The Problem of Public Shaming[*]

<Pre-proof version: Final version in *Journal of Political Philosophy*, 22(8): 188-208>

Harrison Frye
Department of Political Science
University of Georgia

Even in its abstract and generalized form, the following should be a familiar script: A person appears to do something wrong. That something wrong gets broadcast on social media. This particular bit of wrongdoing gets traction and is spread through social media alongside messages condemning the wrongdoer (condemnation takes on varying degrees of severity). Shamers call for the firing of the wrongdoer in light of the wrongdoing. The employer obliges. The wrongdoer not only suffers the emotional pain of facing public shaming, but also incurs significant material costs. Worse still, the wrongdoer now bears stigma by search engine. Any potential friend or employer can see the public shaming should they search the person's name on the internet.[1]

---

[1] For a journalistic overview of many cases of online shaming, see Ronson 2015.

The details of such cases obviously matter. Nonetheless, I take it that most find

public shaming on the internet unnerving and unsettling. Even if you agree that the

target of such shaming did something wrong, you may nonetheless think something is

amiss in these cases.

The common response to public shaming on the internet is that it often goes too

far. The psychologists Paul Bloom and Matthew Jordan, for example, point out that,

"Even though any particular comment on its own may have caused you little or no

pain, the aggregate effect is far more severe."[2] This is a problem of proportionality.[3]

Proportionality undoubtedly captures part of what is wrong in many cases of online

shaming. However, I believe proportionality is ultimately a symptom of the underlying

problem. If we focus on symptoms, this will lead to misdiagnosis. Mistaking symptom

for cause will lead us astray when trying to fully understand and address the problems

of public shaming.

This paper grounds the problem of public shaming in problems of scale. My

starting assumption is that public shaming is a form of social norm enforcement.

Specifically, a social norm confers upon individuals the standing to act in the name of

enforcing that norm, or – to adapt a term of Stephen Darwall's – social norms grant

---

[2] Bloom and Jordan 2018.

[3] For a thorough discussion of the different ways online shaming implicates concerns of

proportionality, see Billingham and Parr 2020b, pp. 378-383.

individuals "representative authority."[4] I will argue that we have good theoretical reasons to believe that increasing the numbers of people who can become aware of and potentially sanction a given norm violation will lead to problems that arise from those individuals acting in the name of enforcing a norm. Unsurprisingly, among these problems is disproportional treatment of norm violators. But other pathologies can arise as well: increased risk of error in enforcement and alienation. Once we see how problems of scale may arise in relation to the enforcement of social norms, it becomes unsurprising public shaming on the internet generates problems. The internet presents a massively scaled context for social interaction, providing individuals the opportunity to sanction far more people than they could before. Similarly, the internet also renders norm violators liable to sanction by far more possible individuals than before.

The upshot of this argument is *not* that we should shrink the scale of morally valuable social norms so that they apply to a smaller group of people (e.g. subjecting fewer people to norms like those that prohibit sexual assault would be a mistake). Instead, my point is that scaling up social norms in ways that avoid these problems requires careful attention to the authority to enforce these norms. In particular, we need ways of limiting the scale of who actually enforces the norms in question. In the later sections of this paper, I will suggest two avenues of achieving this end. The first

---

[4] Darwall 2013, pp. 23-24.

approach focuses on using second-order norms to limit illegitimate public shaming – roughly, we should shame the wrongful shamers. The second approach takes its cue from the late 18th and 19th century, the last time public shaming was a major topic of public concern. This more radical approach suggests that public shaming in mass contexts (like the internet) is almost always an illegitimate form of norm enforcement.[5] Ultimately, I will provide some reasons to believe the second, more radical approach is preferable.

Prior to beginning my argument, it is worth pausing to clarify the nature of my main claim. I am providing an elaborated empirical theory as to why we should expect increasing scale to vitiate the legitimacy of norm enforcement. This argument depends on various empirical assumptions, which I will make explicit as the argument develops and make a case for their prima facie acceptability. To this extent, the argument is open to falsification on empirical grounds. Just because we should expect scale to create certain problems based on the arguments here does not mean it will. However, I take this to be a virtue of the analysis, not a vice, as it shifts our attention away from handwringing about whether or not any particular case of public shaming is wrong

---

[5] I say "almost always" to preserve the possibility that perhaps there are situations where mass public shaming is called for and does not run into the sorts of problems noted below. Specifically, mass-scale shaming may be an appropriate way to respond to mass-scale wrongs by public figures.

towards thinking through what sorts of contexts are likely to lead to problems.[6] Even if

my proposed analysis and prescription fails to hold up after more rigorous empirical

analysis, thinking of public shaming as a social phenomenon as opposed to an

individual one is more likely to lead to progress, or so I hope to show.

I. The Limits of Proportionality

       Let us begin with an example of public shaming to get an idea of what public

shaming is and help motivate the intuition behind the concern with proportionality:

> *Calling the Police:* A white woman calls the police to report a Black
>
> man in a public space as "acting suspiciously." The man had every
>
> right to be in the public space. The incident goes viral and the
>
> woman is widely condemned on social media as a racist. An
>
> alarming chorus of individuals reach out to her employer on social
>
> media and pressure them to fire her as punishment for her
>
> wrongdoing. She ends up losing her job, and ultimately withdraws
>
> from much of public life as she finds herself constantly subject to
>
> public expressions of disapproval as "that" woman.

---

[6] In this way, I am taking what Jacob Barrett calls a "problem solving" approach to

moral and political theory: Trying to understand the causal structure of an apparent wrong and

proposing a solution to that wrong based on that understanding. Barrett 2020, p. 103. See also

Wiens 2012.

Based on this case, I assume in this paper that public shaming is a social practice constituted by the following features. First, public shaming involves public expressions of moral disapproval of the wrongdoer. Put another way, public shaming requires an audience. Second, these expressions of moral disapproval involve either an implicit or explicit claim that the wrongdoer is someone this audience should not associate with to some degree.[7] I say dissociate "to some degree" to point out that dissociation can range from the mild (temporary emotional distance from the norm violator) to the severe (permanent ostracism). Third, the audience dissociates to some disagree with the wrongdoer on the basis of observing these public expressions of moral disapproval. On this view, public shaming requires some degree of uptake – people can attempt and fail to publicly shame others. To be clear, this is not a full account of public shaming, but a working account to fix ideas and initiate our discussion.

A very common response to many cases of public shaming like *Calling the Police* is, "She may have done something wrong, but she didn't deserve all that!" The implicit idea is that, even if the target of shaming did something wrong and deserves some sanction, the sanction the wrongdoer received was out of proportion with the wrongdoing. In *Calling the Police*, perhaps the woman deserved *some* berating, but getting her employer to fire her? That is too much!

---

[7] See also Aitchison and Meckled-Garcia 2021, pp. 9-10.

On this view, the primary problem of public shaming is that it is disproportional. In the case of criminal punishment, it is widely accepted that there is a proportionality constraint on legitimate punishment, or that punishments must not be overly severe relative to the crime. While there is dispute surrounding what exactly a proportionality constraint entails, that need not cause problems here.[8] Difficulties with a proportionality principle often revolve around how proportionality underdetermines the appropriate punishment. Even granting this problem, agreement on particular cases of *dis*proportional punishment remains possible. Thus, proportionality can act as a negative test for evaluating the permissibility of enforcement even if it is insufficient as a positive one.

Thinking of public shaming as tending towards disproportional treatment clearly captures part of what goes wrong in many of these cases. Just because of its mass nature, public shaming tends towards overpunishment (a point I will come back to later). However, proportionality keeps our analysis at the surface, and, because of this, obscures the underlying problem of public shaming. This is because proportionality concerns attach to modes of enforcement in general. Because a proportionality-based approach focuses on whether or not a given act of enforcement is too much, proportionality gives minimal insight into the mechanisms of public shaming. In *Calling the Police,* we could keep constant the consequences that befall the woman (she suffers

---

[8] Bedau and Kelly 2017, Section 3.

intense moral disapproval by others and loses her job as a result of this moral disapproval) while changing the various social pathways that lead to those consequences. In such a scenario, proportionality would provide the same analysis as in the original case. For example, suppose it is the case that this outcome occurs via *private* shaming by someone close to the target as opposed to public shaming. Proportionality appears to have nothing to say about this difference.[9]

Because proportionality is blind to underlying cause of overpunishment, it is difficult to see the structure of public shaming as a social phenomenon when viewing it through this lens. This matters because, if we are interested in understanding the phenomenon of public shaming in a way that is potentially actionable rather than simply lamenting yet another unjust feature of the world, we need a deeper insight into public shaming beyond the claim that is wrong because it violates a proportionality constraint.

----

[9] One might object that you cannot change from public to private without changing the consequences. Part of what makes public shaming so awful is its public nature. But I can grant this and still get my point: Proportionality recognizes a difference in *degree* but not a difference in *kind* between the two cases. While private and public shaming are cousins, they raise different types of concerns. Understanding and dealing with, say, cyber bulling (which could be in some cases interpreted as a kind of private shaming) requires a different set of tools than dealing with public shaming.

II. Public Shaming, Social Norms and Problems of Scale

To reorient our thinking on this issue, I will argue that public shaming is most fruitfully seen as a form of social norm enforcement. [10] In *Calling the Police,* for example, we can interpret public shaming as a way of enforcing a norm against subjecting people to police suspicion on the grounds of race. By social norm, I mean a general rule accepted by a given population. As Geoffrey Brennan and his collaborators note, this involves two features.[11] The first feature is that a social norm is, as the term suggests, *normative* in nature – the existence of a norm purports to supply a reason for action. The second feature is that social norms are social facts. A social norm exists to the extent that enough of a given population accepts the norm and/or believe that others accept the norm.[12]

It is worth briefly describing how informal social norms differ from formal legal norms.[13] Social norms are informal in at least two respects: in regard to *source* and in regard to *enforcement*.

---

[10] For similar, yet distinct, readings of public shaming, see Billingham and Parr 2020a and Aitchison and Meckled-Garcia 2021.

[11] Brennan et al 2013, pp. 3-4.

[12] There is some dispute regarding the distinction between social norms and moral norms. Foot 1972; Southwood 2011; cf. Heath 2017. This is not important for my purposes here, and so I use social norm in a broad, inclusive sense.

[13] See also Brennan et al 2013, pp. 40-56.

To use H.L.A. Hart's terms, social norms involve primary rules, or rules that apply directly to people.[14] Formal norms, in contrast, involve a rich set of secondary rules in addition to primary rules. Secondary rules, simply put, are rules about rules. Who makes the rules? How do we know what the rules are? And so on. I will set this distinction aside to focus on the more central aspect of social norm's informality for this paper's subject: enforcement.

Social norms are informal in enforcement in at least two senses. First, the enforcement of norms is informal in the sense that what counts as appropriate punishment is not necessarily codified. In particularly well-defined norm groups, like a mafia, violating norms come with well-known punishments. However, for the most part, there tends to be less clarity surrounding what the appropriate response to cutting in line is. This is not to say there are not clear cases of inappropriate norm enforcement. It is obviously inappropriate to explode at the dinner table over improper use of a soupspoon. But the fact that some cases of enforcing a norm are obviously inappropriate does not eliminate the lack of clarity in norms surrounding the enforcement of most norms. This discussion also helps render clear that social sanctions and enforcement can take a variety of forms not limited to public shaming. The mafia case suggests a particularly dramatic form of enforcement, but there are more mild and commonly accepted forms of social sanction.

---

[14] Hart 2012, p. 117.

Second, when one looks to the law, there tend to be demarcated agents responsible for enforcing the law. Vigilantism is not permitted. For example, while the state can punish a murderer by throwing him in jail, I cannot throw a murderer in my basement. In contrast, any individual who witnesses someone jumping the line has standing to call that person out. Perhaps the people wronged by the line-cutter have more standing to do so, but there is nothing obviously wrong with an on-looker sharing some strong words with line-cutter. As I put it in the introduction, any person in a group that shares a norm can – all things being equal – represent the community in enforcing the norm against a given violator.[15] This informality in the enforcement of social norms is where problems arise.

To say that *anyone* may act in the name of enforcing a social norm does not mean that *everyone* may act in the name of enforcing that norm.[16] Properly enforcing a norm requires meeting some success conditions: presumably one of which is that the norm has not already been successfully enforced by others. Similarly, you cannot enforce a norm against someone who has not violated the norm. And so on. People claiming to enforce the norm in such contexts are failing to properly enforce the norm. They are simply inflicting social sanctions on a person illegitimately.

---

[15] Darwall 2013, pp. 23-24.

[16] I thank Bob Goodin for suggesting this way of thinking about representative authority.

My claim is that increasing the scale of community can exacerbate problems of social norm enforcement. More specifically, increasing the number of individuals who can act in the name of enforcing a norm may lead to certain pathologies of enforcement. I will describe three such problems here:

Disproportional Treatment: Increasing scale tends towards disproportional treatment of wrongdoers.

Increased Risk of Error: Increasing scale introduces increased risk of punishing the innocent.

Alienation: Increasing scale introduces the likelihood of being subject to enforcement that one reasonably disagrees with.

To say that these are problems of scale is not to say that these problems do not arise in smaller-scale scenarios. We could imagine all these problems within a community of two. Instead, my intention is two-fold: First, to highlight how scale exacerbates and can introduce these problems in contexts where before they did not exist. Second, to suggest how attention to scale reveals limitations towards certain proposals in response to public shaming. Specifically, proposals that individuals simply refrain from engaging in illegitimate social sanctioning out of the good of their heart fail to appreciate how the problem of public shaming comes about from the social context, not necessarily the character of individuals who engage in such behavior. To solve the problem of public shaming requires attention to social dynamics.

Over the next three sections, I will use various hypothetical cases (loosely based on real-life cases) of online shaming to explain these forms of dysfunctional norm

enforcement and the role scale plays in them. With diagnosis in hand, I will turn to treatment for the remainder of the paper.

A. Disproportional Treatment

As noted, Bloom and Jordan focus on disproportional treatment in their treatment of public shaming. However, they are not interested in individual acts of disproportional treatment. Instead, they are interested in the aggregate effects of individual acts. To illustrate this, Bloom and Jordan use Derek Parfit's case of the harmless torturers.[17] Imagine that a thousand individuals push a button that contributes to a significant harm to some person. However, no one person's pushing a button generates a perceptible difference in the tortured person's state. It is only together that these thousand torturers produce the harm. In the case of public shaming, Bloom and Jordan suggest, it is through collective behavior that we get troubling consequences. This is one way that scale introduces problems into the enforcement of social norms. If all the people who can claim to act in the name of enforcing a social norm actually act on that claim, this will end up in illegitimate enforcement in a large enough group. That being said, scale does not just introduce problems in the form of aggregation.

Because of the possibility of such pile-ons, would-be shamers have to coordinate with others if they are to avoid disproportionate norm enforcement. A problem facing any individual deciding whether to sanction a norm violation is how to know whether

---

[17] Bloom and Jordan 2018; Parfit 1984, p. 80.

someone else has already adequately sanctioned a wrongdoer. Unless each person is able to coordinate with others, this may lead to situations where a norm violator is sanctioned multiple times for the same wrong. In this way, a failure to coordinate can lead to disproportional treatment.

People can resolve these coordination problems relatively easily on small scales. Imagine a somewhat idealized account of *Calling the Police* where the wrongdoing occurs without being picked up on the internet. Nonetheless, a number of people physically witness the incident. One of these people chastises the woman for her wrongdoing, and explains why what she did was wrong. The woman apologizes for her wrongdoing and commits to doing better in the future (perhaps she also donates to a non-profit that combats racial inequality in the justice system as penitence). The other people who witnessed her action also witness her being chastised and feel like she has paid an appropriate price. The coordination problem is solved in this case by a small-scale: Because each person can literally *see* the sanctioning and its results, coordinating is very easy. Were someone to double-down on shaming, others in the group could appropriately respond as they observe that as well.

Large scales exacerbate the coordination problem. As scale increases, it becomes more difficult to track the behavior of others and its consequences. For example, when you expand our toy case from a small group that can all see each other simultaneously to a city where you cannot see all simultaneously, it becomes difficult for each person to keep track of (i) if anyone has already sanctioned the norm violator, (ii) if the sanction was sufficient, (iii) if any invalid sanctioning has occurred, (iv) if sufficient sanctioning

of any invalid sanctioning has occurred, and so on. All of these steps require coordination between individuals.

This might appear less worrisome in the case of public shaming on the internet, as there are clear records of behavior. If considering shaming someone, you can just see who and how everyone else has shamed that person by looking at the relevant trends online. In this way, the internet appears different than a big city. However, I believe this feature of online shaming is less advantageous than it initially appears. First, the nature of the internet is such that it is difficult to gauge the consequences of the behavior of others even if we are able to observe their shaming behavior. For example, when one sees a thread of people shaming someone on the internet, it is difficult to infer from that with any degree of certainty what that actually means for the norm violator. Is she sufficiently shamed at 10 comments? 100? 1,000? 10,000? At a certain point, the differences in magnitude become meaningless to any individual considering engaging in shaming behavior, as the marginal contribution of one comment is negligible. Second, and relatedly, scale makes it difficult for individuals to stop other shamers from piling on in a way that is not true in a small-scale context. For example, assume that the appropriate amount of shame a person should receive is 1,000 comments. Any comment beyond that counts as invalid enforcement of the norm. Suppose there are 1,000 additional people who pile on after the threshold of appropriate punishment has been met. Do you go after each and every one of the inappropriate shamers? Even if you do, there is a chance that a few could slip by unnoticed. What do you do if shaming of the shamers itself has spilled over into the territory of illegitimate enforcement? And so on.

Finally, the ease of observing public shaming on the internet may be overstated because of the existence of multiple platforms and media for public shaming. Coordination across platforms could be a problem.

For these reasons and more, increased scale introduces a variety of pathways to disproportional treatment of norm violators.

B. Increased Risk of Error

A second pathology of public shaming that can be exacerbated by scale is *increased risk of error.* Scale makes it more likely for a variety of mistakes to be made about norm violations. To see this, consider the following cases:

> *Baseball:* A baseman at a professional baseball game tosses a baseball towards a young child. An adult grabs the baseball and appears to celebrate in front of the child. A video of this interaction trends on social media. Many publicly denounce the adult on social media, calling him a monster. However, it turns out that adult in question had already caught a game ball for the child, and was going to give the ball in question to a different child who had not yet gotten a game ball.
>
> *Name Sharing*: John Smith commits some heinous act and deserves all the shame heaped upon him through social media. However, many shamers end up targeting various *other* John Smiths. Many of

these John Smiths receive messages and some of them even find

that their employers have been contacted.

These cases each involve a different kind of mistake. *Baseball* involves a mistake regarding what actually happened, while *Name Sharing* involves a case of mistaken identity. You could imagine other similar epistemic errors being made in the course of people claiming to enforce a norm, thereby vitiating claims to enforcement. While epistemic mistakes are commonplace, we have good reason to expect that increasing scale will lead to increased risk of error when particular conditions are met.

The first step involves the claim that increasing the number of potential enforcers increases the probability of individual mistakes in enforcement. Because humans are fallible creatures, there is always some probability that an individual considering engaging in norm enforcement will make a mistake. When scales are small, the number of individuals who make this sort of mistake will undoubtedly be small. To the extent illegitimate norm enforcement arises due to these kinds of idiosyncratic errors, it will be only of concern to particularly thin-skinned individuals. However, when the number of individuals increases, so does the likelihood of idiosyncratic mistakes. These can add up and wear on the target of this shaming, even if the target of such mistakes knows (and most reasonable people know) that he or she violated no norms. One stray comment is annoying, but a hundred or a thousand can overwhelm.

Second, people often do not enforce norms in a vacuum. The concern regarding idiosyncratic mistakes assumes that would-be shamers make their decisions to shame entirely independent of what others decide to do. However, for many people, the

decision to shame also includes looking at what those before us did in the same situation. Oftentimes, this can help individuals economize on information, as seeing what another person does takes advantage of the information that person has without having access to that information. [18] Due to this imitative feature of human decision-making, we have good reason to believe that individual decisions to shame will correlate with how other individuals previously decided. The public nature of public shaming makes it quite likely that individuals will be looking around to get cues as to what to do. If you have shamed a person, I might imitate you and shame that person. However, copying others does not always lead to legitimate enforcement. If a person makes a mistake in deciding to publicly shame a person, and other people copy that person, this leads to the propagation of illegitimate norm enforcement. Because of this, an idiosyncratic mistake may lead to a chain of mistakes if others simply follow the idiosyncratic shamer. This copy-cat problem can be exacerbated by *prestige bias*, or the tendency to focus on high status individuals in deciding who to imitate.[19] If a celebrity shames a person on social media, this makes it much more likely that others will follow

[18] For a classic discussion of informational cascades, see Bikhchandani et al 1992 and Bikhchandani et al 1998.

[19] Heinrich and Gil-White 2001; Richerson and Boyd 2005, pp. 124-126.

suit than if a random anonymous person does the same. This is true even if the celebrity

has made an error in judgment.[20]

A comparison with gossip, another informal mode of social control distinct from

public shaming, may help illuminate the mechanics of this problem. Gossip, like public

shaming, is a tool of social control. Indeed, many speculate that gossip helps small

groups evade free rider problems by allowing sanctioning of shirkers.[21] Gossip,

however, ends up being counterproductive towards that end as scales increase.[22] Gossip

spreads information about people's behavior in chain-like fashion. This is useful for the

purposes of cooperation to the extent that gossip tracks a person's past behavior. The

fear of being spoken ill of can motivate people to pull their own weight. But, as we all

know, gossip is a noisy process. Embellishment and manipulation, both intentional and

unintentional, accompany the story each time it is told. To put it more technically,

gossip is a low fidelity mode of transmitting information about reputation. Thus, scale

----

[20] As an illustrative example, consider when the film director Spike Lee tweeted out the

home address of George Zimmerman. However, Spike Lee had the wrong George Zimmerman.

Instead of tweeting out the address of the George Zimmerman who shot and killed Trayvon

Martin, he tweeted out the address of an innocent third party – who then proceeded to receive a

high degree of negative attention.

[21] Abramitzky 2018, pp. 101-102. To be clear, the story of gossip is much more complex

than this sort of functional analysis suggests. See Gambetta 1994.

[22] See also Gaus 2011, pp. 90-96.

limits the effectiveness of gossip to the extent that, the larger the group, the more error

is introduced into the gossiping process. Like a game of telephone, information gets

more garbled the more it is shared. The more error, the less useful gossip is as a mode

of tracking the behavior of others.

Public shaming, unlike gossip, occurs in public. Because of this, public shaming

may be more robust than gossip against error (third-parties can step in to potentially

correct unlike in gossip). However, if I am correct about the partial dependence of the

decision to shame on the behavior of others, this robustness is not infinite. If a particular

cascade of erroneous shaming begins, it is quite difficult to stop, or at least stop before

the damage has already been done. To the extent that larger scales make it more likely a

person is going to make a mistake that others may follow (or, enough others may follow

before the mistake is contained), larger scales push the probability of error in informal

enforcement mechanisms like gossip and public shaming closer to one.

C. Alienation

I have discussed public shaming as form of social norm enforcement. Such

enforcement can serve many moral functions. For example, public shaming might

generate collective awareness of wrongdoing.[23] Another possible use of public shaming

may be to assure conditional norm-followers by deterring those who do not accept a

given norm, as well as help produce the behavioral and normative expectations that

---

[23] Thomason 2018, pp. 181-182; Billingham and Parr 2020b, pp. 1-2, 6-7.

establish norms in the first place.[24] I do not dispute that these are potentially valuable

functions of shaming. In this section, I only want draw attention to a certain kind of

alienation, and how that alienation is a predictable outcome of increasing scales.

Consider cases of norm enforcement when there is reasonable disagreement

regarding whether or not a norm violation has occurred:

> *Offensive Joke*: An extremely distant acquaintance of yours on social
>
> media makes a joke that you think is offensive. You post a public
>
> response telling your acquaintance that the joke is inappropriate,
>
> and that they should delete the joke and apologize. They respond
>
> saying that they see nothing wrong with the joke and resent your
>
> attempt to shame them. They block you on social media.

One valuable function of social norm enforcement is to help people live up to the

ideals they themselves hold. On this view, shaming should try and induce shame.[25] This

is not what happened in *Offensive Joke*. Rather, the target of shame felt no shame

whatsoever, and resented your attempt to shame them. They felt alienated by your

actions as they did not take themselves to violate the norm in question: A norm against

---

[24] Rasmussen and Yaouzis 2020.

[25] On the link between shame as an emotion and perceived deficiency, see Nussbaum

2005, pp. 269-270.

offense. I take this alienation to be a predictable result of the particular way that the internet has scaled up social norm enforcement.

To begin, let me draw attention to the idea of *reasonable* disagreement. I do not mean here the Rawlsian notion. Instead, I wish to describe how many norms contain a conventional element to them.[26] By this, I mean that many norms that people accept on an abstract enough level need specification, and there is no obvious way of showing that any one specification of the norm in question is the definitive one. Take the notion of *offense*. It seems relatively uncontroversial to say that most accept a social norm against causing needless offense. However, as Sarah Buss argues, communicating respect for others (and, by implication, offense to others) is mediated through etiquette.[27] What counts as a token of the abstract type of "offense" depends in large part on local conventions. A thumbs up may mean "OK!" around these parts, but in others it is a sign of disrespect. There is no a priori reason to prefer seeing a thumbs up as a sign of approval to seeing it as a sign of disrespect – we ultimately need a convention. To be clear, I am not saying what counts as disrespectful is *merely* conventional. What side of the road we drive on is *merely* conventional. This is not true of norms of respect and offense. It makes a great deal of difference what norms we land

---

[26] Heath 2017, p. 284. To say that norms have a conventional element is not to say that norms are themselves conventions, see Bicchieri 2006, pp. 8-41.

[27] Buss 1999.

on. Nonetheless, there is a conventional *element* to such norms. Even if we have independent reasons to avoid norms of respect rooted in unjust hierarchical social structures, such as those found in a culture of chivalry, egalitarian norms of respect still require specification. We still need to know what observable behaviors count as violating such a norm, and this will to some extent be conventional. For example, does a particular gesture or word or phrase denote the relevant disrespect? In all contexts or only some? There are no strictly a priori answers to these questions – how we answer depends in part on convention even if their significance is not merely conventional. What counts as an offensive joke will be tinged by local color, and this may generate alienating norm enforcement. An implication of this view is that norms that rely heavily on conventionalization, such as those against offense, are more likely to give rise to problems of alienation on the internet than norms that seem to rely less on convention, like those against murder.

This discussion has not explained how scale matters yet. The claim, so far, could be summarized as follows: Some norms are partially indeterminate. As a result, people come to adopt different conventions to specify what counts as a violation of the norm. However, this means that, when people enforce norms, they sometimes may enforce that norm based on a convention that is not shared by the target of enforcement. The issue of alienation so far appears to be a result of diversity, not scale.[28] It is true that

---

[28] I thank an anonymous reviewer for drawing my attention to this issue.

scale by itself does not create alienation. However, scaling up creates alienation when done in particular ways.

Increasing the number of people in a community, as the internet does in a massive way, brings together more and more subgroups of people together who can evaluate the behavior of each other. Further, each subgroup may their own conventions in filling out various widely shared norms. When you group people together who do not necessarily share the same conventions specifying some abstract norm, this predictably leads to disagreement about whether some behavior constitutes a violation of the norm in question. Further, such disagreement is *reasonable* in the sense that, by the lights of the respective conventions used by the disagreeing parties, each person has a reason to think they are in the right. Importantly, it is not scale by itself that creates problems here, but rather how scaling up in the particular way the internet does that leads to increasing diversity among would-be shamers. Thus, it is in principle possible to scale up norm enforcement and avoid alienation – so long as scaling up preserves homogeneity. However, realistically speaking, scaling up norm enforcement will involve incorporating different groups together as opposed to enlarging a pre-existing group while simultaneously preserving shared conventions and limiting deviation.

It is important make a few clarifications of the point here to avoid potential misunderstandings.

First, I am not arguing that large mass groups could not have shared norms against offensive jokes. This is obviously false. Indeed, I suspect that there *are* some cases of nearly universally shared norms within our society about offense (consider the

24

use of at least some racial slurs). The problem of alienation occur at the margins of offense – which is where at least some cases of online shaming reside. The margins are where there is likely to be variation among particular social groups in the conventions surrounding offense, giving rise to difficulties across groups arising from differences in norms.

Second, not all shaming of offensive jokes fall within the category of alienation as described in this section. Take the case of unquestionably sexist or racist jokes. Such jokes function to remind people that certain groups are not worthy of equal respect, and, as Merrie Bergmann puts it, "adds insult to the injury" of these groups' oppression.[29] In such cases, a response to shame as found in *Offensive Joke* is not a result of individuals reasonably disagreeing that their behavior counts as a violation of the norm, but instead a case of post hoc rationalization. The person realizes they did something wrong, and ex ante would acknowledge as such. Nonetheless, they deny wrongdoing once called out on their bad behavior. This is not a case of alienation, but rather attempted norm evasion. But the existence of norm evasion does not imply that alienation of the sort I am interested in here does not occur.[30]

---

[29] Bergmann 1986, p. 78.

[30] Shaming of such cases might nonetheless run into the prior noted problems of disproportionality and error.

III. Solutions

A. The Impotence of Conscience-Based Solutions

I have argued that appropriate diagnosis of the ills of public shaming relates to problems of scale: The more people who can possibly claim authority to enforce a social norm, the more likely various pathologies are to enter that vitiate that claim. In contrast to the traditional proportionality-based lens, this broadens our view to seeing a number of problems latent in much public shaming. Perhaps more importantly, this helps understand the source of wrongdoing and can aide us in responding to the problem of public shaming. I will provide two reasons in support of this claim. In this section, I will argue a scale-based approach to public shaming can help us avoid solutions that are unlikely to be successful. Specifically, I will argue that appeals to the individual conscience of would-be illegitimate shamers are unlikely to solve any of the problems noted before. In the next section, I will argue that a scale-based approach can help direct us towards promising treatments by focusing our attention to representative authority.

One somewhat intuitive approach to solving the problem of public shaming is to simply ask people to do the right thing and stop engaging in illegitimate norm enforcement. This approach appeals to the *conscience* of would-be shamers. This is the tack Bloom and Jordan take. At the end of their brief piece, Bloom and Jordan suggest "[I]f we want to be decent people, we should try."[31] On this view, the problem of public

---

[31] Bloom and Jordan 2018.

shaming is solved by encouraging people to be better by thinking about whether or not they are about to engage in illegitimate shaming. If so, they should refrain.

There is nothing wrong with appeals to conscience as such. The problem with such appeals is that they are likely to be impotent in this case. Perhaps a few less people will engage in illegitimate norm enforcement on the margin, but such calls for decency are highly unlikely to solve the problem because they direct our attention to the fact that much public shaming is *wrong* without an understanding of the underlying problem. That is, these appeals focus on symptom as opposed to cause. Seeing the problem of public shaming as a problem of scale, as I have argued we should do, provides two reasons to believe that appeals to conscience are unlikely to be successful.

First, when we make an appeal to conscience we essentially ask a particular individual to aim at doing the right thing. But this assumes that the problems of public shaming arise from individuals not aiming at doing the right thing. While this might be to some degree true (more on this in a bit), a scale-based approach provides reasons to believe that even individuals who are making good faith efforts to engage in legitimate social norm enforcement will run into these problems. Consider the problems of coordination noted in the above discussion of disproportional punishment. The issue with coordinating punishment is *not* that individual shamers are not trying to do the right thing. The issue is that uncoordinated individuals will lead to problems, *even if* each individual is trying to do the right thing and avoid illegitimate public shaming. Similarly, appeals to conscience do not address the epistemic problems exacerbated by

the scale of much public shaming today. A well-meaning person can act on the false information she sees spreading throughout her group.

Second, once we see that people have other reasons for engaging in public shaming, it becomes less clear that an appeal to conscience will solve the problem. As psychologists, Bloom and Jordan are quite aware that part of what motivates public shamers is a "desire to show off our goodness to others."[32] To use Justin Tosi and Brandon Warmke's term, public shaming can act as a form of *moral grandstanding*.[33] This sort of uncouth motivation might seem ripe for an appeal to conscience to work on, but a closer investigation into the nature of grandstanding suggests a thornier issue. Tosi and Warmke point out that the "recognition desire," or desire to be recognized as morally respectable, that drives moral grandstanding is comparative in nature.[34] In the case of public shaming, people who do not publicly condemn some wrongdoer appear *less* morally respectable if there are others who *do* publicly condemn the wrongdoer. It is this comparative aspect of public shaming as grandstanding that creates problems for the appeal to conscience.

If people's desire to engage in public shaming is partially driven by a concern with appearances, and such appearances are relative in nature, this suggests that

---

[32] Bloom and Jordan 2018.

[33] Tosi and Warmke 2016.

[34] Tosi and Warmke 2016, p. 201.

refraining from public shaming has the structure of a collective action problem. We

might all be better off with less public shaming. That being said, each of us has an

interest in engaging in public shaming while others refrain (I get to look better than

those who remain silent in the face of wrongdoing!). Further, each of us has an interest

in not being silent when others shame, for fear of being seen as morally worse than

others. But this just means that we create a kind of moral arms race that is difficult for

individuals to unilaterally escape from. And that is the problem for the appeal to

conscience. Part of what makes collective action problems so thorny is that the

individual perspective tends to be insufficient to solving them.[35] Merely appealing to

individual conscience is unlikely to change behavior in such cases. Further, as Hume

noted long ago, scale makes it far more difficult to keep in check the dynamics of such

problems.[36] Something more needs to be done.

---

[35] Sankaran 2020.

[36] "Two neighbours may agree to drain a meadow, which they possess in common;

because 'tis easy for them to know each others mind; and each must perceive, that the

immediate consequence of his failing in his part, is the abandoning the whole project. But 'tis

very difficult, and indeed impossible, that a thousand persons shou'd agree in any such action;

it being difficult for them to concert so complicated a design, and still more difficult for them to

execute it; while each seeks a pretext to free himself of the trouble and expence, and wou'd lay

the whole burden on others." Hume 1975, p. 528, [3.2.7.8]. See also Olson 1971, pp. 5-52.

B. Limiting the Scale of Enforcement

The failure of appealing to conscience is driven by seeing the problem of public shaming as one of individual behavior and character. Approaching public shaming as a problem of scale directs our attention to the social dynamics behind the problem of public shaming, and makes it more likely to find a solution. If the problem is that the representative authority can lead to pathologies on a large scale, we ought to focus on how to manage that authority. In what space I have left, I will describe two such potential ways of doing so. The more moderate response focuses on limiting representative authority to those cases where people engage in legitimate public shaming. The other, more radical approach opts for a general presumption against public shaming in mass contexts such as the internet. Both have their merits and weaknesses, and I suspect the only proper test of these two responses would be a practical one. That being said, I will provide some theoretical reasons to think the more radical position is more promising.

Both responses begin with the problem noted above facing an appeal to conscience: the interaction of scale and moral grandstanding that generates a moral arms race dynamic at the heart of much public shaming. If people are engaging in public shaming as a form of moral grandstanding, the natural solution is to change the incentives. Specifically, if people engage in public shaming to gain the warm glow of moral recognition, we should not give such recognition to people when they shame illegitimately.

i. A Moderate Approach

The moderate approach seeks to change the incentives by enacting second-order norms that keep in check illegitimate norm enforcement. As Billingham and Parr suggest, "It is desirable that those who shame others can themselves be held to account."[37] Second-order norms solve the arms race problem as it creates a point at which would-be shamers can no longer one-up others. In fact, trying to claim the mantle of enforcing a social norm when doing so would be illegitimate would not just fail to improve one's moral standing, but would damage such standing. Using a social norm against improper norm enforcement like this acts as a disincentive against illegitimate shaming.

The moderate approach has its virtues: it addresses the social context of public shaming, it focuses on whether or not people act on their representative authority legitimately or not, and, similarly, it creates a space for public shaming to play a role in the enforcement of social norms. However, I have some concerns with such an approach.

My major concern with the moderate approach is that it depends on using shame to combat shame.[38] All the problems of disproportionality, coordination, error, and

---

[37] Billingham and Parr 2020a, p. 15.

[38] Billingham and Parr are certainly aware of this, but I do not believe sufficiently alive to how much this might affect the moderate approach to dealing with public shaming. Billingham and Parr 2020a, p. 15.

alienation may arise at the second-order level, and are especially likely to once we take into account the scaling issues noted above. Take the case of alienation: Imagine a person, Albert, engages in public shaming, but another person, Bertha, calls Albert out for engaging in illegitimate shaming. Suppose Albert disagrees with Bertha that the target of shaming has been sufficiently shamed – how do we deal with this disagreement? Because informal social norms lack secondary rules, there is no authoritative way to adjudicate this dispute. You could imagine it possible for a small group of individuals to arrive at a shared understanding of when second-order norms surrounding the enforcement of norms are violated. I would think here of a group the size of a family or group of friends or a church group or a commune. However, larger scales make it less likely for such a shared understanding to arise. Or, take the case of coordinating in such a way to avoid disproportional punishment. Suppose that on one online platform a person has been sufficiently shamed (I am assuming away the problem of determining "sufficiently shamed"). However, people on *another* online platform are unaware of this, and continue to shame. It is not obvious how the secondary norm coordinates behavior across media, which is a problem the scale of the internet is likely to exacerbate. For these reasons, I am worried about a moderate approach to the problem of public shaming that focuses on shaming the shamers. I am not sure if these problems are decisive for the moderate response to public shaming, but they suggest it may be worth looking elsewhere for a solution to the problem of public shaming.

ii. A Radical Approach

The radical approach takes it cue from the history of public shaming. Today is not the first time public shaming has been a subject of concern, and, because of this, it might be worth looking to past efforts. Famously, Nathaniel Hawthorne's *The Scarlet Letter* renders the public shaming of an adulterer in a 17th-century Puritan town. *The Scarlet Letter* was published in 1850, and his work reflects a growing concern with public shaming at this time. Prior to the 19th century, public shaming was an important form of criminal punishment in America. Such practices included use of the pillory, the stocks, public whipping, and even placing the convict on the gallows with a rope around his neck for hours.[39] The fact that public shaming historically resulted from a criminal proceeding generates complications for comparison with today's online cases. Nonetheless, these punishments depend on social norms latent in the relevant community. Public shaming as criminal punishment does not work if the public will not shame, and so relies on members of the public to engage in specific enforcement acts grounded in social norms. To this extent, making all the relevant caveats, it is not inappropriate to compare the public shaming of earlier times with that found online today.

---

[39] Hirsch 1982, p. 1225.

The historian Peter Stearns documents that public shaming largely disappears by the second half of the 19th century in America.[40] This decline in public shaming did not just occur incidentally without notice or remark, but rather involved public criticism of such practices.[41] In a 1787 polemic against public shaming, Benjamin Rush suggests that, "Ignominy is universally acknowledged to be a punishment worse than death."[42] Of note, and in line with my analysis in this paper, there are reasons to believe that the push against public shaming in the 19th century is in part a result of increasing scales of people interacting. The legal scholar Adam Hirsch documents in Massachusetts that most cities (barring Boston and Salem) were no larger than 2,000 people as late as 1690.[43] By the mid-18th century, nearly half of all towns in Massachusetts grew past 1,000 people with thirty towns containing more than 2,000 people.[44] Under such conditions, my view suggests that it would not be surprising if there were some pressure against social norm enforcement in the form of public shaming – and indeed there was. By the 20th century, the vast majority of states outlaw common forms of public shaming punishments.[45] Delaware stands as the last holdout, outlawing public

---

[40] Stearns 2017, pp. 59-65.

[41] Ronson 2015 p. 54.

[42] As cited in Stearns 2017, p. 62.

[43] Hirsch 1982, p. 1223.

[44] Hirsch 1982, p. 1228.

[45] Stearns 2017, pp. 62-65

whipping in 1952. Delaware's obstinacy in this matter did not go unnoticed. In 1876, the

New York Times ran an editorial criticizing Delaware's failing to give up forms of public

shaming.[46]

The point of this brief dip into the history of public shaming is to suggest that

perhaps the right response is the radical one endorsed the last time public shaming was

a subject of public concern: A call for a strong presumption against public shaming in

mass contexts such as the internet. Obviously, there are important differences between

today and the past. For example, making public shaming illegal as it appears on the

internet today is almost certain to generate more problems than it solves. Thus, the

radical approach should not go that far. But my thought is that public shaming in mass

contexts such as the internet will almost always run into the problems of scale noted

above. Because of this, there should be a strong presumption against public shaming in

mass contexts.

iii. Objections and Clarifications

Being a somewhat radical proposal, this is likely to prompt a number of

questions and objections, so let me provide some clarifications.

It is important to stress that being against public shaming does *not* involve letting

people off the hook for bad behavior. In the past, it was not as though critics of public

shaming like Benjamin Rush thought murderers were not subject to norms against

---

[46] Stearns 2017, p. 62.

murder. Rather, the implication of being against public shaming in mass contexts is looking for alternative modes of norm enforcement that are less likely to give rise to problems. In the 19th century, this involved placing the primary responsibility to punish wrongdoers in the hands of the formal legal system. Today, we might consider alternative modes of sanctioning to public shaming – either informal or formal – that are less likely to run into the problems of scale noted above. Space does not permit a full investigation of such alternatives to public shaming in mass society, but possibilities include mild forms of social exclusion.[47] Of course, social exclusion raises its own set of problems and questions, but it might be a place to start when thinking about alternatives to public shaming. Another possibility is that we might opt for leaving the application of social sanctions to smaller scales, even if we wish the norms themselves to apply on large scales. This dovetails nicely into my next point.

Being against public shaming in mass contexts is not the same as being against public shaming in *all* contexts. When considering the moderate solution above, I noted that second-order norms against illegitimate norm enforcement may be easier to arrive at in smaller contexts. Related to my first point, the idea would be to limit representative authority to enforce a norm to a smaller circle around the norm violator. Again, this does not mean the norm does not apply on a large scale, only that the standing to enforce the norm is more limited than currently presumed. This would have

[47] E.g. Molho et al 2020.

the benefit of allowing people to coordinate more easily as well as come to shared understandings of what sort of second-order norms should govern first-order norm enforcement.

One might wonder how a radical proposal does not run into some of the problems noted above with the moderate proposal, or problems related to shaming the shamers. Suppose the idea of being against public shaming in mass contexts picks up steam. In these conditions, suppose it were also the case that someone or a group of people were to engage in public shaming on the internet despite wide-spread consensus against public shaming – how do we deal with this case other than by responding in kind?

I have a few thoughts about this objection. The first thing to say is that the possible alternatives to mass public shaming noted above could be mobilized against would-be shamers. But, perhaps more importantly, if mass public shaming is widely seen as illegitimate, there is really not much needed to do in response. The target of the shaming might be a little rattled, but they could easily comfort themselves by noting these people are engaging in illegitimate behavior. Similarly, third parties, such as employers, could safely ignore the mass shamers as not properly engaged in enforcing social norms. Indeed, I suspect if mass public shaming is seen as illegitimate, this would do a great deal to limit its occurrence without any need to shame the shamers. This is because, for both those who are trying to do the right and those who simply wish to look like they are doing the right thing the way is clear: mass public shaming is off limits.

IV. Conclusion

      With the rise of social media, public shaming has returned to the fore as both a mode of social response to wrongdoing and as an object of public concern. Most analysis of public shaming seems content to remain at the surface – there is something wrong going here! I have argued that this leaves us without a proper understanding of the phenomenon. In particular, I have argued that the problem of public shaming arises from problems of scale. Once we see this, this not only helps us understand why public shaming on mass scales like the internet gives rise to various pathologies, but also helps think through how to solve the problem of public shaming. I have suggested two possible responses, but my own preference is that we should follow the path of critics of public shaming in the 18th and 19th century: reject public shaming as a mode of dealing with wrongdoing in mass social contexts. In order to be successful, however, this path requires investigating alternative approaches to enforcing social norms. What those might be, I have only suggested. However, if we are to solve the problem of public shaming, I believe it is deeply important to find alternatives in dealing with norm violations.

References

Abramitzky, Ran. 2018. *The Mystery of the Kibbutz.* Princeton, Princeton University Press.

Aitchison, Guy and Saladin Meckled-Garcia. 2021. Against online public shaming:

      Ethical problems with mass social media. *Social Theory and Practice*, 47, 1-31.

Barrett, Jacob. 2020. Social reform in a complex world. *Journal of Ethics and Social Philosophy*, 17, 103-132.

Bedau, Hugo Adam and Erin Kelly. 2017. Punishment. *The Stanford Encyclopedia of Philosophy* (Winter 2017 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/win2017/entries/punishment/>.

Bergmann, Merrie. 1986. How many feminists does it take to make a joke? Sexist humor and what's wrong with it. *Hypatia,* 1, 63-82.

Bicchieri, Cristina. 2006. *The Grammar of Society*. New York: Cambridge University Press.

Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch. 1992. A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of Political Economy,* 100, 992-1026.

Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch. 1998. Learning from the behavior of others: Conformity, fads, and informational Cascades. *Journal of Economic Perspectives*, 12, 151-170.

Billingham, Paul and Tom Parr. 2020a. Enforcing social norms: The morality of public shaming. *European Journal of Philosophy*, 28, 997-1016.

Billingham, Paul and Tom Parr. 2020b. Online public shaming: virtues and vices, *Journal of Social Philosophy,* 51, 371-390.

Bloom, Paul and Matthew Jordan. 2018. Are we all harmless torturers now? *New York Times*, August 9, 2018. https://www.nytimes.com/2018/08/09/opinion/are-we-all-harmless-torturers-now.html. Accessed January 10, 2019.

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

Buss, Sarah. 1999. Appearing respectful: The moral significance of manners. *Ethics,* 109, 795-826.

Darwall, Stephen. 2013. *Morality, Authority & Law: Essays in Second-Personal Ethics*. Oxford: Oxford University Press.

Foot, Philippa. 1972. Morality as a system of hypothetical imperatives. *The Philosophical Review,* 81, 305-316.

Gambetta, Diego. 1994. Godfathers' gossip. *European Journal of Sociology / Archives Européennes de Sociologie*, 35, 199-223.

Gaus, Gerald. 2011. *The Order of Public Reason.* Cambridge: Cambridge University Press.

Hart, H. L. A. 2012. *The Concept of Law*, Third Edition, ed. Leslie Green. Oxford: Oxford University Press.

Heath, Joseph. 2017. Morality, convention and conventional morality. *Philosophical Explorations* 20, 276-293.

Heinrich, Joseph and Francisco J. Gil-White. 2001. The evolution of prestige: Freely conferred deference as a mechanism for enhancing the benefits of cultural Transmission. *Evolution and Human Behavior,* 22, 165-196.

Hirsch, Adam J. 1982. From pillory to penitentiary: The rise of criminal incarceration in early Massachusetts. *Michigan Law Review,* 80, 1179-1269.

Hume, David. 1975. *A Treatise of Human Nature*, 2nd Edition, eds. L.A. Selby-Bigge and P.H. Nidditch. Oxford: Clarendon Press.

Molho, Catherine, Joshua M. Tyber, Paul A. M. Van Lange, & Daniel Balliet. 2020. Direct and indirect punishment of norm violations in daily life. *Nature Communications,* 11, https://doi.org/10.1038/s41467-020-17286-2.

Nussbaum, Martha. 2005. Inscribing the face: Shame, stigma, and punishment. Pp. 259-302 in *Nomos XLVI: Political Exclusion and Domination*, ed. By Melissa S. Williams and Stephen Macedo New York: New York University Press.

Olson, Mancur. 1971. *The Logic of Collective Action.* Cambridge: Harvard University Press.

Parfit, Derek. 1984. *Reasons and Persons*. Oxford: Oxford University Press.

Rasmussen, Katharina Brendt and Nicolas Olsson Yaouzis. 2020. #MeToo, social norms, and sanctions. *Journal of Political Philosophy*, 28, 273-295.

Richerson, Peter J. and Robert Boyd. 2005. *Not by Genes Alone: How Culture Transformed Human Evolution.* Chicago: University of Chicago Press.

Ronson, Jon. 2015. *So You've Been Publicly Shamed*. New York: Riverhead Books.

Sankaran, Kirun. 2020. What's new in the new ideology critique? *Philosophical Studies* 177, 1441-1462.

Southwood, Nicholas. 2011. The moral/conventional distinction. *Mind,* 120, 761-802.

Stearns, Peter. 2017. *Shame: A Brief History.* Urbana: University of Illinois Press.

Thomason, Krista K. 2018. *Naked: The Dark Side of Shame and Moral Life.* New York, Oxford University Press.

Tosi, Justin and Brandon Warmke. 2016. Moral grandstanding. *Philosophy & Public Affairs,* 44, 197-2174.

Wiens, David. 2012. Prescribing institutions without ideal theory. *Journal of Political Philosophy*, 20, 45-70.