for online public shaming to be justifiable.

Interestingly, some online fora do not permit anonymity. Some newspaper websites require commenters to use their real names, and Facebook requires that users use their authentic name and identity. Facebook explicitly justify this policy on accountability grounds: "When people stand behind their opinions and actions

with their authentic name and reputation, our community is more accountable".[27] We should certainly be skeptical of the sincerity of this justification, given the way in which this policy enhances Facebook's ability to generate more accurate marketing data. Nonetheless, the policy might serve to increase accountability. For example, Samrat Hanif was easily identifiable as the source of the post office video, and was asked to comment on his decision to post the video online by several media outlets.

However, there are also considerations in favor of anonymity. The ability to write anonymously allows individuals to express unpopular views with less fear of being personally attacked or criticized. People can have legitimate reasons for wanting to be able to express views or reveal facts about themselves without being identifiable. Those who are more vulnerable might be more likely to interact online if they can do so anonymously. In support of this, there is evidence that links anonymity to increases in the number of participants and the range of views articulated within online discussions (McCluskey and Hmielowski 2011).

Perhaps ironically, these considerations in favor of anonymity are persuasive partly because online discussion and criticism is so often conducted in an objectionable way. There would be less need for anonymity if disproportionality were not so rife. Though anonymity has been shown to contribute to these problems, it might also be justified in light of those same problems. At the very least, anonymity's benefits are greater in the unjust world in which we live.

---

[27] Facebook's 'Community Standards', available at

https://www.facebook.com/communitystandards.

In some contexts, there may be a middle way. The popular philosophy blog Daily Nous allows anonymous comments, but only on two conditions.[28] First, anonymous commenters must still submit a real email address, which is not displayed publicly but is seen by the website moderator. Second, anonymous commenters must use a consistent name (or "handle") in all of their comments.[29] These conditions ensure that individuals can be associated with all of their comments and that the moderator can contact even anonymous commenters, if necessary. This policy attractively combines the benefits of anonymity with the virtues of accountability. An obvious drawback, however, is that it requires an active moderator, who can enforce the policy. This might not always be possible. Nonetheless, we would enhance accountability if the prevailing social norm encouraged individuals to adopt a consistent online identity.

5.3. *Collective Harms*

Even putting all of these problems aside, a further issue remains. Online public shaming is carried out by a large number of disparate individuals, many of whom do little more than re-tweet another's criticism. While the overall burdens suffered by the shamed might be great, the contribution of each individual shamer is very small. Whereas Locke's public authority provides a single institution that can be held accountable for its enforcement of the law, the decentralized nature of online public

---

[28] See http://dailynous.com/comments-policy/.

[29] This prevents so-called "sockpuppeting". See
https://en.wikipedia.org/wiki/Sockpuppet_(Internet).

shaming means that no particular participant can be held to account for the full extent of an instance of shaming.

Public shaming is not the only phenomenon that has this feature. In cases of so-called "collective harms", the actions of a large group of individuals cause harm, yet the contribution of each individual is negligible (Kagan 2011; Nefsky 2012; Spiekermann 2014). In many cases of collective harm, it seems that no individual's action makes any difference to the outcome. Whether or not one more individual participates in online public shaming probably makes no perceptible difference to the negative consequences for the shamed. Yet the overall effect of these individuals' actions is large. As several philosophers have pointed out, some of the most important problems facing humanity today, such as anthropogenic climate change and unjust working conditions, share this "collective harm" structure.

There are at least two questions that arise when we consider cases of collective harms. First, can we hold collectives accountable for harmful outcomes when they are not organized and lack structures for collective decision-making (Miller 2007, ch. 5; Lawford-Smith 2015)? Second, can we view an individual as having acted wrongly in cases where a wrong occurred but her individual action did not make any difference to the outcome caused by the group as a whole? We lack space to examine the rich philosophical literature discussing these questions. Suffice to say that these debates are highly relevant to a normative assessment of online public shaming, and one's views on these questions will determine whether one believes that shamers can be held accountable for the effects of instances of shaming. If they cannot be properly held accountable, then this might lead one seriously to question

whether online public shaming is ever justifiable. This is a particularly strong conclusion because the nature of public shaming as a collective harm is a structural feature of the practice, rather than something that can be altered through regulation or better social norms. The importance of the collective harm issue with respect to public shaming should thus be obvious.

6. Conclusion

We have not been able to discuss all of the normatively salient features of online public shaming in this paper. However, we have highlighted some of the central elements that any assessment of online public shaming must contain, and have indicated the ways in which these elements affect the justifiability of the practice.

As we have sought to emphasize throughout, social norms are vital to our living civilized lives together. Their enforcement through public shaming, including online, can serve justified purposes, by upholding valuable norms. However, online public shaming also brings great risks of unaccountable, unrestrained, mob-like attacks on individuals. Such condemnation can be disproportionate, and thus unjustified, even if its target has culpably violated a morally authoritative social norm. Creating the conditions where social norms can be upheld online in valuable ways is itself a matter of creating new and better social norms regarding our online conduct. Social norms could thus provide the solution to the problem of their own enforcement. Whether such norms will in fact develop, however, is very difficult to predict. Changing the practices of thousands, or even millions, of participants in online interactions is plainly a huge task – but it is not an impossible one. If we are correct,

then, one of the central implications of this paper is that further research into the way that online norms change and develop is urgently needed.

References

BBC News. 2016. "Stephen Fry "Quits" Twitter Over Bafta Jibe." February 15, 2016. http://www.bbc.co.uk/news/entertainment-arts-35577913.

BBC News. 2018. "Racist Rant at Camden Post Office Staff Caught on Film." February 9, 2018. https://www.bbc.co.uk/news/uk-england-london-43003120.

Braithwaite, John. 2000. "Shame and Criminal Justice." *Canadian Journal of Criminology* 42 (3):281–98.

Brennan, Geoffrey, Lisa Eriksson, Robert E. Goodin, and Nicholas Southward. 2013. *Explaining Norms*. Oxford: Oxford University Press.

Fry, Stephen. 2016. "Too Many People Have Peed in the Pool." February 15, 2016. http://www.stephenfry.com/2016/02/peedinthepool/.

Gardiner, Becky, Mahana Mansfield, Ian Anderson, Josh Holder, Daan Louterand, and Monica Ulmanu. 2016. "The Dark Side of Guardian Comments." *The Guardian*, April 12, 2016. https://www.theguardian.com/technology/2016/apr/12/the-dark-side-of-guardian-comments.

Gheaus, Anca. 2012. "Gender Justice." *Journal of Ethics and Social Philosophy*, 6 (2): 1–24.

Gheaus, Anca. "Gender." In Serena Olsaretti. ed., *The Oxford Handbook of Distributive Justice*. Oxford: Oxford University Press, 389–411.

Gillis, Charlie. 2015. "He's Fired. Who's Next?" *Maclean's*, May 22, 2015. http://www.macleans.ca/work/jobs/hes-fired-whos-next/.

Hern, Alex. 2016. "Angry about Facebook Censorship? Wait Until You Hear about the News Feed." *The Guardian*, May 11, 2016. https://www.theguardian.com/technology/2016/may/11/facebook-censorship-news-feed-trending-topics.

Howard, Jeffrey W. 2017. "Punishment as Moral Fortification." *Law and Philosophy* 36 (1): 45–75.

Huffington Post Canada, The. 2016. "Shawn Simoes, Man Fired For FHRITP Incident, Rehired By Hydro One." November 2, 2015. http://www.huffingtonpost.ca/2015/11/02/shawn-simoes-hydro-one_n_8453012.html.

Kagan, Shelly. 2011. "Do I Make a Difference?" *Philosophy & Public Affairs* 39 (2): 105–41.

Klonick, Kate. 2016. "Re-Shaming the Debate: Social Norms, Shame, and Regulation in an Internet Age." *Maryland Law Review* 75 (4): 1029–1065.

Lawford-Smith, Holly. 2015. "What 'We'?" *Journal of Social Ontology* 1 (2): 225–49.

Locke, John. 1689. *Two Treatises of Government.*

Massaro, Toni M. 1997. "The Meanings of Shame: Implications for Legal Reform." *Psychology, Public Policy, and Law* 3 (4): 645–704.

McCluskey, Michael, and Jay Hmielowski. 2011. "Opinion Expression During Social Conflict: Comparing Online Reader Comments and Letters to the Editor." *Journalism* 13 (3): 303–19.

McMahan, Jeff. 2009. *Killing in War*. Oxford: Oxford University Press.

McMahan, Jeff. 2016. "Proportionality and Necessity in Jus in Bello." In Helen Frowe and Seth Lazar, eds., *The Oxford Handbook of Ethics of War*. Oxford: Oxford University Press, 418–437.

McTernan, Emily. 2014. "How to Make Citizens Behave: Social Psychology, Liberal Virtues, and Social Norms." *Journal of Political Philosophy* 22 (1): 84–110.

Mill, John Stuart. 1859. *On Liberty*.

Miller, David. 2007. *National Responsibility and Global Justice*. Oxford: Oxford University Press.

Nefsky, Julia. 2012. "Consequentialism and the Problem of Collective Harm: A Reply to Kagan." *Philosophy & Public Affairs* 39 (4): 364–95.

Rawls, John. 1996. *Political Liberalism*. New York: Columbia University Press.

Ronson, Jon. 2015a. *So You've Been Publicly Shamed*. London: Picador.

Ronson, Jon. 2015b. ""Overnight, Everything I Loved Was Gone": The Internet Shaming of Lindsey Stone." *The Guardian*, February 21, 2015. https://www.theguardian.com/technology/2015/feb/21/internet-shaming-lindsey-stone-jon-ronson.

Ronson, Jon. 2015c. "How One Stupid Tweet Blew Up Justine Sacco's Life." *The New York Times*, February 12, 2015. https://www.nytimes.com/2015/02/15/magazine/how-one-stupid-tweet-ruined-justine-saccos-life.html.

Santana, Arthur D. 2014. "Virtuous or Vitriolic: The Effect of Anonymity on Civility in Online Newspaper Reader Comment Boards." *Journalism Practice* 8 (1): 18–33.

Simmons, A. John. 1993. *On the Edge on Anarchy: Locke, Consent, and the Limits of Society*. Princeton, NJ.: Princeton University Press.

Solove, Daniel J. 2008. *The Future of Reputation: Gossip, Rumor, and Privacy on the Internet*. New Haven: Yale University Press.

Spiekermann, Kai. 2014. "Small Impacts and Imperceptible Effects: Causing Harm with Others." *Midwest Studies in Philosophy* 38 (1): 75–90.

Tadros, Victor. 2011. *The Ends of Harm: The Moral Foundations of the Criminal Law*. Oxford: Oxford University Press.

Valentini, Laura. Forthcoming. "Respect for Persons and the Moral Force of Socially Constructed Norms." *Noûs*.

Warman, Matt, and David Barratt. 2014. "Google must delete your data if you ask, EU rules." *The Telegraph*, May 13, 2014. http://www.telegraph.co.uk/technology/google/10827005/Google-must-delete-your-data-if-you-ask-EU-rules.html.