

So You Want to Publicly Shame: The Ethics of Public Shaming

James Edgar Lim Tjen Yao

January 2024

A thesis submitted for the degree of Doctor of Philosophy of
The Australian National University

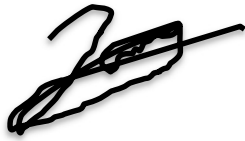
© Copyright by James Edgar Lim 2024

All Rights Reserved

Declaration

This thesis is solely the author's work. No part of this thesis has been submitted for another degree. To my knowledge, any help received in preparing this thesis and all sources have been duly acknowledged.

Chapter IV of this thesis is a modification of my article, "Public Shaming as Moral Self-Defence," forthcoming in *Social Theory and Practice*.

A handwritten signature in black ink, appearing to be 'James Edgar Lim', with a stylized, overlapping 'J' and 'E'.

James Edgar Lim

Acknowledgements

The work in this thesis is the product of many discussions, rounds of feedback, and other informal contributions over the years. I want to thank my primary supervisor Christian Barry, who has spent many hours reading my work, providing valuable feedback, and engaging in productive discussions with me. He has read and provided guidance on each and every chapter of my thesis, in various stages of development, and has certainly left a positive mark on this piece of work. I thank Nic Southwood and Philip Pettit, members of my supervisory panel, who have also read and provided valuable and insightful advice on improving various chapters of this thesis. This project would be poorer if not for their contributions.

Over the years, other people have also read various pieces of my work which have added up to this thesis in one way or another. I'm grateful to Garrett Cullity, Seth Lazar, Jennifer Munt, Yeo Shang Long, Nicholas Carroll, Nicholas Drake, Chris Lernpaß, Alexandra Oprea, and Kate Goodfruit for comments on these pieces of writing. I also thank members of the Australian National University, especially members of the Moral, Social, and Political Theory Work-in-Progress group for their feedback on previous versions of some of the chapters. I gratefully acknowledge the comments provided by three anonymous referees for *Social Theory and Practice* on a version of chapter IV of this thesis. I thank Matthew Liao for the mentoring he provided while co-authoring a paper, not related to this project, with me. And I thank Szymon Bogacz for blazing a trail in the forest of various administrative issues just six months before me throughout this program, thus making it so much easier to navigate them.

Finally, I thank my friends and family who have stood with me and provided invaluable support as I pursued this project. In particular, I thank my parents, Helena Lim and Philip Lim, who have been a pillar of support in my pursuit of this project.

Abstract

Public shaming is a powerful and widespread phenomenon. People use public shaming to punish racist behaviour, sexual misconduct, academic misconduct, dishonesty, inconsiderate public behaviour, violations of public health norms, and so on. But public shaming also inflicts severe harm on targets and their intimate relations, like their families and friends. In recent years, commentators and some philosophers have raised concerns regarding a number of ethical issues, like proportionality, standards of proof, and whether public shaming constitutes a form of objectionable vigilantism. In my thesis, “The Ethics of Public Shaming,” I address some of these issues by developing an account of public shaming, what justifies it, and the rights and duties held by various individuals with respect to public shaming.

Because of the harm imposed by public shaming, human beings have rights against being publicly shamed. This is a key claim I defend in this project. Nevertheless, I also argue that some cases of public shaming can be justified by the contributors’ need to defend certain vital interests. These interests include important constituents of wellbeing, like personal safety and autonomy, and contributors’ interests in maintaining their moral standing within their communities – the subjective perception of their moral value within those communities. Public shaming is sometimes necessary to enforce norms which in turn safeguard these vital interests. As a result, public shaming can be justified, because it can be a form of collective self or other-defence. Drawing upon existing frameworks of justified self-defence, I go on to examine the conditions under which contributions to public shaming can be considered necessary, proportionate, and establish the relevant stakes for agents deliberating under conditions of uncertainty.

Contents

Declaration.....	2
Acknowledgements	3
Abstract	4
Contents.....	5
Chapter I: Public Shaming.....	6
Chapter II: The Right against Public Shaming	24
Chapter III: Justifying Public Shaming – Public Shaming as Norm Enforcement	64
Chapter IV: Justifying Public Shaming – Public Shaming as Moral Self-Defence.....	103
Chapter V: The Moral Responsibility of Individual Contributors	126
Chapter VI: Knowing and Shaming – Public Shaming in Uncertainty.....	146
Notes for Chapter VI.....	184
Bibliography	190

Chapter I: Public Shaming

I. Introduction

As I shall use the term, public shaming roughly refers to the collective public criticism of a target in response to the target's perceived wrongdoing. Public shaming as a topic is ripe ground for philosophical intervention, for it poses a number of puzzles. To illustrate, here is one: On one hand, public shaming seems similar in many respects to paradigmatic forms of punishment like criminal legal sanctions. It typically imposes some very significant harm on its targets; the agents who impose it act intentionally; and those agents often aim at imposing a harmful outcome (although they may not intend every outcome associated with public shaming). It is typically done in response to some perceived wrongdoing, and conveys the disapproval or condemnation of the shamers. These features are also typically found in paradigmatic forms of punishment.¹ On the other hand, it does not seem right to think about public shaming in the same way we treat those paradigmatic forms of punishment. Few people, for instance, would claim that the state should maintain a monopoly on public shaming, even though most of us would agree that the state should have the sole right to incarcerate criminal wrongdoers. Or again, it is commonly held that public shaming is justified or appropriate in cases where criminal punishment is not. For example, in response to a racist joke. Thus, the attention of moral philosophy seems worthwhile. It might explain why our intuitions about public shaming are different from other forms of punishment, or tell us why and how we should revise our views on public shaming.

In addition, public shaming is relevant and important, because of its widespread use in our modern world. In addition to traditional forms of public shaming, like putting someone in the

¹ Radzik, Linda, *The Ethics of Social Punishment* (Cambridge: Cambridge University Press, 2020), 9-10.

stockades, there are newer, far more wide-reaching forms of public shaming. The internet and social media allow thousands of individuals from all over the world to focus their anger and fury at individuals perceived to have violated norms. And the issue has already received some attention from academics and non-academics alike. A number of authors, including philosophers, have theorized the benefits of public shaming, at least within certain contexts.² Yet other philosophers have argued that public shaming should not be used, or that there should be strong presumptions against it.³ This debate is replicated in non-academic commentary, with journalists and other commentators taking sides in favour of public shaming and against it. Public shaming as a topic thus involves a number of socially salient and unresolved issues – issues that philosophical analysis can resolve or at least clarify.

In this thesis, I carry out a systematic analysis of public shaming: why it can be wrong, what justifies it, the duties that individual actors have in cases of public shaming, etc. Roughly in order, here are the broad questions I address in this project:

- 1) What is public shaming?
- 2) Do human beings have rights against being publicly shamed?
- 3) What, if anything, can justify public shaming?
- 4) Given the collective nature of public shaming, what is the relationship between the morality of a case of public shaming and an individual contributor's moral responsibility?

² Billingham, Paul, and Tom Parr, "Enforcing social norms: The morality of public shaming," *European Journal of Philosophy* 28, no. 4 (2020): 997-1016; Braithewaite, John, *Crime, Shame, and Reintegration* (Cambridge: Cambridge University Press, 1989); Jacquet, Jennifer, *Is Shame Necessary? New Uses for an Old Tool* (New York: Pantheon Books, 2015).

³ Meckled-Garcia, Saladin, and Guy Aitchison, "Against Online Public Shaming," *Social Theory and Practice* 47, no 1 (2021):1-31; Frye, Harrison, "The Problem of Public Shaming," *The Journal of Political Philosophy* 30, no. 2 (2022): 188-208.

- 5) Under conditions of uncertainty, what standards of proof should agents hold themselves and each other to if they want to contribute to public shaming?

There are two main goals of this introductory chapter. First, I want to define public shaming, as I intend to use the term. The first goal is the subject of section II of this chapter. The second goal of this chapter is to provide an outline of the rest of this thesis, which will be done in section III.

IIa. The Definition of Public Shaming

As Radzik points out, a definition can help us coordinate our attention on a key set of issues.⁴ My interest in this project is in a distinct set of moral issues and questions that arise from the collective and informal imposition of harm on accused wrongdoers, via discrete individual acts like criticism, in response to perceived wrongdoing. Thus, while I recognize that the phrase “public shaming” has been used to refer to different phenomena by various philosophers, I define public shaming the following way (From this point, any use of the phrase “public shaming” will refer to this definition unless otherwise stated):

Public Shaming: A (a) collective activity in which (b) each contributing agent expresses criticism, blame, resentment, disgust, anger, contempt, or other negative attitudes towards another person (c) in a public forum (d) in response to a perceived violation of some norm.

What is notable about this definition is that it focuses our attention on a certain kind of public shaming – the kind of mass shaming or outrage that we typically see online, though it is not necessarily seen online. The kind of behaviour has also been referred to sometimes as “cancelling”,

⁴ Radzik, Linda, *The Ethics of Social Punishment*, 6.

online shaming, online vigilantism, or outrage culture. It thus captures some paradigmatic examples of public shaming:

Justine Sacco: Justine Sacco, a former PR executive, made the following tweet in December 2013: “Going to Africa. Hope I don’t get AIDS. Just kidding. I’m white!”⁵ By all accounts, this was a part of a long series of off-beat, sarcastic jokes she made to 170 Twitter followers. But within her eleven-hour flight to Africa, thousands of other internet users had somehow caught wind of her tweet. Upon landing, she was bombarded with comments like “No words for that horribly disgusting, racist as fuck tweet from Justine Sacco. I am beyond horrified”, “All I want for Christmas is to see @JustineSacco’s face when her plane lands and she checks her inbox/voicemail”, and “we are about to watch this @JustineSacco bitch get fired. In REAL time. Before she even KNOWS she’s getting fired”.⁶ Popular website BuzzFeed chimed in as well, publishing an article entitled “16 Tweets Justine Sacco Regrets”.⁷ Sacco lost her job soon after, and described the experience as “extremely traumatic”.⁸ Further, she reported alienation among family – an aunt (reportedly) said “This is not what our family stands for. And now, by association, you’ve almost tarnished the family”.⁹ She also reported worries about finding a new job and dating.¹⁰

The Time’s Up & #MeToo Movements: These two movements began in 2017 in response to a series of revelations regarding influential and powerful men in the entertainment industry,

⁵ Ronson, Jon, *So You’ve Been Publicly Shamed* (London: Picador, 2015), chapter 4.

⁶ *Ibid.*

⁷ Smith, Rohan, “Justine Sacco speaks about her experience one year after the tweet that ruined her life,” *News.co.au*, 16 Feb 2016, <https://www.news.com.au/technology/online/social/justine-sacco-speaks-about-her-experience-one-year-after-the-tweet-that-ruined-her-life/news-story/993b0788fdbda58794abb7ee53be04ed>.

⁸ *Ibid.*

⁹ *Ibid.*

¹⁰ *Ibid.*

who had been accused of sexual assault, harassment, and misconduct. The movements quickly evolved to be about sexual misconduct in other industries as well. One aspect of these movements is that they encouraged victims of sexual harassment and assault to speak up about their experiences, sometimes on public platforms like social media.¹¹ As a result, many prominent figures were subject to intense public criticism and scrutiny, resulting in them losing their positions, and in some cases, criminal charges. Those accused of sexual misconduct and subsequently lost roles include producer Harvey Weinstein, Fox News chairman Roger Ailes, commentator Bill O'Reilly, comedian Louis CK, Journalist Matt Lauer, US Senator Al Franken.¹²

Amy Cooper, the "Central Park Lady": Amy Cooper is a white woman who threatened to call the cops on Christian Cooper, a black man who was watching birds in central park (he had asked her to leash her dog), and falsely claimed that he was threatening her life.¹³ Christian Cooper videoed the event and posted it online. Predictably, she was dubbed the "Central Park Lady", and was heavily punished for her actions. According to the New York Times, "internet sleuths" dug into her personal life for more damning information.¹⁴ She lost her job, her dog (she appears to treat the dog roughly in Christian Cooper's video), and was charged in court for lodging a false report.¹⁵

¹¹ Langone, Alix, "#MeToo and Time's Up Founders Explain the Difference Between the 2 Movements — And How They're Alike," *Time*, 22 Mar 2018, <https://time.com/5189945/whats-the-difference-between-the-metoo-and-times-up-movements/>.

¹² Carlsen, Audrey, Maya Salam, Claire Cain Miller, Denise Lu, Ash Ngu, Jugal K. Patel, and Zach Wichter, "#MeToo Brought Down 201 Powerful Men. Nearly Half of Their Replacements Are Women," *New York Times*, 29 Oct 2019, <https://www.nytimes.com/interactive/2018/10/23/us/metoo-replacements.html?nl=top-stories&nid=72995439&ref=cta>.

¹³ Walsh, Joan, "Birding While Black: Just the Latest Bad Reason for White People to Call Police," *The Nation*, 26 May 2020. <https://www.thenation.com/article/society/amy-cooper-birding-police/>.

¹⁴ Jashinsky, Emily, "Amy Cooper Doesn't Deserve Sympathy, But Social Media Shaming Is An Unhealthy Norm," *The Federalist*, 27 May 2020. <https://thefederalist.com/2020/05/27/amy-cooper-doesnt-deserve-sympathy-but-social-media-shaming-is-an-unhealthy-norm/>.

¹⁵ Cooper, Christian, *Better Living Through Birding* (New York: Random House, 2023).

Here are some key terms I will use throughout this project. A *case* of public shaming refers to the aggregate set of individual acts of criticism directed towards a person/group in response to an act of wrongdoing, or series of wrongdoings. A *contribution* to public shaming refers to an individual act which is a part of a case of public shaming. The *practice* of public shaming refers to the way in which public shaming is usually carried out in society, the expectations regarding public shaming, and the social norms regulating how public shaming ought to be carried out.

The term “public shaming” is also commonly used to describe a number of other activities not captured by my definition. These include:

- a) The use of public humiliation by the state or other institutions.¹⁶
- b) A single agent, or small group of agents, humiliating someone in public (imagine a bully making fun of someone for being a nerd in the school canteen).

In addition, some philosophers might believe that “public shaming” should refer to phenomena more specific than collective criticism in public (although collective criticism in public could include the more specific phenomena these philosophers have in mind). Such phenomena might be:

- c) Trying to make someone feel shame in public.
- d) The use of public criticism to cause stigmatization.¹⁷

¹⁶ See Braithwaite, John, *Crime, shame and reintegration* (New York: Cambridge University Press, 1989); Brooks, Thom, “Shame On You, Shame On Me? Nussbaum on Shame Punishment,” *Journal of Applied Philosophy* Vol. 25, No. 4 (2008), 322-334; Kahan, Dan, “Shame, Stigma, and Crime: Evaluating the Efficacy of Shaming Sanctions in Criminal Law,” *Harvard Law Review* Vol. 116, No. 7 (2003), pp. 2186-2207; and Nussbaum, Martha, *Hiding from Humanity: Disgust, Shame, and the Law* (Princeton: Princeton University Press, 2004).

¹⁷ See Aitchison, Guy, and Saladin Meckland-Garcia, “Against Online Public Shaming: Ethical Problems with Mass Social Media”, 8.

- e) The public expression of particularly nasty forms of communication, such as “abusive and vitriolic attacks”.¹⁸

While these alternative definitions raise interesting topics in their own right, they either fall outside the scope of my project, or are limited to a narrower set of activities than what I am interested in. Definitions (a) and (b) involve the use of humiliation by single agents or small groups of agents. They therefore do not raise the same set of issues as the kind of public shaming I want to focus on, which we will discuss throughout this thesis. Meanwhile, definitions (c)-(e) are narrower than my definition. They refer to public shaming as a set of acts with very specific features, like acts intended to inflict shame or cause stigmatization. I depart from these definitions because I am interested in the harm, generally construed, that groups of people inflict in cases of public shaming. While shame, stigmatization, and experiencing abuse may represent some of this harm, targets of public shaming often experience other morally significant costs as well.

I can make this point clearer by considering the relationship between public shaming and the emotion of shame. In my view, public shaming should not be defined in such a way that it necessarily causes shame, or that it necessarily aims at causing shame. Why? For one, the public shaming of an individual may impose serious harm on that individual, but fail to cause shame. Imagine if Justine Sacco, for instance, was simply immune to feeling shame, and that everyone knows she is immune to feeling shame (and thus will not aim to cause her shame). Even so, she was still vulnerable to the other costs of public shaming – like losing her job, becoming socially isolated, suffering the indignity of being insulted, etc. – which are causes for moral concern. The moral concerns that arise from these costs are among the key interests in this project, which is why I want a definition of public shaming that is not necessarily connected with shame. Nevertheless,

¹⁸ Billingham, Paul, and Tom Parr, “Enforcing social norms: The morality of public shaming,”

public shaming is contingently connected with shame, because it reliably causes shame. As James Fritz notes, “the paradigmatic grounds for shame are experiences in which one is (or sees oneself as) exposed before, and judged to be deficient in the eyes of, a group of others”.¹⁹ This analysis of shame is corroborated by other philosophers.²⁰ Because public shaming necessarily involves a high degree of publicity, it is very likely to expose a target to the eyes of others. As a consequence, most people who experience public shaming will feel shame. And this feeling of shame is a cost that we have to take into account when we think about the moral status of public shaming.

The goal of this project is not conceptual analysis, but the analysis of public shaming in the sense that I’ve described it. Those who dislike my use of the term “public shaming” are invited to substitute my use of the term “public shaming” with alternate phrases, like “online shaming” or something similar.²¹ Nevertheless, my use of the term “public shaming” to refer to the kind of collective shaming we typically see online is not entirely idiosyncratic either. Billingham and Parr define public shaming as “a practice of public moral criticism in response to violations of social norms”,²² and in their analysis, they pay particular attention to collective forms of public shaming, like online public shaming. They also point out that “public shaming” in folk usage has recently become associated with mass shaming, like online shaming. This can be seen in the works of non-philosophy academics like Jennifer Jacquet, or journalists like Jon Ronson.²³ In fact, a quick Google

¹⁹ Fritz, James, “Online Shaming and the Ethics of Public Disapproval,” *Journal of Applied Philosophy* 38, no.4 (2021): 691.

²⁰ See, e.g. Stephen Bero, ‘The audience in shame’, *Philosophical Studies* 177 (2020): 1283–1302; Cheshire Calhoun, “An apology for moral shame,” *The Journal of Political Philosophy* 12, no. 2 (2004): 127–146.

²¹ See Fritz, James, “Online Shaming and the Ethics of Public Disapproval,” *Journal of Applied Philosophy* 38, no.4 (2021): 686-701, who uses “Online Shaming”. Or Aitchison and Meckled-Garcia who refer to it as “Online Public Shaming”. I personally prefer not to use these terms, because the fact that shaming occurs online is not itself morally significant. A mass shaming event that occurs offline, such as on the radio or in the streets of a closely-knit town, seems morally concerning in the same ways that paradigmatic online public shaming does. The online nature of shaming is only instrumentally significant because its being online makes it easier for large numbers of people to participate in the shaming.

²² Billingham, Paul, and Tom Parr, “Enforcing Social Norms,” 2.

²³ Jacquet, Jennifer, *Is shame necessary? : new uses for an old tool*; Ronson, Jon, *So You’ve Been Publicly Shamed*.

search for “public shaming” mostly brings up articles related to mass shaming, online shaming, “call-out culture”, and “cancel culture”.²⁴

IIb. Other Accounts

I want to briefly compare my definition of public shaming to some existing definitions which aim to address the kind of mass shaming relevant to this project. These definitions are the ones provided by Billingham and Parr, and Guy Aitchison and Saladin Meckland-Garcia. My goal here is not to critique these accounts. Rather, a brief analysis of their definitions can help highlight some key features of my own use of the term “public shaming”.

Billingham and Parr define public shaming as “a practice of public moral criticism in response to violations of social norms”.²⁵ This definition, of public shaming as public criticism, captures many of the paradigmatic cases of public shaming I listed at the beginning. It is a starting point for my own view, which aims to precisify their definition. My definition effectively distinguishes between *collective public shaming* and *shaming in public* – both possible interpretations of the phrase “public moral criticism”. Collective public shaming refers to the shaming carried out by large numbers of people in public, and is the activity picked out by my definition. In contrast, shaming in public refers to any shaming carried out in a public setting. For example, a father berating his son on a public bus for failing to do his homework. This distinction is potentially important, because the moral considerations associated with collective public shaming may not be present in cases of shaming in public. For one, collective public shaming typically imposes certain kinds of harm that shaming in public typically doesn’t impose. Collective public shaming involves

²⁴ Max, D.T. “The Public Shaming Pandemic,” *The New Yorker*, 21 Sep 2020, <https://www.newyorker.com/magazine/2020/09/28/the-public-shaming-pandemic>; Blackford, Russell, “The Shame of Public Shaming,” *The Conversation*, 4 May 2016, <https://theconversation.com/the-shame-of-public-shaming-57584>; Mohammad, Farah, “The Danger of Public Shaming in the Internet Age,” 25 Jan 2018, <https://daily.jstor.org/the-danger-of-public-shaming-in-the-internet-age/>.

²⁵ Billingham, Paul, and Tom Parr, “Enforcing Social Norms”, 2.

criticism from many members of a community, which can easily in turn dissuade other members of the community from interacting with the target. As a result, collective public shaming can often result in social isolation, or feelings of social isolation. In contrast, the criticism of one person in public is less likely to cause isolation, unless others join in the shaming (in which case shaming in public turns into collective public shaming). Consider Peter Singer, who has criticized, very publicly, the consumption of meat. He has published papers and books on animal liberation, comparing factory farming with the oppression of racial minorities. But without the concerted effort of many members of society to criticise meat-eaters, Singer's criticisms alone do not cause meat-eaters to feel socially isolated. Some moral considerations associated with collective public shaming are not as salient with respect to shaming in public.

Another account of public shaming, proposed by Aitchison and Meckland-Garcia, is that (online) public shaming involves the public derision of "another person's moral character as having a feature that renders that character, in the shamer's view, as transgressive".²⁶ Aitchison and Meckland-Garcia do characterize public shaming as a collective activity, "in the sense of people sharing an aim that requires others to join them to be effective, and that they act to bring about".²⁷ But unlike my own account, according to Aitchison and Meckland-Garcia, public shaming is *intentionally stigmatizing*. They state that people participating in public shaming "aim to publicly present their [target's] identity as being not worthy of participating in (certain) normal social relations, civil conversation, or debates as a moral equal".²⁸ The public shaming is "pinned to a

²⁶ Aitchison, Guy, and Saladin Meckland-Garcia, "Against Online Public Shaming: Ethical Problems with Mass Social Media", 8.

²⁷ *Ibid.*

²⁸ *Ibid.*

specific norm transgression by the target (though not necessarily illegal ones) which is treated as evidence of the morally faulty character that renders the target excludable”.²⁹

Aitchison and Meckland-Garcia’s account seems too restrictive to capture the examples that we saw earlier. In many cases of public shaming, it is unlikely that *all* people engaging in it have the intention to portray their targets as unworthy of participating in normal social relations as equals. Many shamers may be happy for their target to participate in society, at least once the target has apologized or paid their dues. There are also many other things that can motivate an agent to contribute to public shaming. While many contributors to public shaming desire that wrongdoers suffer some costs for perceived wrongdoing, those costs need not come in the form of stigmatization (as opposed to a loss in income or influence).³⁰ Contributors to public shaming may also be motivated by a desire to let others know their commitment to certain norms. This behaviour is called by Justin Tosi and Brandon Warmke “moral grandstanding”, where agents are motivated by the desire “that other people recognize her as morally respectable”.³¹ Contributors may also be motivated to raise awareness about an issue and educate others,³² to pressure employers or legal institutions to punish perceived wrongdoers,³³ or simply to express their own feelings about a target and their wrongdoing. Alternatively, as I shall argue in chapter IV, contributors may be motivated by the desire to shore up their own self-respect, or the desire to engage in *moral self-defence*. When a contributor is motivated primarily by one of these goals –

²⁹ *Ibid.*

³⁰ Tandoc, E. C., Tan Hui Ru, B., Lee Huei, G., Min Qi Charlyn, N., Chua, R. A., & Goh, Z. H. “#CancelCulture: Examining definitions and motivations,” *New Media & Society*, published online, (2022), <https://doi.org/10.1177/14614448221077977>.

³¹ Tosi, Justin, and Brandon Warmke, “Moral Grandstanding,” *Philosophy and Public Affairs* 44, no. 3 (2016): 199.

³² Tandoc, E. C. et. al. “#CancelCulture: Examining definitions and motivations”.

³³ Also see Skoric, Marko M., Jia Ping Esther Chua, Meiyan Angeline Liew, Keng Hui Wong and Pei Jue Yeo, “Online Shaming in the Asian Context: Community Empowerment or Civic Vigilantism?” *Surveillance & Society* 8, no. 2 (2010): 181-199.

grandstanding, expressing commitment, moral self-defence, etc. – the stigmatization of the wrongdoer may just be a side effect of their behaviour, not an aim.

Further, I think it is valuable to present an analysis of public shaming which begins from a more dialectically neutral starting point than Aitchison and Meckled-Garcia. What I mean is that they define public shaming in a way which focuses their attention on a particularly negative form of public shaming. They claim that public shaming aims “to publicly present their [target] as being not worthy of participating in (certain) normal social relations, civil conversation, or debates as a moral equal”.³⁴ As a result, it is not surprising when they conclude that public shaming is “a failure of fundamental respect for people as separate human beings with distinct lives to live and inherently incompatible with due process”.³⁵ What of online criticism that aims at expressing indignation at wrongdoing, or raising awareness of injustice, like the criticism within the #Metoo movement? For Aitchison and Meckland-Garcia, these phenomena simply do not count as online public shaming. In their view, online public shaming is distinct from criticism that is intended to raise awareness about important issues, legitimate forms of exclusion, “online actions aimed at warning others about a person's behavior”, and even some “deprecatory characterisations of personality”.³⁶ Aitchison and Meckled-Garcia thus exclude these sorts of cases from their analysis of “online public shaming”. But these sorts of cases fall squarely within the scope of my project, and the set of issues which I’m interested in. Thus, I need a definition of public shaming which is different from Aitchison and Meckled-Garcia’s account. I find that the most philosophically interesting cases of public shaming are cases which aren’t obviously unjust or disrespectful. This includes cases like those within the #Metoo movement, which can seem morally justified (all-

³⁴ Aitchison, Guy, and Saladin Meckland-Garcia, “Against Online Public Shaming: Ethical Problems with Mass Social Media”, 8.

³⁵ *Ibid.*

³⁶ *Ibid.*, 29-30.

things-considered), despite many worrying features like the harm and stigmatization imposed on the shamed. If we define those cases away as not cases of public shaming, then we may be left with cases which are obviously wrong and philosophically less interesting.

III. The Rest of the Thesis

Having done some definitional work, I shall provide a summary of the rest of the thesis. In a sense, chapters II, III, and IV form the “core” of this thesis, wherein I investigate the right that people have against being publicly shamed, and what, if anything, justifies public shaming. By determining the rights of the relevant parties, and the conditions under which those rights may be infringed, we gain a foothold with which we can pursue further inquiry. In chapters V and VI, I turn to further issues associated with the issue of public shaming: the moral responsibility of individual contributors, and what individuals should do when they face uncertainty in cases of public shaming.

In the second chapter of my thesis, I claim that human beings have a right against public shaming. Such a right is defeasible – it can be outweighed, forfeited, or waived. Nevertheless, this right entails that we may not contribute to public shaming unless there are considerations which outweigh or undermine the right against being publicly shamed. The chapter will involve an explication of the various harms that are caused by public shaming, including deleterious effects on a target’s mental and physical health, the loss of esteem, which leads to a loss of access to important social networks (like friends, family, and professional networks), and a loss of trust and the ability to personate (the ability to make commitments to others). I argue that absent some special justification, we are not generally permitted to impose these harms on a person, or contribute to an imposition of these harms. Therefore, people have rights against being publicly shamed. Another reason for why contributing to public shaming can be wrong is because it constitutes a form of *exclusion*. As I shall claim, exclusion is the act of expressing to others that you

are unwilling to form, or even entertain, certain relationships associated with particular individuals. I argue that because human beings have strong interests in enjoying associative goods, like close relationships, professional networks, and friendly relationships with members of their communities, they also have claims against being excluded from the formation of valuable associations for no good reason.

This brings us to the third and fourth chapters of the thesis. Both of them ask: given that human beings have rights against being publicly shamed, what (if anything) justifies public shaming? Rather than survey every possible justification for public shaming, my approach is to focus on a few particularly salient justifications. In chapter III, I examine a view endorsed by a number of philosophers who have written on this topic: that public shaming is (sometimes) justified because it enforces valuable norms by deterring violations, expressing support for norms, and so on. While I ultimately agree that norm enforcement plays an important role in justifying public shaming, I think that this view, which I call the Norm Enforcement View, is incomplete. It faces a challenge, or a gap in the argument supposed to justify public shaming: private individuals are typically not permitted to impose serious harms in order to enforce norms. Thus, we need an explanation for why a group of private individuals may collectively impose serious harms *via public shaming*. To my knowledge, most philosophers writing on this topic have not supplied this explanation satisfactorily.

There are a few potential answers to this challenge. We might claim that certain wrongdoers forfeit their rights against being harmed, that some wrongdoers waive their rights against being publicly shamed, or that there are conceptual links between a norm and a community member's right to enforce the norm. I argue that none of these answers satisfactorily show that private individuals may contribute to public shaming. Ultimately I defend the view that a person's right to contribute to public shaming is grounded in their natural right to protect their vital

interests. Because certain norms can protect a person's vital interests, or can express the value of a person's vital interests, agents may be permitted to impose harm on each other to enforce norms which protect their vital interests. As a result, some cases of public shaming can be considered forms of collective self-defence, where a group of agents impose collective harms on an individual in order to enforce a valuable norm, which in turn protects the vital interests of the group's members.

In the fourth chapter, I expand on the claim made in the fourth chapter that public shaming can function as a form of collective self-defence. I develop my own account of justified public shaming: public shaming as a form of *moral self-defence*. As I shall claim, public shaming can sometimes be an important tool for individuals to defend their moral standing: the perception of their moral value by others and themselves. In other words, individuals can contribute to public shaming in order to protect their interests in maintaining recognition respect from others, and to preserve their self-respect. How so? Many wrongdoings, but particularly serious wrongdoings related to systemic injustice (like racism), can constitute claims that some individuals are of lower moral standing. Public criticism, expressions of blame, resentment, contempt, or anger are often necessary for individuals to "stand up for themselves" and protect their moral standing. Thus, a contribution to a case of public shaming can be characterized as a form of self-defence: an act carried out by an agent in order to protect her vital interests. While these acts can sometimes cause or contribute to harm, they can be justified as long as certain conditions are met. This is a powerful way of thinking about public shaming, because it is uncontroversial that self-defence can sometimes justify imposing very severe harms on a culpably aggressor. In turn, this fact gives us pretty good reasons to think that the harms caused by public shaming can be justified under some circumstances. However, there are some challenges that have to be overcome. Namely, it must be shown that public shaming as moral self-defence meets the criteria of permissible self-defence –

the existence of an unjustified threat, grounds to prefer the defender's interests, proportionality, and necessity.³⁷ Accordingly, some of the fourth chapter will be dedicated to arguing that some cases of public shaming do involve conditions that meet these criteria.

In the fifth chapter of this thesis, I examine the relationship between the overall morality of the case of public shaming and the moral responsibility held by individual contributors to that case. Public shaming is, in part, what may be called a *collective harm case*. A case of public shaming involves contributions from many individual agents, and many of those individual contributions by themselves will not make a significant difference to the overall effects of the case of public shaming. Given that individual contributions will often not themselves make a significant difference to the overall outcome of public shaming, what can we say about the moral responsibility held by individual contributors to public shaming?

The topic of collective harm has been extensively written on in contemporary philosophy, addressing a variety of topics like climate change, voting, and consuming animal products. Multiple accounts have been offered to explain how individual contributors to such cases can be morally responsible for harmful outcomes, or how they can act wrongly. My goal in this chapter is not to evaluate all of these accounts. Rather, what I want to do is highlight an aspect of public shaming which sets it apart from many other collective harm cases in the literature. In cases of public shaming, many agents possess shared intentions to help bring about collective outcomes – outcomes that are realized through the contributions of many agents. Some contributors to cases of public shaming may believe that their targets deserve to suffer harm, and may intend for the targets to be harmed. Others may want to enforce valuable norms, and intend to deter violations of those norms by way of causing harm to violators. In contrast, contributors in many other

³⁷ This list is adopted from Seth Lazar, "Necessity in Self-Defense and War," *Philosophy and Public Affairs* Vol 40, no. 1 (2012): 3-44.

collective harm cases, like climate change, do not typically intend to bring about the collective outcome (in cases like climate change, many people contribute in spite of disvaluing the collective outcome). I argue that when a person has a shared intention to help bring about harm together with others, and acts on this shared intention, they are morally responsible (in a way that they would not otherwise be). The implication of this discussion is that contributors to public shaming can be held to account for the harms caused by that case of public shaming, even if they do not individually make a significant difference to the amount of harm caused.

Finally, chapter six is about the relationship between public shaming, knowledge, and evidence. Specifically, it is about the standard(s) of proof that we ought to hold ourselves to when we want to contribute to public shaming; if you want to publicly shame, how much evidence do you need that your target deserves to be shamed? Some commentators have suggested that in cases of public shaming, we should use specific standards used in various legal contexts, like the Beyond a Reasonable Doubt standard or the much lower Preponderance of Evidence standard. I argue that these standards should not be applied to public shaming. Among other reasons, these standards fail to take into account the different moral stakes that exist in different cases of public shaming. I further argue that the use of a single principle, like “a person’s guilt must be proven beyond a reasonable doubt” cannot adequately take into account the moral stakes present in the variety of cases of public shaming. Instead, I propose that we rethink the SOP in cases of public shaming, conceiving of it as a framework for deliberation. Such a framework is a set of principles which together helps individuals and communities deliberate on a set of issues and achieve good outcomes. I sketch out and defend a version of a framework for deliberation, arguing that it is more helpful in enabling communities and individuals to take into account the moral stakes in each case of public shaming.

Chapter II: The Right against Public Shaming

I. Introduction

When a group of people publicly shame an individual, they make life significantly less pleasant for the individual. In his book, Ron Jonson details his encounters with victims of public shaming. One of them, the author Jonah Lehrer who was shamed for plagiarism, reported that he felt “like he’d been inside some spinning-machine designed by aliens to test the effects of stress on humans”.³⁸ Lehrer describes “all-consuming” regrets.³⁹ Another, Justine Sacco, who posted a racist tweet on Twitter (which she alleges was intended as an ironic joke), received threats and struggled to “avoid falling into ‘depression and self-loathing’”.⁴⁰ The costs to those publicly shamed can be very high indeed.

But just because public shaming is bad for the shamed, doesn’t automatically mean that it is immoral or unjust. After all, it is sometimes permissible to act in ways which foreseeably cause harm to others. For example, it seems permissible to set up a business competing with someone else’s, even if doing so will put them out of business and leave them destitute. It also seems permissible to voice your criticism of a movie, a work of art, or a politician, even if doing so may make the targets of your criticism worse off. In these sorts of cases, we typically say that individuals do not have rights against being harmed in these ways. Hence, it’s worth asking: do human beings have a right against being publicly shamed? And what are the duties associated with this right?

In this chapter, I argue that human beings have a (defeasible) right against being publicly shamed. Absent sufficiently strong countervailing considerations, it is impermissible or

³⁸ Ronson, Jon, *So You’ve Been Publicly Shamed* (London: Picador, 2015), Chapter 3.

³⁹ *Ibid.*

⁴⁰ *Ibid.*, chapter 4.

unjustifiable for a person to contribute to public shaming (I shall use the words “impermissible”, “unjustifiable”, and “objectionable” as synonyms). I argue that there are two main reasons to affirm such a right. First, when you contribute to a case of public shaming, you contribute to a collective outcome which can be expected to impose a set of severe harms on a target. I detail some of the harms that come about as a result of public shaming, and explain why individuals have rights against the imposition of those harms. Second, a contribution to a case of public shaming can also be wrong because it *constitutes* an objectionable way to treat another person. Your contribution to public shaming can be a form of what I shall refer to as *exclusion*. That is, you are communicating to someone that you are not willing to associate, or consider associating with them. As I shall argue, we have general duties not to exclude others for irrelevant reasons. As a result, contributing to public shaming can not only contribute to a larger set of unjustified harms, but also involve a violation of a duty not to exclude another person for irrelevant reasons.

My focus here is on duties that people have with respect to the target of public shaming, and how contributing to public shaming can wrong that person. A case of public shaming, and the broader practice of public shaming can also have adverse effects on other people, or on other important societal practices. We may also have reasons not to publicly shame that are grounded in those effects. But I shall set aside those reasons in this piece.

This chapter shall be structured as follows. In part II, I examine the harms and costs imposed by public shaming, and argue that human beings have rights against these harms and costs. In part III, I examine a claim that comes from the philosophical literature on privacy and rights to privacy, that individuals have the right to control their public image or social identities. I argue that this right can only ground duties not to publicly shame in some cases. In part IV, I argue that public shaming can also be wrong when it excludes others for no good reason.

But first, a quick note on the way I talk about “rights” in this chapter. I will use the Hohfeldian analysis of rights, in which agents have “claim-rights” which correspond to duties that others have with respect to them.⁴¹ When a person has a claim-right (or simply, right) against the imposition of some harm, there is either some agent who has a duty not to impose the harm, or there is a group of agents who collectively have duties not to bring about the harm. Thus, if a person has a right against being publicly shamed, there is some group of agents who together have duties not to publicly shame that person. Another aspect of the Hohfeldian analysis relevant here is “privileges”: when an agent has no duty with respect to some sort of behaviour.⁴² In this chapter, I consider arguments that people have rights to engage in certain behaviours, like public criticism. In my analysis, I treat some of these arguments as saying that people have privileges with respect to those behaviours.

IIa. Public Shaming Causes Harm

I have two aims in this section. First, I shall describe how public shaming can cause harm to its targets. Second, I shall argue that these harms can be wrongful, and that people can have rights against the imposition of those harms. For the purposes of this discussion, I shall adapt Joel Feinberg’s definition of “harm”, which identifies harm as a “setback to interests”.⁴³ We have interests in many things, like the fulfilment of our desires, pleasure and avoiding pain, and having access to valuable goods and opportunities. “Harm” is thus a fairly broad category which includes the imposition of suffering, the frustration of morally relevant desires, or the deprivation of valuable goods.

⁴¹ Wenar, Leif, "Rights," *The Stanford Encyclopedia of Philosophy* (2023), Edward N. Zalta & Uri Nodelman (eds.), <<https://plato.stanford.edu/archives/spr2023/entries/rights/>>.

⁴² *Ibid.*

⁴³ Feinberg, Joel, *The Moral Limits of Criminal Law Volume 1: Harm to Others* (New York: Oxford University Press, 1984), 31.

Public shaming typically imposes a number of harms, including:

- Adverse psychological and physiological effects associated with humiliation, social isolation, and shame.
- The loss of esteem from others.
 - In turn, this results in the loss of important friendships, professional networks, access to political community, and other associations.
- The loss of the ability to personate.

While I present these different types of harm separately, this presentation is in large part a matter of clarity and organization; many harms related to public shaming are related to each other. For instance, a victim who experiences feelings of isolation may shy away from social opportunities, which in turn may cause further isolation. So it is worth noting that in presenting my arguments in the following way, I do not suggest that each category of harm is distinct and unrelated to each other.

The Psychological and Physiological Effects of Public Shaming

When we assess the costs and risks that public shaming can impose on its targets, what stands out are dramatic mental health impacts. Victims of public shaming, like Lehrer and Sacco, report feeling stress, fear, and depression.⁴⁴ Lindsey Stone, a woman shamed for making a disrespectful gesture at the Arlington National Cemetery, reported that even a year after the shaming, she was constantly “terrified” that others would find out about what she did.⁴⁵ Sacco reported worries and anxiety about meeting new people; “it’s not like I can date, because we google everyone we might date. So that’s been taken away from me too. How am I going to meet new

⁴⁴ Ronson, *So You’ve Been Publicly Shamed*.

⁴⁵ *Ibid.*

people? What are they going to think of me?”⁴⁶ These feelings about the future are, by themselves, very painful, and can lead to further problems, like a loss of hope and motivation. In some cases, these feelings can be so consuming that they result in self-harm and suicide.⁴⁷ For example, following the #Metoo movement, a number of suicides were attributed to the public shaming and outrage that targets received in response to allegations of sexual misconduct.⁴⁸

Public shaming is associated with being humiliated, which in turn results in serious health effects – as Torres and Bergner note, “Suffering severe humiliation has been shown empirically to plunge individuals into major depressions, suicidal states, and severe anxiety states, including ones characteristic of posttraumatic stress disorder”.⁴⁹ Victims of humiliation report “feelings of hopelessness and helplessness”, as well as “profound feelings of worthlessness”.⁵⁰ They also note that some victims of humiliation become more prone to “powerless rage” and violence, which transfers some of the costs of public shaming to others in physical proximity to the target.⁵¹ Marit Svindseth and Paul Crawford report that public shaming can result in losses of self-esteem, depression, feelings of worthlessness, self-mutilation, or suicide.⁵² Public shaming can also result

⁴⁶ *Ibid.*

⁴⁷ Pridmore S, McArthur M, “Suicide and Reputation Damage,” *Australasian Psychiatry* 16, no. 5 (2008):312-316; Thompson, J. D., and Cover, R, “Digital hostility, internet pile-ons and shaming: A case study,” *Convergence* 28, no. 6(2022): 1770-1782.

⁴⁸ Peretz, Evgenia, “How Jill Messick’s Suicide Reflects the Tragedies of the #MeToo Era,” *Vanity Fair*, 6 Sep 2018, <https://www.vanityfair.com/hollywood/2018/09/how-jill-messick-suicide-reflects-the-tragedies-of-the-metoo-era>; “Seoul’s mayor found dead in presumed suicide after #MeToo allegation,” *France 24*, 9 Jul 2020, <https://www.france24.com/en/20200709-body-of-seoul-s-mayor-found-after-massive-search>; “Swedish opera singer Anne Sofie von Otter blames #MeToo for her husband’s suicide,” *The Sydney Morning Herald*, 31 Jul 2018, <https://www.smh.com.au/entertainment/music/swedish-opera-singer-anne-sofie-von-otter-blames-metoo-for-her-husbands-suicide-20180731-h13de0.html>; “Carl Sargeant: Welsh politician who resigned amid sexual misconduct claims is found dead,” *ABC News*, 8 Nov 2017, <https://www.abc.net.au/news/2017-11-08/welsh-politician-who-resigned-amid-misconduct-claims-dies/9128596>.

⁴⁹ Torres, Walter J., and Raymond M. Bergner, “Humiliation: Its Nature and Consequences,” *Journal of the American Academy of Psychiatry and the Law* 38, no. 2 (2010): 201-202.

⁵⁰ *Ibid.*

⁵¹ *Ibid.*

⁵² Svindseth, Marit F., and Paul Crawford, *Humiliation: Mental Health and Public Shame* (Bingley: Emerald Publishing Ltd, 2019).

in stigmatization, social isolation or perceived social isolation, due to the fact that people do not usually want to associate with those perceived to be immoral. This in turn can result in further health effects. In a philosophical analysis based on empirical studies,⁵³ Kimberley Brownlee observes that the effects of social isolation include “anxiety-inducing physiological reactions”, elevated blood pressure, increased depressive symptoms, “obesity, the progression of Alzheimer’s disease, increased vascular resistance, diminished immunity, reduced independent living, alcoholism, suicidal ideation and behaviour, over-expression of genes bearing pro-inflammatory responses, and mortality in older adults”.⁵⁴

Many of these health effects seem bad in themselves. But in addition, Torres and Bergner note that these effects can “pre-empt” other motivations in a person’s life.⁵⁵ The feelings and motivations that surround a case of humiliation can overcome a person’s desire or capacity to engage in other valuable activities. As they note, “The individual cannot focus on these matters, cannot concentrate, and is constantly distracted by the humiliating situation and its implications”.⁵⁶ They report the case of “Jane”, who “found herself so completely consumed by her situation that she could not concentrate at work, in social situations, or even while trying to watch a television show. Many evenings she could not sleep because of her obsessive preoccupation with her circumstances”.⁵⁷ Thus, the mental health effects of public shaming are also instrumentally bad, in that they can prevent agents from pursuing other valuable activities or goals.

⁵³ Hawkey, L.C. et al, “Loneliness Predicts Increased Blood Pressure: 5-year Cross lagged Analyses in Middle-aged and Older Adults,” *Psychology and Aging* 25 (2010): 132-141; Cacioppo, J.T. et al, “Loneliness as a Specific Risk Factor for Depressive Symptoms: Cross-sectional and Longitudinal Analyses,” *Psychology and Aging* 21 (2006): 140-151.

⁵⁴ Brownlee, Kimberley, “A Human Right against Social Deprivation,” *The Philosophical Quarterly* Vol. 63, No. 251 (2013): 211.

⁵⁵ Torres and Bergner, “Humiliation: Its Nature and Consequences”: 201.

⁵⁶ *Ibid.*

⁵⁷ *Ibid.*

James Fritz argues that public shaming also reliably causes shame, and defends this conclusion by referring to some empirical studies.⁵⁸ I shall not dispute those studies here. If they are right, then shame is yet another one of the harms that public shaming can inflict on its targets. As Fritz points out, shame is “associated with dispositions to hiding, isolation, alienation, depression, and anger”.⁵⁹ These dispositions can be harmful. A disposition for hiding, isolation, and alienation can lead to actual isolation and alienation. As we’ve seen, these are very unhealthy outcomes for the isolated individual. Furthermore, anger, another effect of shame, can harm both the angry individual and those around her. In extreme cases, anger manifests itself in obvious violence, both physical and non-physical. But even in milder cases, a person disposed to anger can be quite unpleasant, and the quality of her relationships can degrade. In short, by causing shame, we cause a person to suffer a number of other harms.

Public Shaming and the Loss of Esteem

Public shaming involves expressing a withdrawal of esteem. That is, when you publicly shame someone, you express that you’ve either lost some pre-existing esteem that you had for them, or that you now hold them in negative esteem. For most humans, the withdrawal of esteem is instrumentally and non-instrumentally bad. Let us start with the instrumental badness. The esteem of other individuals greatly shapes our social interactions, such that being in the good graces of others makes it much easier to navigate the social world, and vice versa. So losing the esteem of others, particularly in public, makes our lives harder. One reason for this, as pointed out by

⁵⁸ Fritz, James, “Online Shaming and the Ethics of Public Disapproval,” *Journal of Applied Philosophy* Vol 38, no. 4 (2021): 6. Studies cited by Fritz include: Smith, Richard H., J. Matthew Webster, W. Gerrod Parrott & Heidi L. Eyre, ‘The role of public exposure in moral and nonmoral shame and guilt’, *Journal of Personality and Social Psychology* 83, no. 1 (2002): 138–159; June P. Tangney, Rowland S. Miller, Laura Flicker & Deborah Hill Barlow, ‘Are shame, guilt and embarrassment distinct emotions?’, *Journal of Personality and Social Psychology* 70 (1996): 1265–1269; H. G. Wallbott & K. R. Scherer, ‘Cultural determinants in experiencing shame and guilt’ in June P. Tangney & Kurt Fischer (eds) *Self Conscious Emotions: Shame: Guilt, Embarrassment, and Pride* (New York: The Guilford Press, 1995), pp. 465–487.

⁵⁹ *Ibid.*

Geoffrey Brennan and Philip Pettit, is that the esteem of others has an “evidential value” for myself and others.⁶⁰ When you enjoy the esteem of others, individuals (including yourself) have some reasons to think that you are “living up to the ideals assumed in the background evaluations”.⁶¹ For example, someone esteemed as a very truthful person will be assumed to be truthful in new interactions. This is mostly good for her, as it will incentivize other agents to engage in positive interactions with her. Conversely, someone disesteemed as an immoral person will be assumed to be immoral by other agents. This results in downstream harms. When someone disesteems me, they are less likely to associate with me, in either professional or more personal settings. And when there is common knowledge among my community that most people disesteem me, members of the community will be reluctant to associate with me. Thus, victims of shaming often lose social networks, like family, friendships, professional relationships, and access to political communities. In many cases of public shaming, targets lose their jobs and other career opportunities, like presenting at conferences. These social networks can be both intrinsically and instrumentally valuable, and so losing them can both be intrinsically and instrumentally bad for a person.

Even when a victim of public shaming does not lose access to her social networks, she may find the nature of their relationships changed, which can in turn affect the value she attains from the relationship. For instance, a woman shamed for having an affair may become extra wary of all her interactions with married men, lest others interpret her actions as flirtatious. This in turn may limit her ability or willingness to pursue further professional opportunities or expand her social network in valuable ways.

⁶⁰ Brennan, Geoffrey, and Philip Pettit, *The Economy of Esteem: An Essay on Civil and Political Society* (Oxford: Oxford University Press, 2004), 26-27.

⁶¹ *Ibid.*

Suffering the disesteem of others can also be intrinsically bad for the person suffering it. Human beings value being esteemed even when “there is little or nothing to be gained in pragmatic or evidentiary terms”.⁶² Think, for example, about how environmental activists make appeals along the lines of “what will future generations think of us?” In making this appeal, activists rightly assume that we desire the esteem of future generations, even though they cannot interact with or benefit us. Likewise, most of us desire the esteem of other human beings, and in particular the esteem of those close to us – our friends, family, and even members of our community. Public shaming, which causes a loss of esteem for the target among audience members, frustrates the desire to enjoy a certain level of esteem.

Public Shaming, Stigmatization, and Personation

Another harm that can be caused by public shaming is that it can prevent people from *personating*. As defined by Philip Pettit:

Agents personate insofar as they assume a special authority in communicating their attitudes to others and, at least in general, live up to their words in practice: they act as the attitudes communicated would warrant. They speak for themselves, inviting others to trust their words: to rely on their displaying the profile or persona that they project in what they say.⁶³

Personating is thus the set of acts where an agent expresses to others elements of their moral and rational agency – their ability to hold attitudes and make intentions, ability to take on responsibility (of the morally relevant sort), their ability to act on those intentions, and make plans to act on the intentions. Agents express these elements by making commitments to others and themselves,

⁶² Brennan and Pettit, *The Economy of Esteem*, 29.

⁶³ Pettit, Philip, “My Three Selves,” *Philosophy* Vol 95, no. 3 (2020): 369.

which in turn express an intent to act in a certain way, and expose themselves to “a higher expense than normal – and so makes the communication more credible than normal – should they fail to act on the advertised disposition”.⁶⁴ These commitments can either be explicit, through practices like promising, or tacit.

According to Pettit, a person’s ability to personate is constitutive of personhood.⁶⁵ While I will not attempt to defend this particular claim, it at least seems that the ability to personate is of significant instrumental value. Personating is a crucial ingredient in gaining the trust of others. As a result, many of our important and valuable interactions with others depend on our ability to personate. Close personal relationships, like family and friendships, depend on us making commitments to each other (either explicitly or tacitly), avowing certain values, and endorsing the expectations we might have of one another. Other important relationships, like professional relationships, also depend on these practices. Even our ability to relate to others in casual contexts (eg. making a transaction at a store) depends on people personating in some ways. For example, when I purchase a good from a vendor, I assume and expect that the vendor will deliver the good as promised, and that the good will be of decent quality. I have these expectations based on either explicit (eg. guarantees) or tacit (eg. opening a business in an environment with certain social norms and consumer protection laws) commitments from the vendor that they will abide by the expectations that I have of them.⁶⁶

⁶⁴ *Ibid.*

⁶⁵ *Ibid.*

⁶⁶ Pettit (2020) claims that even living in certain societies, under the right conditions, can constitute as making some commitments. He writes,

Suppose you and I live in a society, for example, where it is manifest to all that certain regularities hold and that everyone is expected in general to live up to them without question . . . I will know in the range of everyday interactions that others manifestly expect me, without any question, to accept those norms and to be willing to conform; indeed, I will also know that, should I fail, then appealing to excuses of the misleading- or changed-mind sort will not wash . . . And by acquiescing in those expectations, I will effectively avow acceptance of the norms and pledge conformity to them.

A person can lose some of her ability to personate when she cannot express to others her commitments, beliefs, and intentions. This can happen if she loses certain physical or mental capacities. But it can sometimes also happen when a person has been stigmatized, which has been associated with being discredited or degraded.⁶⁷ Attempts made by a stigmatized individual to express a commitment to the relevant norms are often met with scepticism or mockery. One reason for this is that because being stigmatized is very unpleasant, stigmatized individuals have obvious prudential incentives to apologize, ask for forgiveness, and promise to change their behaviour. As a result, we might suspect that when a stigmatized person expresses commitment to a norm that they previously violated, they may be doing so out of self-interest rather than a sincere belief in the norm. Another reason for why stigmatization can result in a loss in the ability to personate is because of a loss of fundamental trust – the trust that some other agent will behave in accordance with basic norms regarding respecting and refraining from harm or interfering with others. In serious cases of stigmatization, a person can be presented as “being not worthy of participating in (certain) normal social relations, civil conversation, or debates as a moral equal”.⁶⁸ Being stigmatized in this way can be particularly hard to come back from, because if a community views an agent as unworthy of participating in normal social relations, or fundamentally immoral in some way, it can be extremely difficult for the agent to convince others that they are committed to the norms held by the community. For example, if everyone believes that you are a compulsive liar, or an amoral psychopath, or that you share no common ground with them, they will not take your expressed commitments seriously. This effect can be exacerbated by common knowledge, where agents believe that you are immoral, and believe that others share this belief as well (and believe that others share the belief that everyone has that belief, and so on). When there is common

⁶⁷ Loughlin, M, Dolezal, L, Hutchinson, P, Subramani, S, Milani, R, Lafarge, C. “Philosophy and the clinic: stigma, respect and shame,” *J Eval Clin Pract.* 28 (2022): 705- 710.

⁶⁸ Aitchison, Guy, and Saladin Meckland-Garcia, “Against Online Public Shaming: Ethical Problems with Mass Social Media,” *Social Theory and Practice* 47, no.1 (2021): 8.

knowledge that a person has behaved immorally, it is extremely difficult for them to salvage their reputation.

Torres and Bergner make a similar set of observations. Based on their empirical studies, they argue that humiliating events (like public shaming) can not only take away a target's status (as legitimate occupants of certain social positions), but also a target's "status to claim a status".⁶⁹ As they write:

When humiliation annuls the status of individuals to claim status, they are in essence denied eligibility to recover the status that they have lost. They have effectively lost the voice to make claims within the relevant community and especially to make counterclaims on their own behalf to remove their humiliation. (This is evident, not only conceptually but phenomenologically. The person who is suddenly humiliated in a group is typically left feeling stunned and speechless, unable to counterclaim.)⁷⁰

Public shaming often has an effect like this. Some cases of public shaming can present a person as fundamentally immoral, antisocial, or unworthy of attention. As a result, they can take away a target's ability to reaffirm, reconcile, or make new moral commitments, and thus remove their ability to recover the social status that they previously lost.

My claim here is not that public shaming is *necessarily* something that a person cannot come back from. Even in the face of public shaming, it is possible for a person to salvage their reputation enough for others to take their commitments seriously. Nevertheless, there are features of modern, online, public shaming that make such "comebacks" more challenging. One such feature is that public shaming is often *massive*. It is often carried out by large numbers of people. Modern public

⁶⁹ Torres and Bergner, "Humiliation: Its Nature and Consequences," 198.

⁷⁰ *Ibid.*

shaming, often on the internet, could involve thousands of shamers. Imagine trying to convince the members of your community that you are committed to a set of norms, in the face of thousands of others criticizing you for a violation of that very norm. This is an uphill task. Another feature of public shaming is that *first impressions are decisive*. When a person is publicly shamed, many people who interact with the shaming – either by contributing to it or by witnessing it – will see some information about the target: who they are, what they did, etc. Most will form an impression of the target based on this information, and then move on. They will not do research on the target to find out more about them, they will not hang around to hear what the target has to say, nor will they spend evenings pondering the deservingness of the target. The first impressions of the target – who has been portrayed as a racist, sexist, etc. – is decisive for most people who contribute to or witness a case of public shaming. Finally, online public shaming tends to have some *inescapability*. Traces of online shaming, like articles on the internet, can hang around for a long time, and can be accessed from any location. Thus, a target cannot escape the effects of public shaming simply by lying low for a few months, or by moving to a new town.

As a result, while public shaming does not necessarily affect a target's ability to personate, it may often have this effect. People who have been publicly shamed may struggle to express their commitment to moral beliefs and norms, because members of their community may not take them seriously. In some cases, the harm may be localized if a target's credibility has been affected only in relation to a particular norm. For example, someone shamed for academic dishonesty may not be taken seriously when she tries to publish or when she goes to an academic conference. In other cases, the harm is more global when a target has been portrayed as untrustworthy in all domains, amoral, or "beyond the pale". Someone shamed for being a sex offender, for example, will probably struggle to gain the trust of others in a very wide range of interactions, not just in intimate contexts.

IIb. Do we have Rights against Public Shaming?

It is clear that public shaming can set back the interests of its targets, and therefore harm them. But it doesn't immediately follow that individuals have rights against the imposition of those harms. As I noted in the introduction, there are some cases where it is generally considered permissible to contribute to some harm. Let's call these kinds of harms *presumptively permissible harms*. In cases of presumptively permissible harms, an agent is not typically required to offer some special justification, based on some special features of her specific position, for the harm she imposes. For example, if you want to set up a business rivalling mine, you are not required to explain why *you*, of all people, may threaten my revenue in such a way. Neither are you required to explain how I might have forfeited or waived my right not to be harmed, or explain that your actions are all-things-permissible because they serve the greater good. Rather, you could probably just point out that I never had a right to operate my business in a competition-free environment, which can foreseeably lead to losses. I do not have the right not to be harmed via fair competition (perhaps because of the presumptive good of such a competitive system and the privilege of anyone to seek to further their self-interest within it, at whatever cost to the interests of other competitors). In a similar vein, some people might think that public shaming also imposes presumptively permissible harms, because people do not have a right against being publicly shamed.

Why might we think that people do not have a right against being publicly shamed? One thought is that each person has the right to freedom of speech, which is a privilege which entails that agents have rights to say what they desire, which in turn entails that no one has a right against suffering the consequences of speech. The view is an absolutist view about speech rights, which claims that everyone is morally permitted to speak as they desire. But this is an implausible position, as there are clearly some types of speech which are morally impermissible, like hate speech, slander, and speech which results in foreseeably disastrous consequences (like yelling "fire" in a crowded

theatre). These types of speech are very harmful, and agents typically don't have strong, morally legitimate, interests in carrying out these types of speech.

Most people are not absolutists about freedom of speech. In fact, I think that the freedom of speech typically refers to a claim-right against others interfering with you or preventing you from speaking, not a moral privilege. We may not be prevented or coerced from making most kinds of speech acts. Nevertheless, we may still act wrongfully when we engage in such speech. For example, if I want to make risky political speech which may contain harmful ideas, the government (or other private citizens) should not prevent me from making such speech. But if my speech is truly harmful, and I do not have sufficiently strong moral reasons to make the speech, I could wrong others by making it. Or for another example, consider artistic or intellectual criticism. Again, we have a right to engage in artistic or intellectual criticism, in that no one may prevent us from making such criticism, or coerce us by threatening us with severe costs. But we can wrong others by engaging in some criticism, particularly if we lack any justification for the criticism. If I make baseless public critiques about a person, and do not justify these critiques, I can wrong them. Thus, people can have some rights against being publicly criticized, and these rights are compatible with the freedom of speech as a claim-right against interference.

Some may find the claim I have just made too strong. Surely, we do not wrong somebody just by criticizing them without justification. For example, if I say, "I don't know why, but I just don't like Tom Hanks", it seems implausible that I have *wronged* Tom Hanks. But what we're interested in this chapter are instances of public criticism which cause or contribute to significant harm. When I say, "I just don't like Tom Hanks", I am unlikely to cause harm to him. There is no harm or expected harm for which I am on the hook for justifying. But suppose instead that I am an influential public intellectual, and I proclaim that Tom Hanks is a terrible person, and that no one should watch his movies or hire him (and suppose that my actions come with an expectation

of harm for Tom Hanks). In this case, it does indeed seem right that I wrong Tom Hanks, unless I can supply some justification for why I think he is a terrible person, and why we ought to boycott his movies. And if I cannot supply this justification, it would be appropriate for others to react to me with attitudes like blame or disapproval.

Instead of defending a general privilege to contribute to public shaming, an interlocutor might suggest that individuals lack rights against some specific harms caused by public shaming. They may argue that while it is bad for a person to lose esteem and valuable social associations, it is not wrong to withdraw these two kinds of goods from a person. They may argue that in the first place, esteeming someone does not seem to be something you can do at will. And if we don't have voluntary control over how much you esteem someone, it seems that you do not have a duty to give someone a certain level of esteem. It also seems that no particular person has a right to your association. You are not required to impart the good of your friendship on anyone. Nor are you required to associate with someone professionally, or give them a certain amount of esteem. So we might think that it's not wrong to disassociate with a person as well, even if doing so causes them harm like social isolation or a loss of social or professional opportunities.

Let us first examine whether humans have the right against losing esteem. The loss of esteem can refer to both a loss of positive esteem, and a gain in disesteem. A loss of positive esteem may be the withdrawal of approval, admiration, or positive appraisal. For example, if I learn that a famous actor delivers an average performance, I may lose some esteem that I previously had for her. I might think of her badly, I just don't think of her as well as I once did. In contrast, a gain in disesteem involves some negative attitudes, like blame, anger, contempt, derision, pity, or disgust. For example, if I learn that the famous actor was abusive to her assistants, I will probably lose some prior esteem that I may have had for her, and also disesteem her. As I noted earlier, it

is unclear if we can have duties to esteem or disesteem others, because we do not always have control over whether we esteem someone else or not.

However, we can have duties regarding acts associated with esteem or disesteem. In their book, Brennan and Pettit discuss the concept of “esteem services”: “voluntary acts whereby one person may raise the probability of another's enjoying esteem”.⁷¹ In a similar vein, we may characterize contributions to public shaming as “disesteem services”: voluntary acts whereby a person increases the probability of another's suffering disesteem. Contributions to public shaming constitute disesteem services, because they make it more likely for a target to experience disesteem. Intuitively, we are not usually required to perform positive esteem services, like praise or providing high ratings. Individuals generally do not have a right to praise. But it does seem that we wrong others by providing disesteem services, such as criticism, scorn, mockery, or expressions of anger, when those disesteem services are likely to bring about further harm. For instance, if I were to publicly criticize you, claiming that you are immoral in some way, I could raise your chances of experiencing disesteem. In response, you would probably want to demand some justification from me. It would be unsatisfactory and callous if I went on to say “look, you have no right against my disesteem, or my expressing my dislike of you, so deal with it”. Now, my claim is not that all instances of criticism require justification. As we saw earlier, there are some cases where people may be permitted to say, “I just don't like him”. But insofar as instances of criticism (and some other disesteem services) have a high likelihood of causing significant amounts of disesteem and harm, they do seem to require justification. Thus, it seems that human beings can have some rights against disesteem services like public criticism.

⁷¹ Brennan, Geoffrey, and Philip Pettit, *The Economy of Esteem*, 55.

Now let us examine the claim that we are not required to associate with other people. We take it for granted that agents are generally not required to form associations with particular people. But it does not follow from this that human beings do not have rights to certain associations. A number of philosophers have recently argued that human beings have rights to some associations, or the goods associated with associations. These include the right to friendship,⁷² the right to be loved,⁷³ rights against social deprivation,⁷⁴ and the right to opportunities for meaningful relationships.⁷⁵ In particular, I think that one right humans do have is to some existing associations. Consider friendship. Suppose you are not required to befriend any particular person. But once you have established a close friendship with someone, it does seem that your friend can acquire some rights to maintain this friendship. It would be wrong under some circumstances for you to terminate that friendship without adequate explanation or justification. Part of why this is the case is that what it means to be in that relationship is to value that relationship, and to view the other party as entitled to some of your continued time, effort, and abilities. If a close friend were to stop speaking to me for no reason, I would suspect that she was never a true friend in the first place. Furthermore, we grow to depend on relationships after some time. We build our lives, our plans, and our expectations on our relationships. So losing a friend can harm me in a way in which I wouldn't be harmed if I never had that friendship in the first place. Even if no person has a right to a certain association, like a friendship, they can have a right to enjoy the continued existence of the association once it's been formed.

⁷² Wittrock, John, "A human right to friendship? Dignity, autonomy, and social deprivation," *The International Journal of Human Rights* 26, no. 9 (2022): 1590-1607.

⁷³ Liao, Matthew S., *The Right to be Loved* (Oxford: Oxford University Press, 2015).

⁷⁴ Brownlee, Kimberley, "A Human Right Against Social Deprivation," *The Philosophical Quarterly* Vol. 63, No. 251 (2013): 199-222.

⁷⁵ Couto, Alexandra, "A Right to Opportunities for Meaningful Relationships," in *Being Social: The Philosophy of Social Human Rights*, eds. Kimberley Brownlee, David Jenkins, and Adam Neal (Oxford: Oxford University Press, 2022).

A person's right to enjoy her existing associations are held primarily against the participants of the associations. That is, the primary duty-bearer in each case is the person in association with the right-holder. For example, if I have a right to enjoy my existing friendships, my friends are the primary duty-bearers. However, a person's right to enjoy her existing associations can have secondary duty-bearers, which can include many members of the wider community. How so? Members of the wider community can set conditions which make it easier or harder to form and maintain relationships with certain individuals. For example, by spreading nasty rumours about a colleague, I could stigmatize them. If a person is stigmatized badly enough, it could become costly to maintain a (non-secret) relationship with that person. Nobody wants to be seen as being friends with a bad person, lest they be seen as condoning their bad behaviour. Furthermore, as noted earlier, public shaming can also affect a person's ability to personate, which in turn affects their ability to attain the trust of others. By setting conditions which make it easier or harder to form and maintain associations, members of a wider community can contribute to a person's ability to enjoy her existing associations. Consequently, when members of a community make it very hard for a person to maintain her existing associations, they can contribute indirectly to an infringement of her right to enjoy those associations.

Are we morally responsible for the loss of an association, when we contribute indirectly to its loss? Consider this example:

False Gossip: I spread a false rumour that Bridget has been stealing other people's lunches from the pantry. As a result, she loses her friends.

In the above case, it seems I have wronged Bridget (via causing her to lose her friendships), even though it was her friends who made the ultimate decision to cut her off from their friendship. This shows that it is possible to be responsible for causing someone to lose valuable associations, even when other agents ultimately make the decision to dissolve those associations.

You may point out that the case *False Gossip* contains a number of confounding factors, including the fact that I've slandered Bridget, and that I have acted maliciously. This is true. But we can consider a further example:

Social Pressure: Rahul is falsely accused of sexual misconduct. As a result, he is publicly shamed by members of his community. They also shun those who remain friends with Rahul, and threaten to boycott his company unless he is fired. Following this, his friends cut him off from their friendships, and his company fires him.

In the above case, let us hold fixed that each contributor to the case of public shaming acts in good faith, and contributes only because they actually believe that Rahul has committed sexual misconduct. Holding this fixed, it still seems that the contributors bear some responsibility for the harm caused to Rahul, because they set conditions which made it likely that he would lose valuable associations. Specifically, the contributors applied social and economic pressure on Rahul's friends and his bosses to cut him off. This example, *Social Pressure*, seems similar to many cases of public shaming, where public shamers do in fact call on friends and employers to terminate relations with the target, and threaten them with sanctions (like shaming and boycotts) if they do not terminate their relations. In such cases, it does seem that public shamers are at least partly morally responsible for the harms caused by the loss of association.

Must Public Shaming be Necessarily Harmful?

To briefly recap, public shaming is harmful because it can cause social isolation, a number of devastating mental and physical health effects, and feelings of shame. It also causes its targets to lose the instrumental good of esteem, as well as valuable relationships and connections. And in some cases, public shaming can affect a person's ability to personate, which further impacts their ability to build trust and form associations with others. But how reliable is the connection between

public shaming and these effects? It might be thought that there's nothing *necessarily* wrong with public shaming, because it's not necessarily linked to these harms. We can imagine for example a world of perfectly "civil" shamers, who express themselves in very temperate ways like "I disapprove of the way you behaved", or "that was immoral of you". We might also imagine that in this world, there are organizations and associations that actively ensure that publicly shamed individuals don't suffer the effects of isolation – organizations like Alcoholics Anonymous, but for people who have been publicly shamed. We could further suppose that in this world, there are clear and accepted norms regarding reintegration, which make it easier for targets to normalize their relationships with others. It seems likely that perfectly civil shaming will produce far fewer negative effects in this world than the shaming we see in our world. Perhaps we can claim that public shaming is unproblematic as long as it is very civil.

I agree that public shaming isn't necessarily harmful. But my aim in this chapter is not to show that we necessarily have duties to avoid contributing to public shaming. Rather, the goal is to show that we have duties to avoid contributing, given the current practices and norms about public shaming (and the practices we will probably have in the near future). The world that we imagined in the last paragraph is clearly a massive idealization. In our world, there is a reliable connection between being publicly shamed and suffering the harms outlined in this chapter. And individual instances of public criticism can (and often do) contribute to a person suffering from the effects of public shaming, even when those instances of public criticism are civil. While an individual contributor may not be directly responsible for the harms that a target suffers, as I argued earlier, people can be indirectly responsible for some harm if they knowingly set conditions which make it likely for a target to suffer harm. Thus, even if your contribution does not violate any independent moral norms, like norms associated with civility, you can have duties to avoid

making that contribution because of the foreseeable collective harm that results from that contribution and others like it.

Furthermore, it is not even clear that civil shaming can avoid imposing many of the negative effects associated with public shaming. Civil shaming can still impose certain costs for targets of public shaming. These costs involve the loss of a number of instrumental and non-instrumental goods, like the loss of some valuable social connections. For example, someone shamed for racist behaviour may suffer disesteem, and thus have a hard time finding a date, even if the shaming was very civil. It is also hard to believe that civil shaming will cause zero negative psychological effects. For one, public shaming is very likely to cause feelings of shame, regardless of how civil it is. Recall that shame is a feeling paradigmatically caused by a perceived exposure of one's deficiencies to others. As a result, public shaming tends to cause shame.⁷⁶ Most of us cringe even when our small wrongdoings, such as failing to recycle, are pointed out politely by random strangers. So it's not surprising that public shaming, which involves subjecting a person's wrongdoing to a large audience, causes shame even when we try to shame people in civil ways.

III. Public Shaming, Labelling, and the Loss of Control

There is one more possible harm of public shaming that I want to discuss – which I have saved till now, because I think it only generates duties not to publicly shame only in limited situations. We might worry that public shaming is wrong because it *labels* its targets.⁷⁷ Each of us has a *social identity* – who we are to others, which in turn can affect how we view ourselves. Public shaming can impart onto people a certain social identity, which they may not have chosen themselves. In giving someone an unchosen social identity, we label them. For example, Amy

⁷⁶ Fritz, “Online Shaming and the Ethics of Public Disapproval”.

⁷⁷ Thomason, Krista, *Naked: The Dark Side of Shame and Moral Life* (New York, NY: Oxford University Press, 2018), 208-209.

Cooper, the woman who called the police on a Black man, became known as the “Central Park Lady”, effectively defining her social identity according to that incident. Labelling can be bad for a person in a number of ways which we have already discussed. If you are labelled as a racist, for example, your reputation will be harmed. You will lose the esteem of others, and you may lose your access to a number of valuable social goods, like friendships or job opportunities. Because we have already covered these harms earlier, and why it is wrong to impose these harms, these harms are not the focus of this part.

Another thought is that labelling is wrong because it deprives them of the opportunity to define themselves. When we label someone, we give someone a certain social identity. For example, as a liar, or a joker, or a tough guy. Sometimes, this given identity can be so overwhelming that it threatens to overshadow the identity the target sees themselves as having, or would like to have. For example, when Amy Cooper became the “Central Park Lady”, her identity as a racist person, based on this one interaction, became her defining feature to others on the internet, and overshadowed her own attempts to identify herself. As Philip Pettit puts it, “the image pinned on me may not only run counter to how I personate sincerely and reliably; it may silence any attempts to correct the image imposed and to change the expectations raised. It may deprive me of the capacity to speak for myself in a way that commands the credence of others, discounting or recasting the sorts of things I might say”.⁷⁸ Thomason also captures this concern when she argues that when “we shame others we put ourselves in a position to define their identities in social life. We decide what their identity should be, and we invite everyone else to interact with them accordingly”.⁷⁹ When we publicly shame someone, we force onto them an identity – usually a bad

⁷⁸ Pettit, Philip, “My Three Selves”, 383-384.

⁷⁹ Thomason, *Naked*, 205.

one. In the process, we take away any agency they have to control their own social identity, or “who they are” to others.

Thus, the thought is that one reason for why labelling is wrong is because it violates an interest that we have in controlling our social identities. As I shall argue, contributing to public shaming is only wrong *qua* labelling under specific circumstances. For example, if you have a very large following, who pay attention to your contributions on social media. But if you only make an average contribution to public shaming, which doesn’t have a very large effect, you do not wrong your target *qua* labelling.

There are two main claims we need to analyse to assess the claim about labelling. The first is an empirical one – that public shaming is the kind of activity that can force an identity onto someone, and “silence” their attempts to define themselves. It seems plausible that it can. Modern public shaming is often carried out on the internet, which has several features which make a forced identity a likely outcome of public shaming. As I noted earlier, online public shaming is usually *massive*, relative to the size of the relevant community. In many cases today, hundreds or thousands of people participate or contribute to the shaming of one person. It’s quite hard for the shamed to push back against the tide of disapproval, offer an explanation or justification for their behaviour, or salvage their image. Pettit notes that a given identity can “silence any attempts to correct the image imposed” on a person. This is especially the case in online public shaming, when the voice of a single person can be drowned out in a chorus of criticism. As Aitchison and Meckled-Garcia argue, targets of public shaming do not have the ability “to respond in a way that has an equal and proportionate force to the numbers involved in stigmatising them. A “conversation” of

tens, hundreds, or thousands can take place about a person, in which she is labelled with a shaming term and which excludes her from participating in any meaningful way”.⁸⁰

Secondly, public shaming, but especially online shaming, is not conducive to nuanced discussion. On social media platforms, posts are rewarded for being short. The shorter a post, the more likely it is to be shared or retweeted.⁸¹ So there’s not much incentive to make nuanced, balanced, and empathetic claims about someone else’s behaviour. In fact, because of the effective word limits imposed by social media, nuanced claims can be misinterpreted for making “excuses” for the wrong side. Furthermore, cases of public shaming often contain instances of “*moral grandstanding*”, where contributors compete with each other to express their values in the most emphatic way possible.⁸² As a result, some individuals are incentivized to “ramp up” the severity of their condemnation by making sweeping, sometimes essentializing claims about their targets. For example, a mother who leaves her child in a car for five minutes may be characterized as a monster, rather than an otherwise loving parent who made a mistake. A charity worker who repeats a racist joke may be presented as a committed racist. And so on. Thus, a target’s social identity can be reduced to a simple set of traits, which in cases of public shaming, are always negative.

The final feature of online public shaming which facilitates the power of a given identity is its *inescapability*.⁸³ The reputation that you develop on the internet is hard to shake. You cannot move to a new city to run from a reputation on the internet. If you look for a new job, your

⁸⁰ Aitchison, Guy, and Saladin Meckled-Garcia, ““Against Online Public Shaming: Ethical Problems with Mass Social Media,”: 11.

⁸¹ In market studies, it’s been found that the optimal length is 70-100 characters, while the optimal Facebook post is 40 characters. See Lee, Kevan, “Infographic: The Optimal Length for Every Social Media Update and More,” *Buffer.com*, 21 Oct 2014, <https://buffer.com/library/optimal-length-social-media/>.

⁸² Tosi, Justin, and Brandon Warmke, “Moral Grandstanding,” *Philosophy and Public Affairs* 44, no. 3 (2016): 199.

⁸³ See Daniel J. Solove, who remarks that “Being shamed in cyberspace is akin to being marked for life. It’s similar to being forced to wear a digital scarlet letter or being branded or tattooed. People acquire permanent digital baggage”, in *The Future of Reputation* (New Haven and London: Yale University Press, 2007), 94.

employer will Google you. Even people you date are likely to look you up online. Once we've forced an identity onto someone via public shaming, it's difficult for them to meet new people without having the shadow of this identity loom over them. For these reasons, I think that the empirical claim is on solid grounds. We are likely to force an identity onto someone by publicly shaming them, particularly on the internet.

The second claim we must examine is normative – that we have a duty not to label someone, thus pinning on someone an identity that they may struggle to change. Thomason seems to make this claim when she argues that we should avoid shaming, writing:

When we shame others, we make the social part of their identities sovereign: because you have come across as a racist to us, that is what you are. We assume that we have “read” your comments or behavior in exactly the right way such that we feel confident calling on the community to publicly condemn you. How you see yourself does not matter, and we can explain any of your protests away as attempts to alleviate yourself of your guilt or your responsibility.⁸⁴

We need to disambiguate between two claims: that it is bad to give someone a bad label; and that it is wrong to label. We already know that it is bad to give someone a bad label (if they are not liable). Therefore, for the sake of this analysis, I am interested in the latter claim. Those who seem to defend this claim seem to argue that each person has some right to define their social identities to some extent, and that labelling someone can infringe on this right. Andrei Marmor, in discussing the right to privacy, claims that every person has an “interest in having a reasonable measure of control over ways in which we present ourselves to others and the ability to present different aspects of ourselves, and what is ours, to different people”.⁸⁵ As a result, it is wrong even to label

⁸⁴ Thomason, *Naked*, 205.

⁸⁵ Marmor, Andrei, “What is the Right to Privacy?” *Philosophy and Public Affairs* 43, no. 1 (2015): 7.

someone with a neutral or positive image. This is because even a neutral or positive image can reduce a person's control over how they present themselves when an overwhelming number of people view that person through the label.

However, we do not have a right, or even an interest, in having complete control over our social identities. This point is acknowledged by Marmor, Thomason, and Pettit.⁸⁶ Marmor, in particular, makes the specific claim that we have an “interest in having a reasonable measure of control over the ways in which we present ourselves”.⁸⁷ It is this interest that grounds a person's right to privacy, in Marmor's view. There are two key notes worth making here. First, there are many ways we present ourselves, in different contexts. Thus, the interest in controlling one's social identity is context-sensitive. For example, my interests in controlling my social identity at work may be different from my interests in controlling my identity among members of my family. My interest in controlling my social identity in a given situation is a function of whether it is reasonable for me to have control in that situation. Second, control comes in degrees. For any given set of moral practices involving how people talk about each other, each individual will have a degree of control over their social identities. The question then, is *how much* control a reasonable person would want over their social identity, in a given situation.

Since we do not have a right or interest in having *complete* control over our social identities, then we have no duty to refrain from reducing a person's control in some cases. It is only wrong to reduce a person's control *beyond reason, within a given context*. It is notoriously difficult to pin down what it means for what counts as a “reasonable” measure of control. Presumably, a reasonable amount of control over my social identity in a situation depends on the strength of my interests in exercising control in that situation, compared to the relative strength of competing interests.

⁸⁶ Pettit, “My Three Selves”, 385-386; Marmor, “What is the Right to Privacy?” 7; Thomason, *Naked*, 205.

⁸⁷ Marmor, “What is the Right to Privacy?” 7.

Marmor makes a similar claim, suggesting that reasonableness in the context of his discussion on privacy is a function of two main factors: “the relative importance of the type of activity in question, and the drastic way in which the policy restricts people’s ability to have any control in that domain over ways in which they can present themselves to others”.⁸⁸

With this in mind, let us start by asking a fairly broad question: can individuals reasonably expect a high degree of control over their social identities in public spaces like the internet? Let us suppose that a “high degree of control” entails that others must ask for my permission before talking about me on the internet. Given this starting point, I do not think individuals can reasonably expect a high degree of control. Many instances of online speech acts, including sharing videos and photos of others, mentioning others, retweeting or sharing posts made by others, telling stories about others, and sharing articles written by others, can affect the social identity of other people. Even stories that are somewhat anonymized (eg. “I have a colleague who said...”) can affect the social identity of someone, because the internet gives us the tools to make inferences about the identity of a person. If each person has a right to a large degree of control over their online reputations or social identities, it would be usually impermissible to carry out most of online speech acts without obtaining the permission of relevant parties. The practice of online speech, or public speech, would have to be very different from the way we practice it today. This counterfactual outcome comes with a high social cost. While we might not have strong interests in each instance of talking about another person, or sharing a story involving another person, we do have strong interests in having the general permission to talk about others. Sharing stories about our friends, family, and colleagues is an integral part of the way we communicate, both in person and on social media. We tell these stories not just because they are fun, but because they are a big part of how we relate to others. It seems unreasonable for me to demand that others usually refrain

⁸⁸ *Ibid.*, 12.

from talking about me without my permission. Thus, we do not have a right (and possibly, an interest) in having a high degree of general control over online speech about us.

But even though we do not have a right against a practice, broadly construed, we might still have rights against sub-practices or particular acts which are parts of the practice. Just because the practice of public speech by itself is valuable, doesn't mean that all instances of public speech are morally justified. How can we go about deciding which sub-practices are permissible, and which aren't? Sometimes, a sub-practice is central to a broader practice – if we imposed a rule against the sub-practice, we would lose much of what's valuable about the broader practice. For example, the practice of education as we know it would not continue, and would likely be far worse, if we did not hire teachers. So the sub-practice of hiring teachers is central to the broader practice of education. In contrast, the sub-practice of imposing detention on misbehaving children is not as central to the practice, as a ban on detentions would not change the practice of education beyond recognition, or decrease the quality of education dramatically (I suspect). It seems that if a practice is very valuable, a sub-practice which is central to the practice is more likely to be very valuable as well.⁸⁹

With this in mind, we can see why we do not have a general right against being publicly shamed, on the grounds that public shaming exerts an unreasonable amount of control over our social identities. A case of public shaming is primarily constituted by individual acts of public shaming, expressions of anger, blame, etc. These individual acts are also central to the broader practices of public/online speech. One distinctively valuable aspect about our practices of public

⁸⁹ There are further questions about how we can delineate sub-practices from one another, for the purposes of assessing whether a sub-practice is central to a practice. For example, we clearly shouldn't be asking: "may people with last names starting with C contribute to public shaming?" This particular way of delineating a sub-practice is arbitrary. Presumably, there is a set of criteria of morally relevant features which we may use to delineate relevant sub-practices from each other. However, it is beyond the scope of this piece to develop this criteria.

or online speech is our ability to make public criticism. It is also constitutive of our practice of public speech that we can make public criticism. Thus, to assert a right against being criticized is to assert a right against our current practices of public speech. Because it is not reasonable to expect control over public speech about us, it is unreasonable to expect control over public criticism about us. And given that public shaming is constituted by instances of public criticism, we do not have a general right against being publicly shamed (on the grounds of controlling our social identities).

However, there may be particular kinds of contributions to public shaming which people can have rights against, because of the relative strengths of the interests involved. One example is publicizing certain kinds of wrongdoing done in private. Consider someone who overhears two other people making racist or sexist jokes, and publicizes their actions by taking a picture of them or recording them (or “calling out” others for their conversations). Adria Richards, an attendee at a tech conference, did this to two men who reportedly made sexist jokes while waiting for a talk to begin.⁹⁰ While reporting bad behaviour is a valuable practice which helps communities to enforce important social norms, we also have very strong interests in being able to carry out some conversations in private. In my view, the potential benefits of a norm which would allow people to listen in on others’ conversations in order to catch potentially bad behaviour is outweighed by the costs of such a norm: not being able to carry out quiet conversations in public places without the comfort of privacy.

People might also have rights against particular contributions which have the power to influence many others. For example, imagine a celebrity who joins a case of public shaming, thus influencing many of her fans to pile on as well. Her contribution might have a disproportionate

⁹⁰ Jon Ronson details this event in *So You’ve Been Publicly Shamed*.

impact on the target's social identity, because of the sheer amount of influence she has. I think that we have very strong interests against being subject to the whims of very powerful or influential people. This interest may be strong enough to overcome the interests that very influential people have in making public criticisms, at least in some cases.

The point I've made in the last few paragraphs is that an interest in having a reasonable amount of control over my social identity does not ground a general right against public shaming. This is because public criticism is central to our practices of public speech, and it seems very unreasonable to expect a high degree of control over the way people talk about me publicly. But my interest in controlling my social identity could ground a right against certain ways of contributing to public shaming.

IV. Public Shaming and Excluding

Here's one more way in which people have a right against being publicly shamed. When you contribute to public shaming, you can *exclude* them. As I shall use the term, exclusion has two components. First, there's a dispositional component, which involves the pre-emptive "ruling out" of others from potential associations. For example, if you've decided, based on my weird laugh, that you can never work with me, you've ruled out the possibility that I can become your professional associate. You are thus disposed not to undertake any activity, and perhaps oppose any activity that will bring us closer to becoming friends. Second, exclusion comes with an active component, which involves acting on the disposition not to associate with someone in some relevant way. This can often involve some kind of communicative act, like letting a person know that you are not willing to associate with them. For example, a teenager who puts their bag on an unoccupied seat to prevent a specific person from sitting there communicates that they want no association with that person. Other times, the active component can involve deliberately removing

someone from a decision-making process. For example, a hiring manager who throws away someone's job application without reading it.

We can and often do exclude people for good reasons. But sometimes people are excluded for no reason, or for irrelevant reasons. I claim that human beings have rights against being excluded from certain associations for no reason or irrelevant reasons. Because contributing to public shaming can constitute exclusion, a contribution can wrongfully exclude someone when there is no relevant reason in favour of the contribution. Now, people often contribute to cases of public shaming because they perceive that the target has violated serious norms. This seems like a relevant reason to exclude someone from a large range of associations. Thus, the key claim in this section is conditional: a human being has a right against being excluded via public shaming *iff* they have not in fact made themselves liable through some relevant wrongdoing. I shall start by establishing that we have duties not to exclude for irrelevant reasons, and then explain how contributions to public shaming can constitute forms of exclusion.

It seems wrong to pre-emptively rule someone out from being a friend, a professional opportunity, or some other associations, without any good reasons. This appears to be the case even regardless of actual or likely outcomes. Imagine, for instance, that upon meeting me, Zander deliberately rules me out from being his friend for no good reason (not because he's just not in the market for friends, but because of some immediate, irrational dislike). Unbeknownst to both of us, he and I don't share any hobbies, and would have never become friends anyway. Despite this, I am still wronged by Zander's original snub. Or imagine that you apply to give a talk at a research conference. The conference organizer is having a bad day, and throws away your application without looking at it. As it happens, your application was significantly worse than all the other applications, and never stood a chance at being accepted. Regardless, it seems that you are wronged by the conference organizer. We can also think about real-life cases like boycotts

against individuals. Take, for example, the boycott of Israeli academics, where participating boycotters refuse to associate with Israeli institutions and their members.⁹¹ It's not my goal to make an all things considered judgment about such boycotts. However, I think that journal publishers and conference organizers cannot defend a boycott by saying something along the lines of "this Israeli academic was not entitled to publish in the journal anyway". While it may be true that the Israeli academic is not entitled to publish in a specific journal, or attend a specific conference, she has a right to be considered for the journal on the basis of the academic merit of her work. It's possible that this right can be defeated or outweighed by some other consideration. But it's incumbent on the person ruling out to explain why this consideration is relevant to ruling out the academic.

We might raise a couple of questions about the duty not to exclude. First, what is the scope of the duty not to exclude for irrelevant reasons (ie. What kind of associations are we required not to exclude people from)? The duty not to exclude, in my view, covers most of the potential associations which a human being needs to enjoy a decent life. These include some close personal relationships, like friendships, important professional associations, and some forms of membership in political communities. That is to say, you should not rule out other people for no good reason when you want to form these sorts of relationships. The reason for this, as I shall argue shortly, is that it is extremely important for human beings to enjoy some of these relationships. Close personal relationships are non-instrumentally important for a person's well-being, and for a person's mental and physical health. Professional relationships, on the other hand, are instrumentally very important for a person's ability to make a living and for them to obtain meaningful work. Membership in some political communities help safeguard a person's

⁹¹ "Academic Boycott," *BDS*, retrieved 13 Dec 2023, <https://bdsmovement.net/academic-boycott#guidelines>

fundamental interests and rights. Because relationships are extremely important for human beings, they have strong interests in having adequate opportunities to make these relationships. These interests generate claims on other people to provide these opportunities by not excluding for no good reason.

Another question is what counts as a relevant reason to rule someone out of a potential association. The answer to this question varies from case to case, depending on the type of association in question, the importance of the association to all relevant parties, and the level of partiality that we typically deem appropriate for the association in question. For example, in romantic associations, a lot of things can count as relevant reasons, because many features can affect our sexual preferences. If you want to date only people with bushy eyebrows, small noses, or large earlobes, that's your prerogative.⁹² On the other hand, you cannot rule someone out pre-emptively from a job opportunity because of their hair colour. Therefore, what counts as exclusion is different in different contexts. Fortunately, we do not need to resolve this particular set of issues for our purposes. What I want to establish at this point is that human beings have a right against being publicly shamed, which is partly grounded in the right against being excluded. Like many other rights, the right against pre-emptive exclusion can be defeated by various considerations. I will explore those considerations in subsequent chapters.

A third issue is that it might be thought that given my earlier definition of "harm" as a setback to interests, being excluded is merely another category of harm. This might be correct. But the last section considered ways in which *a case* of public shaming could cause various harms, and explained how individuals can wrong others by contributing to those harms. In contrast, this

⁹² There may be exceptions to this. It may be wrong to exclude potential romantic partners on the basis of their race, for instance. Why this could be so is a difficult question. But perhaps it's got something to do with the fact that stereotypes about the attractiveness of members of certain races play a role in perpetuating racism in general.

section argues that individual contributions to public shaming can constitute a form of exclusion, given our current practices regarding public shaming. My arguments, if true, imply that the acts which contribute to public shaming, like public criticism or expressions of anger, can be wrongful in and of themselves, regardless of what the overall effect of the case of public shaming is.

What explains our presumptive duties not to exclude? As I mentioned earlier, most human beings need some associations to live a decent life. We thus have important associative needs, such as the need for friendship and love. As Brownlee summarizes, our social connections are “central to our nature as social creatures”, “vital for respect for our autonomy”, and “necessary prerequisites for the development and maintenance of cognitive, physical, emotional and social ability”.⁹³ These claims are supported by numerous empirical studies on the effects of social isolation, a few of which I’ve already mentioned. Even perceived social isolation has been noted to result in the production of stress hormones, impulsive behaviour, “increasing inflammatory biology”, and “the decreased salubriousness of sleep”, which together can contribute to “early morbidity and mortality”.⁹⁴

However, we cannot immediately claim that the associative needs of human beings generate claim-rights on other agents to fulfil those needs. We also have associative freedoms, which entail some privileges to choose whether or not to associate with others. Generally, it is thought that we are morally free to associate with whomever we choose, with a few constraints, such as a constraint against refusing to associate for racist reasons, or requirements to maintain associations with certain people like family and close friends. But for the most part, we don’t think it’s wrong not to make friends with a person. We are even permitted to make no friends, if we so

⁹³ Brownlee, Kimberly, “Ethical Dilemmas of Sociability,” *Utilitas* 28, no. 1 (2016): 72.

⁹⁴ Cacioppo, John T., and Stephanie Cacioppo, “Social Relationships and Health: The Toxic Effects of Perceived Social Isolation,” *Soc Personal Psychol Compass* 8, no. 2 (2014): 58-72.

choose. Unfortunately, our associative freedoms lie in tension with our associative needs. If we are not required to associate with anyone specifically, many people may have their associative needs unmet. For instance, a person generally regarded as “weird” or unfriendly may end up with no friends. But it is commonly held that no one has wronged her, because no one was required to befriend her.

Brownlee resolves this tension by coming down firmly on the side of associative needs. She claims that “we have moral responsibilities to make our social resources available, in principle, to the collective social pool if required . . . [and to] acknowledge the social claims that people can have on us . . . in virtue of our proximity to them, our awareness of their situation and the relative ease for us of relieving their social suffering”.⁹⁵ If I think that John has trouble making friends, and if I know that he is much worse off for his lack of friends, I ought to make my social resources available to him, befriending him if necessary. A similar view is articulated by Stephanie Collins, who argues that each of us has “relationship-forming duties”, at least in the case of friendships.⁹⁶ A duty not to exclude for no good reason falls out of these duties advocated by Brownlee and Collins. After all, it’s very hard to give our social resources to someone if we’ve excluded her from social opportunities. So if you accept their proposals, you should accept a duty not to exclude for no good reason.

Many people might consider the claims made by Brownlee and Collins too revisionary. They are indeed costly claims, because they imply that our freedom of association is much more limited than we might like. Furthermore, the thought that I am morally required to befriend an entirely incompatible person in some situations seems too demanding, and intrudes on an area of my life where it seems important to enjoy a significant prerogative in my decisions. It may also be

⁹⁵ *Ibid.*

⁹⁶ Collins, Stephanie, “Duties to Make Friends,” *Ethical Theory and Moral Practice* 16 (2013): 915-916.

pointed out that a “duty to befriend” is contrary to the meaning of friendship – a person who only calls herself your “friend” because some sense of obligation does not seem like a true friend. A less costly alternative is to endorse a duty not to exclude for no good reasons.⁹⁷ If you discover some relevant reasons not to form an association with someone, you are not required to associate with them. Otherwise, you are required to at least consider them as a potential candidate for association.⁹⁸ This duty takes us some of the way towards meeting the associative needs of individuals. If we do not pre-emptively rule out individuals based on irrelevant traits, many people currently excluded by society will enjoy significant improvement in their social situations. For example, those who are currently left out by society because of their gender, sex, race, or attractiveness can expect to benefit from a duty not to exclude. Although, under this scenario, there may still be some people left without friends or other close relationships, recognizing a duty not to exclude, without also recognizing a duty to make friends, preserves more of our associative freedoms than the proposals made by Brownlee and Collins.

Contributing to public shaming can violate the duty not to exclude because it can constitute a form of exclusion. Recall that one form of exclusion involves an agent communicating that they have ruled someone out from a potential association. When an agent excludes another in this way, she communicates that she has a disposition not to consider the other agent as a candidate for those associations. Public shaming communicates this disposition because of the conventions

⁹⁷ Collins also endorses a similar claim, that ““each human has a right that each agent (i) deliberates equitably about the agent’s intimacy inclinations and (ii) act as those deliberations demand”, in Collins, “A Human Right to Relationships,” in *Being Social: the Philosophy of Human Rights*, eds. Kimberly Brownlee, David Jenkins, and Adam Neal (Oxford: Oxford University Press, 2022).

⁹⁸ The circumstances under which we might be required to treat someone as a candidate for association depend on the type of association, and other contextual factors. For example, in formal associations, like a professional position, we are required to treat someone as a candidate as long as they submit a formal application which meets some basic criteria. In less formal associations, like friendship, there may be a set of not-neatly-delineated factors, like spatial proximity, frequency of contact, or membership within more formal groups (like being members of the same social club).

surrounding public shaming in modern Western societies.⁹⁹ If I publicly shame someone, my act carries the implicit message: “I think what you did was messed up – enough so that I’m willing to tarnish your reputation publicly. Associating with you from now on would only make my criticism seem insincere. So unless certain conditions are met, don’t bother trying to get in touch with me”. On the flip side, it would be very surprising and discordant if I publicly shamed a colleague, before casually asking them to get coffee with me some time later. What explains these observations? When a person publicly shames another person, they usually communicate a severing of ties and goodwill. This is a qualified claim. A parent who is so frustrated with her misbehaving child may scold her child in public. Or I may publicly call out a friend on a public platform for some bad behaviour, if I believe it is the only way to get her to change her behaviour. But in most cases, to publicly criticize someone is to foreseeably contribute to some pain, shame, and humiliation. And most of the time, we are unwilling to subject others to such outcomes unless we bear ill feelings toward them, and no longer want to have a close relationship with them. Thus, we can expect that when I contribute to public shaming, I am severing ties and goodwill with the target. Note that I do not need to intend to communicate this kind of message in order to communicate it, just as I do not need to intend to insult someone in order to insult them. Rather, because public criticism usually conveys ill feelings, we reasonably interpret public criticism as a communication of ill feelings, and expect that in most cases when people publicly criticize each other, they bear ill feelings.

The claim is that contributions to public shaming, via convention in modern Western societies, expresses exclusion. But to be clear, those contributions are not necessarily exclusive in

⁹⁹ I use the word “conventions” roughly in the way David Lewis has defined it: as a kind of behaviour which is recurrent and widely conformed with, supported by expectations that people will conform to the kind of behaviour. See Lewis, David, *Convention: A Philosophical Study* (Cambridge: Harvard University Press, 1969).

the way I have talked about. In principle, there is nothing exclusive about criticism, or anger, or disappointment. We often criticize and blame loved ones, sometimes to their faces. But we don't usually take these as excluding, at least when they occur in private. We don't take them to mean that the blamer intends to cut the blamee off from the relationship, exclude them from further opportunities, or disengage with them. Quite the opposite. Sometimes, when we blame and criticize loved ones, our criticism exhorts them to engage with us in a better way. We call on them to make amends for their mistakes, and work to repair the relationship. Public shaming could in principle be non-exclusive, just like private criticism sometimes is. John Braithwaite, for example, advocates for "reintegrative shaming" as an alternative to our current methods of criminal punishment.¹⁰⁰ In fact, Braithwaite points to societies like Japan, in which reintegrative shaming is supposedly practised. If reintegrative shaming is conceptually and practically possible, then we don't necessarily communicate exclusion when we contribute to public shaming.

V. Conclusion

Let us round up this chapter. In it, I defended two broad reasons for why humans have a right against being publicly shamed. The first broad reason involves the imposition of harms. Most salient of these harms is the health impact associated with humiliation, isolation, and the loss of valuable social goods. But public shaming can also cause the loss of esteem, as well as a person's ability to personate. I argued that we have duties not to impose these harms on others. Public shaming can also cause targets to lose control over their social identities, but I argued that this consideration can only generate duties for some agents under specific situations. The second broad reason involves exclusion, which I define as forming a disposition to rule someone out of a potential association, and acting on this disposition in a relevant way. I argued, based on the

¹⁰⁰ Braithwaite, John, *Crime, Shame, and Reintegration*, (Cambridge: Cambridge University Press, 1989).

associative needs of individuals, that we have duties not to exclude people for no reason or for irrelevant reasons. When we contribute to public shaming, we engage in an act which excludes others, and unless we possess relevant reasons to publicly shame, we wrong them.

Together, these two broad reasons support the claim that human beings have rights against being publicly shamed. That said, these rights are defeasible. This sets up another, perhaps more interesting question: what, if anything can justify public shaming? I shall take this question up in the next two chapters. In the next chapter, I explore various dimensions of one of the more common answers to this question: that public shaming is justified as a form of norm enforcement. And in the following chapter, I make my own significant contribution to this discussion by introducing and defending the idea of public shaming as moral self-defence.

Chapter III: Justifying Public Shaming – Public Shaming as Norm Enforcement

I. Introduction

Public shaming is a (a) collective activity in which (b) each contributing agent expresses criticism, blame, resentment, disgust, anger, contempt, or other negative attitudes towards another person (c) in a public forum (d) in response to a perceived violation of some norm. As we saw in the last chapter, human beings have rights against being publicly shamed, because they have vital interests in things that can be compromised by public shaming. People who have been publicly shamed can suffer a number of serious mental and physical health effects, lose the esteem of their peers, important friendships, professional opportunities, other valuable associations, and lose the ability to personate. In addition, the experience of being publicly shamed, which often involves feeling the emotion of shame, is intrinsically unpleasant. Further, we have a duty not to exclude anyone for no good reason from valuable social opportunities that we may offer. This involves a duty not to communicate to others that you have ruled them out of the relevant opportunities without relevant reasons. Public shaming communicates just that.

But there are some cases where mass public criticism or condemnation are justified, perhaps even required. One such case which comes to mind is that of Harvey Weinstein, an influential film producer who was accused of 87 counts of sexual harassment and rape.¹⁰¹ Public accusations of his behaviour began in 2017 during the #Metoo movement, which culminated in

¹⁰¹ “Harvey Weinstein timeline: How the scandal has unfolded,” *BBC News*, 24 Feb 2023, <https://www.bbc.com/news/entertainment-arts-41594672>; Moniuszko, Sara M, and Cara Kelly, “Harvey Weinstein scandal: A complete list of the 87 accusers,” *USAToday*, 27 Oct 2017, <https://www.usatoday.com/story/life/people/2017/10/27/weinstein-scandal-complete-list-accusers/804663001/>.

his conviction in 2020. Many people credit his eventual conviction to the fact that he was publicly accused and criticized as part of the #Metoo movement.¹⁰² If indeed, public shaming was causally responsible for the eventual downfall of a very dangerous man, it seems that public shaming was justified, if not morally required, in this case.

So what exactly justifies public shaming? For the most part, philosophers who have written on public shaming and other forms of social punishment have pointed out that public shaming (among other social punishments) plays an important role in enforcing valuable social norms. An act enforces a norm if it induces others to follow the norm. Call this the *Norm Enforcement View* on public shaming. Paul Billingham and Tom Parr state this view most explicitly, writing that “public criticism is an important way in which social norms are upheld and reinforced”.¹⁰³ This view is echoed by Harrison Frye.¹⁰⁴ Nicolas Delon writes that shame “seeks to restore compliance with an existing norm or to promote adherence to a new norm”.¹⁰⁵ Linda Radzik, who writes not only about public shaming but informal social punishment more generally, argues that social punishment is justified as a form of moral pressure on wrongdoers to atone for their wrongdoing.¹⁰⁶ The #Metoo movement and the downfall of Harvey Weinstein exemplifies public shaming as norm enforcement very well – by publicly shaming him, members of society reinforced norms against sexual harassment and rape.

¹⁰² Donegan, Moira, “Harvey Weinstein went from untouchable to incarcerated. Thank #MeToo,” *The Guardian*, 25 Feb 2020, <https://www.theguardian.com/commentisfree/2020/feb/24/harvey-weinstein-untouchable-guilty-thank-me-too>; Grady, Constance, “Some say the Me Too movement has gone too far. The Harvey Weinstein verdict proves that’s false,” *Vox*, 24 Feb 2020, <https://www.vox.com/culture/2020/2/24/21150966/harvey-weinstein-rape-conviction-sexual-predatory-assault-me-too-too-far>.

¹⁰³ Billingham, Paul, and Tom Parr, “Enforcing Social Norms: The Morality of Public Shaming,” *European Journal of Philosophy* 28 (2020): 1000.

¹⁰⁴ Frye, Harrison, “The Problem of Public Shaming,” *The Journal of Political Philosophy* 30, no. 2 (2022): 189.

¹⁰⁵ Delon, Nicolas, “Scarlet Letters: Meat, Normality and the Power of Shaming,” *BooksandIdeas.net* (Sep 2019), <https://booksandideas.net/Scarlet-Letters.html>.

¹⁰⁶ Radzik, Linda, *The Ethics of Social Punishment* (Cambridge: Cambridge University Press), 45.

I agree that norm enforcement is an important function of public shaming, and can play a role in justifying many cases of public shaming. But the Norm Enforcement View, as presented by philosophers so far, is incomplete, because it does not explain why random, private individuals are permitted to collectively impose *serious* harms – harms associated with the deprivation of vital interests, rights, or goods – on others to enforce a norm, even very important norms. So it is unclear why those private individuals are permitted to enforce norms via public shaming, which can cause or constitute serious harm to the shamed. The Norm Enforcement View faces a challenge in offering this explanation, which I call the *authority challenge* for the view. My goals in this chapter are to clarify and improve the Norm Enforcement View by explaining how it can respond to the authority challenge. In light of this, there are three subordinate objectives for the chapter. Firstly, to explain the Norm Enforcement View and examine how public shaming enforces norms. Secondly, to present the authority challenge and explain why existing versions of the Norm Enforcement View do not offer satisfying responses to the challenge. Finally, to present my own version of the view, which is that norm enforcement can be used to enforce important norms which protect our vital interests. This leads to my conclusion that public shaming is justified in some cases where it is necessary to enforce important norms which protect our vital interests.

Here is a roadmap for this chapter. In part II, I overview a broad structure for how a contribution to public shaming, or a case of public shaming, can be justified, and lay out a set of desiderata for a theory on public shaming. In part III, I present the Norm Enforcement View, describe the authority challenge, and reject some ways the Norm Enforcement View might meet the problem. In part IV, I present my own version of the Norm Enforcement View, which claims that public shaming is justified when it enforces particular norms which protect our vital interests. Given the connection between my view and self-defence, I mobilise self-defence theories of criminal punishment, and adapt them to the issue of public shaming. I also explain why self-

defence theories can be used to help the Norm Enforcement View overcome the authority problem. Finally, in part V, I consider a number of issues frequently associated with public shaming: necessity, proportionality, and respect for individuals.

II. Justifying Public Shaming

As I shall argue, a theory which seeks to justify public shaming must do at least two things:

- a) Explain why targets of justified public shaming can be *liable* to be publicly shamed.
- b) Explain why at least some of the contributors to justified cases of public shaming (including some random, private members of society) are permitted to contribute to public shaming.

Let us start with desideratum *a*: a theory of public shaming should explain why targets of justified public shaming can be liable to be publicly shamed. For the purposes of this project, a person is liable to some harm when it is permissible to impose that harm on them, in virtue of the fact that they no longer have a claim against the harm, or when the claim has been weakened to a sufficient degree. This happens quite clearly in cases where a right is *waived*. For example, when a boxer signs up for a boxing match, they consent to being harmed in certain ways, and thus do not have a claim against being harmed in those ways. It is thus permissible for another person (their opponent) to harm them in those ways. Paradigmatically, as in this example, a right is waived when an agent freely provides consent.¹⁰⁷

Another way a person can become liable to a harm is by *forfeiting* their right against the harm. Paradigmatically, a right is forfeited when an agent behaves faultily, although Judith Thomson has argued that individuals can forfeit rights without fault.¹⁰⁸ She argues people can

¹⁰⁷ Thomson, Judith Jarvis, *The Realm of Rights* (Cambridge, Massachusetts: Harvard University Press, 1990), 361.

¹⁰⁸ *Ibid.*, 366.

forfeit rights by “letting things lie”, or by being a threat to others.¹⁰⁹ For now, let us focus on paradigmatic cases of forfeiture, which involve fault. Typically, fault implies that an agent is morally responsible for violating a norm. Obviously, morally unjustified acts violate norms, and individuals who carry out morally unjustified acts can forfeit some rights. But norm-violating behaviour can also include the violations of a range of social norms and rules that are specific to the way we set up institutions and even games. For example, if I break a rule while playing a game of chess, I forfeit my claim to keep playing.

Why is it important for a theory of public shaming to explain the liability of targets of public shaming? Broadly speaking, if a person has a right against a harm, and they are not liable to suffer that harm, then they continue to have a claim against the harm. Thus, the harm can only be justified if the person’s claim has been outweighed by some other consideration. A right is outweighed when the reasons to act in a way that would infringe on the right are greater than the reasons to act in a way that would uphold the right. For example, each of us has a right not to be killed. But in cases like the standard trolley case, where the only way to save five people is to divert a trolley to a different track, thus killing one person, it seems permissible to divert the trolley.¹¹⁰ The reasons in favour of saving five people in this case outweigh the reasons not to kill the one person. When a right is outweighed, the right-holder still has a claim that the right be fulfilled. For example, the one person who is killed in the trolley problem still has a claim against being killed, even though this claim is outweighed by the claims held by the five people.

A theory which merely explains how some consideration can outweigh the right against being publicly shamed may be correct, but it will not be very insightful. For it is plausible that in sufficiently large quantities, a large variety of moral considerations can outweigh the right against

¹⁰⁹ *Ibid.*, 361.

¹¹⁰ Thomson, Judith Jarvis, “The Trolley Problem,” *The Yale Law Journal* 94, no. 6 (1985): 1395.

being publicly shamed (and perhaps any right in general). For example, suppose that a powerful wizard promises us that if we thoroughly humiliate, shame, and ostracize an innocent person (who has invoked her wrath), she will cast a spell that ends poverty forever. We could be permitted to do as the wizard asks. Or imagine that by publicly shaming one innocent person, everyone in the world will become extremely happy for the rest of their lives. Again, public shaming in this case could be permissible, though it would otherwise violate the right. These examples reveal that it is undesirable for a theory of public shaming to simply point out how a person's right against public shaming can be outweighed. Firstly, such a theory by itself would fail to pick out a feature of culpable wrongdoers which makes it more justifiable to publicly shame them than innocent people. The theory would have to be supplemented by some other claim (eg. that culpable wrongdoers forfeit their rights) in order to explain why public shaming is more justifiable against a culpable wrongdoer than an innocent person. Secondly, it is unsatisfying for a theory to show that some consideration, in sufficient quantities, outweighs the right against being publicly shamed. Given that a very large list of considerations can outweigh the right against being publicly shamed, it would not be very insightful for a theory to point out a particular item on that list (unless it is surprising for the item to appear on that list in the first place).

Let us now move on to the second desideratum, desideratum *b*: explaining why some private members of society are permitted to infringe on the right against public shaming. Rights are held against duty-bearers. When someone is liable to be harmed, they can be liable with respect to some duty-bearers, but not others. For example, if I enter a boxing ring to fight with George, I waive my right against harm with respect to George. But I still have a right against everyone else that they do not harm me. If a random passer-by jumped into the ring and punched me, he would act unjustifiably. This brings us to the second desideratum; a theory justifying public shaming must tell us why the particular agents carrying out public shaming are justified in acting as they do. As I

established in the first chapter, public shaming is constituted by the contributions of many individuals. In most cases, these individuals will be random, private individuals who may have no connection to the wrongdoer. Thus, a theory of public shaming must tell us why at least some of those random private individuals can have the right to contribute to public shaming.

You may suspect that this second desideratum is redundant. A theory which explains how a person can be liable to some harm might also explain who they are liable *with respect to*. This may be true. But as we shall see in the next part, the fact that a person is liable to harm (because they have forfeited their right against the harm) does not necessarily settle *who* may impose the harm. Therefore, theories which explain how a person can be liable do not necessarily provide any explanation for why private members of society may contribute to public shaming.

III. The Norm Enforcement View and the Authority Challenge

I shall now explain the Norm Enforcement View of public shaming, as well as the authority challenge. Public shaming, along with other forms of social punishments, like criticizing, blaming, shunning, expressing anger, etc., are often used to enforce social norms. A norm, for our purposes, refers to a principle that has corresponding normative attitudes among sufficient members of a social group.¹¹¹ For example, in Australia, there is a norm against smoking indoors, because many Australians have normative expectations (that is, beliefs about what others should do) that others will not smoke indoors. A *social norm* is a norm that is not constituted by formal rules, like laws.¹¹² Public shaming enforces norms via a number of mechanisms.¹¹³ Because public shaming typically

¹¹¹ Geoffrey Brennan, Lina Eriksson, Robert E. Goodin, and Nicholas Southwood, *Explaining Norms* (Oxford: Oxford University Press, 2013), 29.

¹¹² See Brennan et al. for a more complete discussion on formal and non-formal norms. Note that within the category of non-formal norms, Brennan et al. make a distinction between moral norms (eg. don't kill) and social norms (eg. dry your hands before shaking someone else's). I've chosen to use "social norms" to refer to all non-formal norms, because this is consistent with how Billingham and Parr use the term in "Enforcing Social Norms".

¹¹³ Paul Billingham and Tom Parr, "Enforcing social norms: The morality of public shaming".

imposes a hefty cost on perceived norm-violators, most people are disincentivized from violating norms. Public shaming thus deters people from violating norms. Call this the *deterrence* function of public shaming. Public shaming is also an expressive act whereby members of a community express their belief or commitment to a norm by criticizing someone for violating it. Thus, public shaming can be used by a group to express and reinforce their commitment to a norm. Call this the *expressive* function of public shaming. Relatedly, public shaming can also constitute a form of communication to norm-violators. When members of a community publicly criticize someone for a norm violation, they let the target know that they do not condone the violation, potentially changing the behaviour of the target. Call this the *communicative* function of public shaming.

The Norm Enforcement View is thus the view that the value of enforcing some social norms can justify the use of public shaming and the harms that come with it. How then does the Norm Enforcement View fare with respect to the desiderata set out in the last section? While proponents of the Norm Enforcement View have not themselves spent much time addressing these desiderata, we might think that the Norm Enforcement View can be combined with existing theories on punishment and rights forfeiture to tell us why targets can become liable to be publicly shamed in some cases. Let us consider some views on rights forfeiture. John Simmons, for instance, offers one interpretation of Locke's writings, that "Since acting wrongly demonstrates some measure of irrationality (on Locke's view), it involves a loss of some part of the rights enjoyed by the fully rational," although Simmons points out that wrongdoing is "neither a necessary nor a sufficient condition for lack of full rational".¹¹⁴ Simmons' own preferred view is that it would be unfair to give a person the protections and privileges "under rules he refuses to live by".¹¹⁵ Meanwhile, retributivists believe that committing wrongdoing can make a wrongdoer deserving of

¹¹⁴ Simmons, John, "Locke and the Right to Punish," *Philosophy & Public Affairs* 20, no. 4 (1991): 335.

¹¹⁵ *Ibid.*

punishment (with “strong” retributivists claiming that desert is necessary and sufficient for punishment, while “weak” retributivists believe that desert is only necessary for justifying punishment).¹¹⁶ Being deserving of punishment could cause one to forfeit their rights against punishment. On the other hand, Christopher Morris, building on a contractarian theory of morality, claims that our moral standing depends on our willingness to “impose constraints” on our own behaviour, and that “certain humans unwilling to impose constraints on their conduct toward others - will have partial or no moral standing”.¹¹⁷ Because humans unwilling to impose constraints on their own conduct lose some or all of their moral standing, they lose some of their claims against others, thus potentially forfeiting rights against being harmed.

We can apply these theories to public shaming. For example, if Locke is correct that acting wrongly demonstrates irrationality, and in turn that such behaviour causes the wrongdoer to lose some of their rights, then it is possible that wrongdoers can lose their rights against being publicly shamed. Or suppose that retributivists are right, and that being deserving of punishment makes one liable to be punished. We could plug this claim into a theory about public shaming, and state that at least some wrongdoers deserve to be publicly shamed, and thus forfeit their rights against public shaming. Or suppose that Morris is correct that human beings unwilling to impose constraints on their own behaviour have partial or no moral standing. Again, if this is true, it is possible that wrongdoers can lose their right against being publicly shamed because they have partial or no moral standing. Thus, as long as some theories of rights-forfeiture are correct, it seems that the first desideratum for a theory on public shaming can be met by the Norm Enforcement View.

¹¹⁶ Alexander, Larry, Ferzan, Kimberly, and Stephen Morse, *Crime and Culpability: A Theory of Criminal Punishment* (Cambridge: Cambridge University Press, 2009), 7.

¹¹⁷ Morris, Christopher, “Punishment and Loss of Moral Standing,” *Canadian Journal of Philosophy* 21, no. 1 (1991): 62.

The challenge for the Norm Enforcement View is in meeting the second desideratum for a theory about public shaming – explaining why some ordinary, private individuals are collectively permitted to impose the harms associated with public shaming. Let us say that an agent has *authority* over another person if they are permitted to impose harm on the other person in order to enforce a norm regarding how individuals should behave. It is plausible that ordinary people have some authority over one another. We are sometimes permitted to impose small costs on one another in order to enforce norms. For example, in criticizing a colleague for eating noisily in our shared office, I may impose some costs on her. My criticism seems justifiable, at least if the content of the criticism is correct (my colleague does indeed eat noisily) and does not impose disproportionate costs. So I have some limited authority over my colleague. But things are different when we think about the imposition of serious harms – harms that result in a lot of pain, the frustration of vital interests, or the deprivation of fundamental rights. In most cases, ordinary private agents do not have the authority to impose serious harms on one another. We normally think it is impermissible for private individuals to impose serious harms on others even in order to enforce very important and uncontroversial norms, unless doing so is necessary to prevent imminent future violations which threaten someone's vital interests. Put another way, outside of self or other-defence cases, it seems impermissible to impose serious harms to enforce norms. You may not, for instance, take it upon yourself to beat up someone who has committed sexual assault, even if you know that doing so will contribute significantly to enforcing the norm against sexual assault. Or at least, it is unclear that you may act in such a manner.

It also seems impermissible to make small contributions to serious harm in order to enforce norms (again, outside of self or other-defence cases). For example, suppose that a group of a thousand vigilantes learn that Tom has embezzled a lot of money from the government. They kidnap him, and each take turns to inflict a tiny, almost imperceptible cut on him. But together,

they impose serious suffering on him. It seems that each of them acts impermissibly. Thus, it is unclear why private individuals may collectively impose the harms associated with public shaming on a wrongdoer.

The authority challenge, therefore, is for a theory of public shaming to explain how at least some random, private agents have the authority to publicly shame liable targets. I shall now consider some potential responses to the authority challenge.

Rights-Forfeiture

Wrongdoers can forfeit their right against being harmed in certain ways, and this sometimes explains why others may harm them. Can the concept of rights forfeiture explain why public shaming is permissible? I think not. The concept of rights-forfeiture, *by itself*, does not explain why particular agents, much less random private members of the public, have the authority to impose harm. To see why, let us begin by observing that rights can be forfeited with respect to some particular agents, but not others. The fact that a person has forfeited a right does not entail that all other agents may act in a way that would otherwise infringe on the right. When we say that a wrongdoer has forfeited a right, we just mean that there is *some* agent who is now permitted to act in a way that would infringe on the right. Consider cases of self-defence or other-defence. Many philosophers agree that a person can forfeit their right against harm by culpably posing a threat to the vital interests of others. But an aggressor does not forfeit their right against harm with respect to everyone. For example:

Impotent Defence: You see Lawrence attack Michael with intent to cause serious harm. You have a weapon which you can use to harm and incapacitate Lawrence. But your weapon involves a slow-acting ingredient, such that by the time it incapacitates Lawrence, Michael will already have suffered harm.

As far as the goal of defending Michael is concerned, you may not harm Lawrence, because you cannot in fact defend Michael. More generally, in self-defence or other-defence cases, when an aggressor forfeits their right against harm, they forfeit it with respect to specific agents: agents who, by harming the aggressor, can protect the vital interests of the would-be victim. Agents who cannot protect any vital interests by harming an aggressor are not permitted to do so. Thus, the structure of rights-forfeiture in cases of self or other-defence is such that rights are forfeited only with respect to specific agents in certain contexts.

What about rights forfeiture in other kinds of cases, like in criminal punishment? Most people believe that ordinary private individuals are not permitted to enforce criminal sanctions on offenders, and that the state (assuming it is not seriously dysfunctional) has a monopoly on the right to punish. This belief has been questioned by some philosophers.¹¹⁸ My goal here is not to adjudicate between theories of punishment which support vigilantism and those do not. Rather, the two examples we have just considered show that it is not a conceptual truth that rights forfeiture entails rights forfeiture with respect to all. Or, just because a person has forfeited a right, does not mean that the person has forfeited the right with respect to everyone.

To be clear, my argument here is not that we cannot supply an argument for why wrongdoers may forfeit rights with respect to everyone in their community. Indeed, my own view, which I supply later, does imply that wrongdoers can sometimes forfeit their rights against public shaming with respect to most members of their community. Rather, the claim is that the concept of rights forfeiture, by itself, does not support the claim that all members of a moral community may impose harm on a wrongdoer in virtue of the wrongdoer's behaviour. If we want to say that

¹¹⁸ As Christopher Wellman points out, "given that the permissibility of vigilantism is none other than the fundamental debate between anarchism and statism, it seems to me that we should be suspicious of any theory of punishment that did antecedently rule out anarchism", in "The Rights Forfeiture Theory of Punishment," *Ethics* 122, no. 2 (2012): 371-393.

someone forfeits a right with respect to everyone, or with respect to a large portion of their community (as we must to justify public shaming), we need to supply an independent explanation for why this is the case.

Arguments from Analogies

Another set of arguments involves pointing out that random, private individuals are sometimes permitted to impose serious harms on each other in the defence of norms. Such cases of harming, if analogous to public shaming, could show that random, private individuals do have the authority to contribute to public shaming. Here are such cases. It seems permissible to participate in a boycott of a company, to force it to change its practices in a way that will meet some moral norm. In doing so, we can impose some significant harm on the company's employees. They might suffer pay cuts, lose their jobs, or suffer some negative publicity (and all the ills that come with it). It also seems clearly permissible to articulate public criticism of individuals when the criticism is an important part of a set of practices. For example, criticism of artistic performances, or the criticism of a politician's moral character. These kinds of criticism can be seriously damaging to an individual's reputation but are important parts of broader practices. Artistic criticism is an important practice within the broader practice of art, and political criticism is an important practice within democratic systems.

I think there are relevant differences between the cases that we just discussed and public shaming. Let us start with participating in boycotts, which are actually a bit more controversial than they may seem at first (I shall set aside the issue of organizing boycotts and campaigns, which may be morally more complicated than merely participating in one).¹¹⁹ Setting aside these controversies, participating in boycotts seems permitted because consumers generally have a

¹¹⁹ See Goodin, Robert E., "Acting in Combination," *Philosophy & Public Affairs* 45, no. 2 (2017): 158- 194.

prerogative over their purchases. No consumer has a moral obligation to a company or its members to purchase a product. A consumer may choose not to purchase a product for no reason at all. And the fact that someone else has organized a boycott doesn't seem to give you a duty to purchase a product, when you previously had none. One reason for this is because refusing to purchase a product is not so much an imposition of harm, as it is a refusal to confer a benefit. Generally, it is accepted that we have (defeasible) duties not to cause harm. But we do not have duties to confer benefits, except in cases where we have special relationships with the beneficiaries (like our children), or in cases where the potential beneficiaries are in serious need.¹²⁰ Thus, an employee who suffers a wage cut cannot blame a consumer for not patronizing her company, or call on the consumer to justify her actions. If you may refuse to buy a product for no reason, you can hardly be faulted for refusing to buy that product in order to defend an important norm. On the other hand, we cannot make a similar set of claims for public shaming. This is because when you criticize someone and they suffer as a result, you impose a harm. Much more is this the case if you criticize or blame someone publicly, in a way that, together with others, imposes a serious cost on them. You do not have a prerogative to cause harm. Thus, you do not have a prerogative over your public criticisms which contribute to harm, in the same way that you have a prerogative over your consumer choices.

What about cases like criticizing artists and politicians? In those cases, it seems that individuals waive any rights they have against certain kinds of criticism, by partaking in particular activities. As we noted earlier, paradigmatically, a right is waived when the right-holder freely

¹²⁰ For further discussions on the difference between harming and failing to confer benefits, see (among many articles) Shiffrin, Seana, "Harm and its Moral Significance," *Legal Theory*, Volume 18, Special Issue 3: Tort Theory (2012): 357 – 398; Purves, Duncan, "Harming as making worse off," *Philosophy Studies* 176 (2019): 2629–2656; Klockseim, Justin, "Harm, Failing to Benefit, and the Counterfactual Comparative Account," *Utilitas* 34, no. 4 (2022): 428 – 444.

provides consent to another agent to act in a way that would otherwise infringe on that right.¹²¹ But a right can also be waived when an agent undertakes a voluntary act under certain conditions, without intending to waive the right.¹²² For example, if I post a private photo of myself on the internet, I waive my right against others looking at that photo, even if I did not intend to allow them to look at it. One explanation for why I waive my right in this case is the reasonable expectation that a certain behaviour can result in what would otherwise constitute an infringement of the right (although as we shall shortly see, this is not a sufficient condition for waiver). Others might justify looking at my photo, by saying something along the lines of “you could’ve and should’ve known better”. In this vein, it is natural to think that people can waive rights they have against public criticism when they engage in relevant public activities, like creating and publicizing art. When people engage in these sorts of activities, they can be expected to know that being subject to critique is an important and generally accepted part of the practice. These waivers are usually domain-restricted in some way. While it’s perfectly acceptable to criticize a musician for their musical skills, it may be wrong to criticize them for their looks. People who choose to become politicians might waive rights against more general forms of criticism, both of their competence and their moral character (within limits). This is because it is an important part of the practice of democracy that we subject our elected officials to rigorous scrutiny, not just for their political skills, but for other aspects of their character.

Do targets of public shaming waive their rights against being publicly shamed? In some cases, yes. I have already noted that individuals can waive their right against public criticism by undertaking certain projects. When many people publicly criticize a person for some failing or norm violation, this constitutes public shaming. But there could be many cases of public shaming

¹²¹ Thomson, Judith Jarvis, *The Realm of Rights*, 350.

¹²² Thomson, Judith Jarvis, “The Right to Privacy,” *Philosophy & Public Affairs* 4, no. 4 (1975): 302.

which might not be justified by the concept of rights-waiver, because the target could not be expected to know that they would be shamed. Imagine, for example, that Jon, a public figure, privately makes some racist comments to his friends. Unbeknownst to him, those comments get recorded, and released on social media. As a result, he gets publicly shamed. Arguably, public figures should be publicly shamed when they make racist comments. But given that Jon didn't know about the recording (nor could he be expected to know), it is implausible that he waived his right against being publicly shamed simply by making those racist comments.

More importantly, it also seems that there are rights which are so important that they simply cannot be waived unintentionally (there is a separate question of whether they can be waived intentionally or forfeited). These rights include our rights to life, our bodily autonomy, and basic conditions for human wellbeing like security, food, water, shelter, etc. For example, suppose there's an alley near my home called "Robbery Alley", infamous for being patrolled by robbers. Even though I can easily avoid Robbery Alley, I take my chances anyway and walk through it. If a robber robs me, they still violate my right against being robbed. Or again, it seems impossible to waive your right to your life without intending to. You do not waive your right against being killed simply by foolishly entering an area known to be rife with murderers. The more important a right is, the stronger the reasons have to be offered in order to justify the infringement of the right. The concept of waiver is not a strong enough justification for these purposes.

If there are rights that cannot be waived unintentionally, then I think many cases of public shaming cannot be justified via the concept of rights waiver. As I argued in the last chapter, public shaming can set back some very important interests, like our social needs, health, our ability to personate, and against being humiliated. These interests are in my view no less important than many of our interests against being physically harmed. I also argued that these interests ground rights against being publicly shamed. Assuming that the strength or importance of a right is a

function of the importance of the interests they protect, then our rights against public shaming can be very strong indeed. Some kinds of public shaming, like those restricted to specific domains like artistic ability, usually do not cause humiliation or prevent a person from making friends. But other kinds of public shaming, particularly those that attack a person's moral character, can cause ostracization and stigmatization. Because the right against being publicly shamed in such ways is very strong, a person cannot unintentionally waive their right against being humiliated or being stigmatized by their community, any more than we can unintentionally waive our rights against physical harm. Thus, public shaming which is likely to stigmatize a person cannot be justified by the notion of unintentional waiver.

Conceptual Arguments

Some have suggested conceptual links between norms and enforcement, thus implying that agents may participate in things like public shaming just because they are enforcing a morally authoritative norm. According to Harrison Frye, “a social norm confers upon individuals the standing to act in the name of enforcing that norm”.¹²³ For him, this is because of a conceptual relationship between norms and enforcement.¹²⁴ Drawing on the work of Stephen Darwall, he calls this standing “representative authority”.¹²⁵ Frye seems to lean on Darwall's arguments on representative authority here, where the latter claims that there are “conceptual relations” between morality itself and “authoritative claims and demands that must be able to be addressed to those to whom they apply”.¹²⁶

I am sceptical of Frye's claim that social norms confer on individuals the right to enforce norms, at least via the imposition of serious harm. For it is one thing to make an authoritative

¹²³ Frye, Harrison, “The Problem of Public Shaming,” *The Journal of Political Philosophy* 30, no. 2 (2022): 189.

¹²⁴ *Ibid.*, 192-193.

¹²⁵ *Ibid.*, 189.

¹²⁶ Darwall, Stephen, *Morality, Authority, and Law* (Oxford: Oxford University Press, 2013), 135.

claim or demand on another person. It is another to enforce that claim or demand via the imposition of serious harm. Darwall himself does not claim that having the representative authority to make a demand of another person entails the right to enforce norms via imposing harm. He writes that “Someone has practical authority with respect to another if, and only if, the latter has a second-personal reason to comply with the former’s valid claims and demands and is responsible to the former for so doing”.¹²⁷ Second-personal reasons, according to Darwall, are roughly reasons to act that are grounded in what a person owes to another (as opposed to reasons grounded in bringing about a better state of affairs).¹²⁸ That is to say, if I have a second personal reason to perform some kind of act for you, you hold a right against me, that I perform that act. Thus, to say that you have the representative authority to make a demand is just to say that you hold a right against someone that the demand is complied with. It does not say anything about what you may do in order to ensure that the demand is met.

A related idea is the conceptual link between moral norms and “negative reactive attitudes”, like blame, resentment, moral anger, disgust, etc. P.F. Strawson, for example, famously argues that having negative reactive attitudes towards wrongdoers is constitutive of treating them as morally responsible. As he writes, resenting someone is “the consequence of continuing to view him as a member of the moral community; only as one who has offended against its demands”.¹²⁹ On the Strawsonian view, to say that an act is wrong is to say that it is deserving of some negative reactive attitudes. Since public shaming is the combination of many people publicly expressing attitudes like blame or resentment, could a Strawsonian view justify public shaming?

¹²⁷ *Ibid.*, 141.

¹²⁸ *Ibid.*, 136.

¹²⁹ Strawson, P.F. “Freedom and Resentment,” *Proceedings of the British Academy* 48 (1962), retrieved 30 Nov 2023, <https://www.ucl.ac.uk/~uctytho/dfwstrawson1.htm>.

I think not. There is a gap between the expression of negative reactive attitudes and public shaming. For the expression of negative attitudes can be done privately, without risking or contributing to the harms caused by public shaming. This point has been made by Radzik, who notes:

Since I have argued that rebukes are typically punishments, it is a pretty good justification of at least this method of punishment. But notice how limited this victory is. The account defends private rebukes but not public rebukes. After all, public rebukes intentionally impose further harms, such as damage to reputation and the exposure of the wrongdoer to the shaming gaze of an audience. These go beyond the harm that is inseparable from the communication of moral anger to the wrongdoer. The argument also falls short of defending practices of pointed social avoidance, which purposefully withdraws the benefits of cooperation and social contact, and boycotting, which directly targets reputation and revenue.¹³⁰

The key point then, is that just because expressing negative reactive attitudes (rebuking) is an appropriate and justified response to wrongdoing, doesn't mean that all kinds of rebuking are automatically justified. Some forms of rebuke, like public rebuke, impose more harm than others. The Strawsonian argument does not entail that particularly harmful forms of rebuke are justified.

A final idea is that there is a link between morally authoritative norms and our (privilege) right to criticize. We often say that we have reasons to hold wrongdoers "to account", or to hold each other "to account". Holding someone to account involves the imposition of tangible sanctions, like humiliation, or intangible sanctions like disapproval (I discuss this further in chapter V). Among such sanctions, there is criticism. Nicholas Southwood, in discussing the notion of

¹³⁰ Radzik, Linda, *The Ethics of Social Punishment* (Cambridge: Cambridge University Press, 2020), 33.

feasibility, writes that “the constitutive aim of prospective criticism is simply to hold one another (and ourselves) accountable for how we act”.¹³¹ The role of criticism here is not just to deter bad behaviour, nor to get a wrongdoer to amend his ways. Holding someone accountable “requires precisely holding firm and undertaking to criticize him regardless of his inability to bring himself” to do the right thing.¹³² It might thus be thought that we have a right to criticize wrongdoers grounded in our reasons to hold wrongdoers to account. However, this idea does not automatically justify public shaming. Again, even when we accept that we have reasons to criticize a wrongdoer, it doesn’t follow that we are justified in criticizing in public, or criticizing in ways that result in particularly significant harm.

Here’s the takeaway from the last few paragraphs: there is a dilemma for Frye’s claim that “a social norm confers upon individuals the standing” to enforce that norm. On one hand, “standing to enforce norms” can refer to something relatively weak, like the Darwallian kind of practical authority, the right to express negative reactive attitudes, or the right to express criticism. But if that is the case, then it is not clear that social norms confer on individuals the right to collectively impose serious harms in order to enforce norms. Just because you may demand something of someone, doesn’t mean you may impose serious harms on them to enforce those demands. Because public shaming imposes serious harms, this line of thought does not justify public shaming. On the other hand, “standing to enforce norms” could refer to the right to enforce norms via the imposition of serious harm. But if that is the case, then it is not clear why we should believe Frye’s claim. The mere fact that something is a social norm doesn’t seem to be the sort of thing that can give individuals the permission to impose harm on each other. As I have argued,

¹³¹ Southwood, Nicholas, ‘Does “Ought” Imply “Feasible”?’ *Philosophy & Public Affairs* 44, no. 1 (2016): 41.

¹³² *Ibid.*, 42.

concepts like representative authority, reactive attitudes, or holding others to account do not support the claim that private individuals may enforce norms via the imposition of serious harm.

Now for a slightly concessive point. Social norms can play roles in explaining why some people have a right to harmfully enforce a norm, although these roles are secondary to some other justification for the harmful enforcement of norms in the first place. One such role that social norms play is expectation setting. For example, there is a social norm against dialogue in movies that is considered “on-the-nose”. A director who has many lines which are too “on-the-nose” may be publicly criticized for her work, and may even lose her career for particularly bad dialogue. Now, the mere fact that there is a norm against “on-the-nose” dialogue isn’t what justifies public criticism in this case. Rather, the director waives her right against public criticism by creating the movie. And the social norm explains why she waives her right in this instance – if the social norm is widely accepted, the director can be expected to know that she could be liable to public criticism if she makes a movie with “on-the-nose” dialogue. Thus, social norms set expectations, which can sometimes explain why individuals waive their rights against public criticism in some cases.

Another role that social norms play is in coordinating punishment. Take an important norm, like the norm against murder. Clearly, when an agent commits murder, *someone* is permitted to punish the murderer (for independent reasons, eg. the murderer deserves to be punished). If each person were to take it upon themselves to enforce this norm, it could lead to disastrous results. To avoid disaster, we might have legal rules to establish who may enforce the norm, like rules against vigilantism, and rules permitting the state to enforce the norm. We might also have social norms to fulfil this role. For example, in a pre-legal society, there may be a norm that only family members of the victim may enact retribution on the murderer. In these cases, social norms identify specific agents as the executors of punishment. These social norms are in turn grounded in independent normative claims, such as:

- It would be bad if all agents regarded themselves as permitted to punish a wrongdoer.
- There is some agent (eg. the state) who can better achieve the goals of punishment than other agents.¹³³

The point then, is that while social norms do sometimes play a role in explaining why certain agents may enforce norms, these explanations are dependent on other normative claims. It is insufficient just to say that because X is a social norm, all members of a community may impose harm in order to enforce X.

IV. Self-Defence and the Authority to Enforce Norms

I think that at least a part of what justifies norm enforcement is our right to protect ourselves. Within certain constraints,¹³⁴ any agent is permitted to defend themselves against some unjustified threats to their vital interests. Agents are also permitted to protect others, especially those who are incapable of protecting themselves. While philosophers disagree about why defensive harming is justified, and when exactly it is justified, it is uncontroversial (by philosophical standards) that most individuals are sometimes permitted to harm others in self-defence or other-defence.

My approach to answering the authority challenge for public shaming is to claim that public shaming can sometimes function, and be justified, as a form of collective self-defence. This claim is similar to the claims made by self-defence theories of criminal punishment, which claim that criminal punishment is a form of collective self-defence. In developing my view, I will accept and

¹³³ Christopher Wellman makes an argument to this effect in order to justify the state's monopoly on punishment, arguing that the goals of punishment "stand to be adequately realized only in a system with an authoritative body which makes, enforces, and impartially adjudicates promulgated rules", in "Rights and State Punishment," *The Journal of Philosophy* 106, no. 8 (2009): 428.

¹³⁴ For a list of conditions of justified self-defence, see Seth Lazar, "Necessity in Self-Defense and War," *Philosophy and Public Affairs* 40, no. 1 (2012): 3-44.

extend some arguments and claims made by self-defence theories of punishment. My aim in this section is not to defend those arguments, but to show that they can apply to cases of public shaming. Thus, I acknowledge that my account is not entirely neutral with respect to the more established debate on what justifies criminal punishment. That said, my view does not necessarily entail that self-defence theories of punishment correctly explain and justify all instances of criminal punishment. There may be moral considerations associated with cases of criminal punishment that do not similarly apply to cases of public shaming, and vice versa. It is possible for the arguments in this part to succeed in explaining and justifying some cases of public shaming, while failing to explain and justify the institutions of criminal punishment. Thus, my claims in this part do not require that self-defence theories of punishment succeed, only that some of the arguments within those theories be true.

Self-defence theories of criminal punishment make the same claims made by other theories of punishment – that punishment can reduce the likelihood of future wrongdoing either by deterring wrongdoing, expressing society’s commitment to wrongdoing, communicating with wrongdoers, reforming them, or some combination of these approaches. As Wellman points out, these other theories of punishment identify the aim of punishment, but do not explain why particular agents are justified in carrying out the punishment.¹³⁵ Self-defence theories take the step in explaining why agents have the moral right to carry out this punishment. They assert that by reducing the likelihood of future wrongdoing, punishers take a necessary step to protect their own interests. The basic argument, therefore, looks something like this:

1. Agents have the right to take necessary and proportionate steps to defend their vital interests against unjustified threats.

¹³⁵ Wellman, Christopher, “Rights Forfeiture Theory of Punishment,” *Ethics* 122, no. 2 (2012): 371-393.

2. Punishing wrongdoing is (sometimes) necessary and proportionate to protect one's vital interests against unjustified threats.

C. Agents (sometimes) have the right to punish wrongdoers.

It is worth noting a limitation to this argument. It only justifies punishment in cases where vital interests are under threat. On self-defence theories of punishment, we are justified in punishing murderers, rapists, violent offenders, perhaps some trespassers, and perhaps those who violate important coordination norms, like norms regarding speed limits. Self-defence theories do not justify punishing someone for talking loudly in church, swearing during a memorial service, or failing to pay parking fines.

I take premise 1 in the above argument to be true. Premise 2 also seems very plausible – most people will agree that punishment of some kind is sometimes necessary to reduce violent crime to levels that we are willing to live with. The potential problem with the argument is that it is invalid, because of an equivocation about the term “unjustified threats”. The “unjustified threat” in premise 1 refers to agents who are liable to be harmed. In most theories of self-defence, a person is liable to be harmed in self-defence only if they pose a threat to you *and* if they meet some further condition, like having forfeited a right. Some philosophers think that an aggressor forfeits their rights against defensive harms merely by posing a present threat to others.¹³⁶ Other philosophers maintain that an aggressor must be morally responsible for posing a threat.¹³⁷ Yet other philosophers claim that an attacker is liable only if they are culpable for creating a threat.¹³⁸ In any case, for self-defence to be justified, the same person who presents a threat must also be liable to

¹³⁶ Thomson, Judith, *The Realm of Rights* 366-371; Renzo, Massimo, “Rights Forfeiture and Liability to Harm,” *The Journal of Political Philosophy* 25, no. 3 (2017): 324.

¹³⁷ Otsuka, Michael, “The Moral-Responsibility Account of Liability to Defensive Killing,” in *The Ethics of Self-Defense*, eds. Coons and Weber (New York: online edn, Oxford Academic, 2016).

¹³⁸ Ferzan, Kimberly Kessler, “Culpable Aggression: The Basis for Moral Liability to Defensive Killing,” *Ohio State Journal of Criminal Law*, 9, no. 2 (2012): 669–697.

be harmed. In contrast, when we punish past wrongdoing, we might harm someone who has previously met the conditions for liability to harm. But this person may no longer present a threat to us. So the problem is this: when we punish Aaron to prevent Bob from committing wrongdoing, we harm someone who may have met the conditions for liability, but no longer presents a threat.

This problem is a serious one for self-defence theories of punishment, and presenting a satisfying defence of the theories will take us too far afield. For now, I shall present two responses to the problem, both of which strike me as plausible. Warren Quinn proposes that the right to punish wrongdoers comes from the right to create credible threats against having our own rights being violated. As he writes, “it is because one could not object to the threat in the first place that one cannot, later, object to being punished”.¹³⁹ Quinn’s thesis may be understood as a series of premises. First, the right to self-defence entails the right to create threats against potential aggressors, if those threats are necessary or important in self-defence. Second, to create a threat is to “first, deliberately to create a real risk that x will suffer a certain evil if he does or omits a certain specified action and second, to warn x of the existence of this risk, where by these means x may possibly be deterred from the act or omission”.¹⁴⁰ Third, in order to create a real risk that an aggressor will “suffer a certain evil”, a threat has to be sincere. That is, the threat-maker has to have the intention to carry out her threat, or must at least credibly signal her intention to carry out her threat. And the best way to signal this intention is to commit oneself to carrying out the threat, perhaps via an institution of punishment which not only threatens, but punishes.

Quinn analogizes criminal punishment to a group of “fantastically complex devices that can (at least as well as we can) detect wrongdoing in our new community, identify and apprehend

¹³⁹ *Ibid.*

¹⁴⁰ Quinn, Warren, “The Right to Threaten and the Right to Punish,” *Philosophy & Public Affairs* 14, no. 4 (1985), 335.

those who are responsible, establish their guilt, and subject them to incarcerations (and perhaps other evils)".¹⁴¹ These devices punish within limits of proportionality, and are "conservatively programmed that they will never be "fooled" into punishing the innocent".¹⁴² Quinn believes that we are justified in activating these devices, and publicizing the fact that these machines are in operation. The threat that these machines pose to potential criminals, as Quinn points out, is similar to the kind of threat posed by barbed fences and one-way tire spikes. They are threats of *automatic* harm that a wrongdoer faces upon committing a wrongdoing. In this way, punishment is justified because it is a natural consequence of the threat: "our choice to bring about future punishments is derivative, an unavoidable consequence of our more basic choice to set up the threat".¹⁴³

Just as we are justified in activating the punishment machines, Quinn proposes that we are justified in threatening punishment, and then carrying it out. He defends the claim that "Every intuitively justified practice of punishment has as its counterpart a practice of [device-implemented] punishment justified by rights of self-protection, and vice versa".¹⁴⁴ In the same way that activating Quinn's machines is to automatically bring about punishments for those who ignore the threat, to be justified in threatening someone is just to be justified in carrying out that threat. As he claims: "our moral deliberations about setting up a practice of punishment we may regard the creation of the threat as if it amounted to causally determining our wills so that we would in fact try to punish crimes and would do so without raising any further question of right".¹⁴⁵ By

¹⁴¹ *Ibid.*, 337

¹⁴² *Ibid.*

¹⁴³ *Ibid.*, 339.

¹⁴⁴ *Ibid.*, 351.

¹⁴⁵ *Ibid.*, 360-361.

setting up a system of criminal punishments, we commit ourselves to carrying the threats entailed by the system.

If Quinn's theory is correct, it can be used to justify some cases of public shaming. Public shaming can be used as a kind of threat. When we, as a community, set up a practice of public shaming, we effectively say "follow this norm, or else". If a norm is sufficiently important, it may be permissible to protect that norm via some sort of threat. And we might then argue that if no one can object to the threat of being publicly shamed, they cannot object to being publicly shamed. It is worth noting however that while Quinn's theory might be able to justify some cases of public shaming, it does not cover some interesting and compelling cases of public shaming. These are cases in which public shaming is used to shape and change norms, or when public shaming occurs in the context of changing norms, instead of when public shaming enforces existing and widely accepted norms. The first few people shamed for sexual harassment, for example, were not threatened with public shaming. Being among the first, they would not have expected that they would be shamed for sexual harassment. They did not know, or believe, that there was a real risk of being publicly shamed. Thus, they were not threatened. Quinn's thesis, and any theory which justifies punishment on the basis of threats, has little to say about whether we are justified in shaming in such circumstances.

Another option of justifying punishment as self-defence is Victor Tadros' "duty view" of punishment.¹⁴⁶ According to Tadros, agents who are culpably responsible for creating harms to others have stringent, enforceable duties to neutralize those harms, or those threats. For example, if I hire a hitman to assassinate you, I may have a duty to throw myself in front of the hit man's bullet, if doing so is necessary to save you. But if this cannot be done, I have a duty to remedy this

¹⁴⁶ Tadros, Victor, *The Ends of Harm: The Moral Foundations of Criminal Law* (Oxford: Oxford University Press, 2011), 266.

wrongdoing in a way that is “reasonably similar in stringency and kind”.¹⁴⁷ Tadros illustrates this with the following example:

Evelyn hires a hit man to kill Wayne. Fred has also hired a hit man to kill Wayne. Both hit men arrive at the same time. Because of where they are standing, Wayne can only use Fred as a shield against Evelyn’s hit man and Evelyn as a shield against Fred’s hit man. He manages to do that, resulting in the deaths of Evelyn and Fred.¹⁴⁸

Intuitively, Wayne does nothing wrong. Tadros’ explanation for this is that both Fred and Evelyn have duties to stop their hit men from assassinating Wayne. But if neither are in a position to stop their own respective hit men, they still have duties to stop the others’ hit man, if doing so requires the same amount of sacrifice and effort as stopping their own hit man. Just as Fred is required to use himself as a shield to stop his own hit man from assassinating Wayne, he is required to use himself as a shield to stop Evelyn’s hit man from killing Wayne.

Further, Fred and Evelyn’s duties to save Wayne are enforceable. That is, Wayne, and perhaps other agents, may force them to fulfil their duties, even if they do not want to do so. This is not because duties are generally enforceable. Rather a duty is enforceable “if doing the duty is necessary to avert a great harm and compensation will be inadequate, or will be unlikely to be forthcoming, if the duty is breached”.¹⁴⁹ This applies in the cases of self-defence like the ones above. If we do not enforce Fred and Evelyn’s duties to protect Wayne, Wayne will die. This is a very bad outcome for Wayne. And Fred and Evelyn will be unable to compensate him, because he is dead. This principle can also be applied to cases of serious, but not lethal, harm. If Wayne is not killed by the hit men, but severely wounded, no amount of money or apology can suffice as

¹⁴⁷ *Ibid.*, 275.

¹⁴⁸ *Ibid.*, 192.

¹⁴⁹ *Ibid.*, 280.

forms of compensation. So again, compensation is impossible, or very hard to bring about. Tadros' theory may be thought of as a theory about the just distribution of costs. In cases of self-defence, there is some party that must bear some very serious costs. All else being equal, it is at least permissible for an innocent party to make a wrongdoer bear those costs, in virtue of the fact that they have a duty to bear those costs in the first place.

From here, we can see how Tadros justifies some punishment as a form of self-defence. When a criminal violates my rights, he can acquire enforceable duties to protect me from other harms. These duties are similar in kind and stringency to the duties that he would have had to prevent his actions from harming me in the first place.¹⁵⁰ One instance of such a duty involves accepting some harm in order to deter other prospective criminals from violating my rights in a similar way. Because this duty is enforceable, I may harm him in order to deter other prospective criminals from violating my rights.

Supposing that Tadros' theory is successful in justifying criminal punishment, can it be used to justify public shaming? It seems so, at least in some cases. We might claim that people who commit wrongdoing acquire enforceable duties to protect their victims from similar wrongdoing. If the wrongdoing in question is severe enough, the acquired enforceable duties can be stringent enough to require their bearers to incur costs. Under the right conditions, it may thus be permissible to impose costs on the wrongdoer in order to enforce their duties. For example, a

¹⁵⁰ The notion of similarity admits some problematic vagueness. Suppose that I shoot you, causing injury. Many years later, for unrelated reasons, you suffer kidney failure and require an organ transplant. Reasoning that I would have been required to sacrifice my life to save you from my bullet, and that this duty is similar to sacrificing myself to save you in other situations, you decide that I am now required to sacrifice my life to save you. So you kidnap me and forcibly extract my kidney. It seems that in this case you act impermissibly. So more needs to be done to narrow the notion of similarity. Yet, it cannot be narrowed too much, lest we sacrifice the goal of justifying general deterrence. We do not punish axe-murderers just to deter other axe-murders. We punish axe-murderers to deter crime in general. Thus, in order for Tadros' theory to work at justifying general deterrence, it needs to strike a balance between having a narrow enough notion of similarity to exclude implausible results, but a wide enough notion to justify general deterrence.

sexual harasser has enforceable and stringent duties to protect his victims from more sexual harassment. If imposing (proportionate) costs on him, via public shaming, is necessary to prevent his victims from being sexually harassed again by deterring potential sexual harassers, then we are permitted to publicly shame him.

Let us return to our list of desiderata for a theory of public shaming. As we noted, the first desideratum is for a theory to explain how a person can become liable to be harmed. Self-defence theories can provide this explanation. As we saw earlier, multiple theorists on self-defence have explanations for how those who constitute threats to others can forfeit their right against being harmed. These theories can be plugged into a theory on public shaming. Secondly, a theory needs to explain why random, private individuals have the authority to inflict significant harm on others. On self-defence theories, each agent has the moral right to protect themselves and others who require protecting. If an agent needs to contribute to public shaming to defend their vital interests, and if contributing to public shaming will not violate other constraints (like proportionality), then they have the right to do so.

V. Necessity, Proportionality, and Respect

Necessity

For a harmful act of self-defence to be permissible, it must be necessary. This constraint allows us to identify the kinds of cases where public shaming is more likely to be justified, and cases where it is less likely to be justified. An act of harmful self-defence is necessary if and only if it marginally lowers the expectation of the defender (or in cases of other-defence, the potential victim) being harmed by a sufficiently large amount, relative to other courses of action.¹⁵¹ For

¹⁵¹ For a more precise claim, Lazar states that:

Defensive harm H is necessary to avert unjustified threat T if and only if a reasonable agent with access to the evidence available to Defender would judge that there is no less harmful alternative,

example, suppose that someone attacks me with the intention of killing me. If I shoot him in the legs, I have a 51% chance of stopping him, and I will not kill him. If I shoot him in the chest, I have a 90% chance of stopping him, and he will die. Here, shooting my attacker in the chest may not be *strictly* necessary, in the sense that I have another course of action available to me that has a reasonable chance of saving my life. But it still seems that shooting my aggressor in the chest is necessary and permissible, because shooting him in the chest raises my chances of survival by a very significant amount, relative to the increase in costs that I will impose on the aggressor. Or suppose that someone is trying to kill me, and I can either attempt to disarm them or shoot them. In attempting to disarm them, I have an 85% chance of success, while I will have a 90% chance of success if I shoot them. Here, it seems that shooting my aggressor is not necessary, because I impose a much greater harm for the sake of an increased 5% chance of success.¹⁵²

As a result, public shaming is more likely to be necessary in cases where it is used in response to norms which are not already enforced, or poorly enforced, by existing institutions. Conversely, it is less likely to be necessary in cases where a norm is already well and reliably enforced. This is because any cost we impose on wrongdoers is likely to have diminishing marginal effects on how well a norm is enforced.¹⁵³ An additional year of a sentence is less likely to deter potential wrongdoers when the average sentence is already twenty years, than if the average sentence is only one year. We might think that a similar pattern holds for public shaming. For example, suppose that a state punishes a crime heavily and reliably, such that it deters 90% of

such that the marginal risk of morally weighted harm in H compared with that in the alternative is not justified by a countervailing marginal reduction in risk to the prospective victims of T. See Lazar, Seth, "Necessity in Self-Defense and War," *Philosophy and Public Affairs* 40, no. 1 (2012): 13. Also see McMahon, Jeff, "The Limits of Self-Defense," in *The Ethics of Self-Defence*, eds. Christian Coons and Michael Weber (New York: Oxford University Press, 2016), 188.

¹⁵² This example is adapted from Lazar, "Necessity in Self-Defense and War," 12-13.

¹⁵³ See Stemen, Don, *The Prison Paradox: More Incarceration Will Not Make Us Safer* (New York: Vera Institute of Justice, 2017); Johnson, Rucker, and Stephen Raphael, "How Much Crime Reduction Does the Marginal Prisoner Buy?" *Journal of Law and Economics*, Vol. 55, No. 2 (May 2012), pp. 275-310.

prospective criminals from committing that crime. If we were to publicly shame people for that crime, we would at best deter the remaining 10% of potential wrongdoers, no matter how severe the public shaming is. In reality, we will probably deter less than the remaining 10%. While some of the 10% may be deterred by public shaming when they were not previously deterred by the state punishment, other members of the 10% may simply not be responsive to the potential costs that society can impose on them. Given that public shaming is less likely to have a significant marginal effect in cases where norms are well enforced, any given cost imposed by public shaming is less likely to result in a sufficiently large marginal difference to justify the cost. Thus, public shaming is less likely to be justified in such cases.

However, when a norm is not enforced by legal means, or when a norm is poorly enforced by the legal system, public shaming can become easier to justify.¹⁵⁴ Consider hate speech. In a number of countries like the US, hate speech is not illegal. Therefore, informal sanctions against those who make hate speech are likely the main source of deterrence. Public shaming may constitute one of these informal sanctions. As a result, the marginal difference public shaming makes on deterring racist jokes is likely to be larger. Or consider the norm against rape. According to a survey of other studies done by the National Sexual Violence Center, 63% of rapes are not reported to law enforcement.¹⁵⁵ And even among reported rapes, a tiny fraction result in conviction.¹⁵⁶ As a result, while there is a legal norm against rape, it is very poorly enforced. Given that rapists are more likely than not to get away with their crime, current measures do not deter

¹⁵⁴ See Dumsday, Travis, “On Cheering Charles Bronson: The Ethics of Vigilantism,” *The Southern Journal of Philosophy* 47 (2009): 49-67.

¹⁵⁵ “Statistics about Sexual Violence,” *nsvrc.org*, (2012, 2013, 2015), retrieved 30 Nov 2023, https://www.nsvrc.org/sites/default/files/publications_nsvrc_factsheet_media-packet_statistics-about-sexual-violence_0.pdf

¹⁵⁶ Hohl, Katrin, “New scorecards show under 1% of reported rapes lead to conviction – criminologist explains why England’s justice system continues to fail,” *The Conversation*, 2 April 2022, <https://theconversation.com/new-scorecards-show-under-1-of-reported-rapes-lead-to-conviction-criminologist-explains-why-englands-justice-system-continues-to-fail-180345>.

rape very well. As a result, it is possible that even a modest amount of public shaming could result in significant marginal benefits in deterring potential rapists. This conclusion could mean that public shaming in rape cases is likely to be necessary to protect the vital interests of some people, given the limitations of the current legal regime.

I have said that public shaming is more likely to be justified in cases where a norm is not well or reliably enforced by existing institutions. Of course, this means that public shaming can still be justified in some cases where a norm is already well enforced, and public shaming can be unjustified in some cases where a norm is not well-enforced. As in many other issues in morality, the specifics matter. Ultimately, what is key is that a potential contributor to public shaming must take into account the likely harms of that case of public shaming, and weigh it against the marginal effect that case of public shaming will have on enforcing a norm. Additionally, the contributor must also take into account the likely effects of their own contribution (small though they may be), and how many others have contributed to the case. Contributing to a case of public shaming when millions of other people have already contributed is less likely to be necessary, because it is more likely that the case will have achieved their goals.

Proportionality

A common worry raised about public shaming, particularly online public shaming, is that of proportionality.¹⁵⁷ For Billingham and Parr, this is because public shaming – in particular, online shaming – often involves uncertainty over norms (which can mean that wrongdoers are less culpable than they may seem), uncharitable interpretation of the target’s behaviour, and large

¹⁵⁷ Russell Blackford, for instance, articulates this worry by writing that “While the Arlington photograph [a specific case where a woman photographed herself pointing her middle finger at Arlington Cemetery] might seem childish, or many other things, posing for it and posting it on Facebook hardly add up to any serious wrongdoing. It is not behaviour that merited the outcome for Lindsey Stone: destruction of her reputation, loss of her job, and a life of ongoing humiliation and fear,” in “The Shame of Public Shaming,” in *The Conversation*, 4 May 2016, <https://theconversation.com/the-shame-of-public-shaming-57584>.

numbers of disparate individuals, which result in “pile-ons”.¹⁵⁸ Frye also traces the disproportionateness of public shaming to the numbers of individuals involved in public shaming, although he suggests that disproportionateness is a symptom of a deeper problem – that public shaming is often disproportionate because of a mismatch between the scale of norm enforcement and the amount of harm that is proportionate.¹⁵⁹

We can distinguish between two ways a harm can be disproportionate in cases of norm enforcement. A harm can be disproportionate in one sense if it exceeds what a wrongdoer is liable to suffer. Call this *narrow disproportionality*.¹⁶⁰ Given the nature of online public shaming and pile-ons, it seems that many cases of public shaming are disproportionate in this sense. But a harm can also be disproportionate in a second way when it is excessive compared to the marginal value achieved in enforcing a norm. Call this *wide disproportionality*.¹⁶¹ A harm can be narrowly disproportionate, while not being widely disproportionate. For example, suppose that Max makes a racist joke, resulting in him being publicly shamed so severely that it effectively destroys his life. However, because the public shaming of Max was so severe, it effectively ends all racist comments forever. In this case, the harm imposed by the public shaming is disproportionate to what he is liable to suffer, but proportionate to the value gained in norm enforcement.

When a case of public shaming is narrowly disproportionate, but widely proportionate to the value gained in norm enforcement, it can be justified all-things-considered. It seems that the public shaming of Max is justified if it truly ends racist comments forever, although regrettable in

¹⁵⁸ Billingham and Parr, “Enforcing Social Norms: The Morality of Public Shaming”: 1008.

¹⁵⁹ Frye, Harrison, “The Problem of Public Shaming,” *The Journal of Political Philosophy* 30, no. 2 (2022): 204–205.

¹⁶⁰ Billingham and Parr, “Enforcing Social Norms: The Morality of Public Shaming”: 1001.

¹⁶¹ *Ibid.*

the outcome of destroying Max's life. Similarly, consider the case of Amy Cooper, which I rehearse here:

Amy Cooper. Amy Cooper is a white woman who threatened to call the cops on Christian Cooper, a black man who was watching birds in central park (he had asked her to leash her dog), and falsely accused him of threatening her life. Christian Cooper videoed the event and posted it online, together with reflections of his own experiences of race in America. Predictably, she was heavily punished for her actions. According to the New York Times, "internet sleuths" dug into her personal life for more damning information. She lost her job, her dog (she appears to treat the dog roughly in Christian Cooper's video), and was charged in court for lodging a false report.¹⁶²

It is important to note the context of this case. It occurred in May 2020, during the height of the Black Lives Matter movement, a response to a string of episodes of police violence against African Americans. Amy Cooper's behaviour was interpreted not just as an isolated case of injustice, but a representation of a wider, and very serious threat that Black people face.¹⁶³ It seems that while Amy Cooper did something wrong, she didn't deserve to experience the amount of harm she did. But racial injustice is also a serious problem, and violence against Black people in the US is a serious threat. Given the context of the case, it is possible that the marginal value (from spreading awareness about police violence, expressing condemnation of such violence, deterring people from acting like Amy Cooper, etc.) gained from public shaming in this case was proportionate to the harm imposed on her. Of course, it is hard in such real-life cases to make a definitive assessment of all the harms and benefits attributable to a case of public shaming, and then weigh them. But

¹⁶² Jashinsky, Emily, "Amy Cooper Doesn't Deserve Sympathy, But Social Media Shaming Is an Unhealthy Norm," *The Federalist*, 27 May 2020. <https://thefederalist.com/2020/05/27/amy-cooper-doesnt-deserve-sympathy-but-social-media-shaming-is-an-unhealthy-norm/>.

¹⁶³ Cooper, Christian, *Better Living Through Birding* (New York: Random House, 2023).

the point here is just that it is plausible that cases of public shaming are sometimes all-things-considered justified even when they are disproportionate to what a wrongdoer deserves. Thus, while the issue of proportionality should give us reasons to be very cautious of using online public shaming, I don't think we have to conclude that online public shaming should never be used.

Again, I think that cases of public shaming are more likely to be proportionate, in the sense of being proportionate to the value achieved in norm enforcement, when a norm is poorly enforced by existing institutions. As with my point on necessity, this is because of the diminishing marginal value of imposing harm in norm enforcement. If a given harm is going to deter only 1% of potential wrongdoers from committing a wrongdoing, it brings about less marginal value than if it were to deter 10% of potential wrongdoers. The harm that deters 1% of wrongdoers is thus less likely to be proportionate, all else being held equal.

Respect

A common set of worries about public shaming, or shaming in general, is that it involves marginalization or stigmatization which is incompatible with the basic respect that we owe to human beings. For example, Aitchison and Meckled-Garcia claim that characterizations of a target in (online) public shaming “aim to publicly present their identity as being not worthy of participating in (certain) normal social relations, civil conversation, or debates as a moral equal”,¹⁶⁴ and conclude that public shaming “mistreats them, stigmatising or dissuading their adoption of life goals, projects, and priorities that give shape to their lives”.¹⁶⁵ They regard this activity as incompatible with recognition respect within the context of liberal social arrangements, which does not require individuals to have a specific kind of character.¹⁶⁶ Similarly, Thomason describes

¹⁶⁴ Aitchison, Guy, and Saladin Meckled-Garcia. “Against Online Public Shaming: Ethical Problems with Mass Social Media,” *Social Theory and Practice* 47, no. 1 (2021): 7.

¹⁶⁵ *Ibid.*, 3.

¹⁶⁶ *Ibid.*

shaming as “globalizing in a morally objectionable way,” because “We take our own views of her to be the deciding factor in who she is, and we refuse to acknowledge the authority of her own point of view”.¹⁶⁷ She expresses similar worries about stigmatization, which she defines as an act which attempts to marginalize or lower the social standing of another person: “Punishments that are specifically designed to manipulate the social standing of offenders meddle in their identities in ways that threaten their abilities to live their lives”.¹⁶⁸

Contrary to Aitchison and Meckled-Garcia, I think that contributors to public shaming may not always aim at causing stigmatization or marginalization. Nevertheless, it is hard to escape the fact that public shaming often does cause stigmatization and marginalization. Still, this fact does not mean that public shaming cannot be justified. As I have argued, a justification for public shaming is that it enforces important norms, and in turn, the enforcement of those norms can protect our vital interests. So public shaming can sometimes function as a kind of collective self or other-defence. Whether or not an act of self or other-defence is justified turns on whether the attacker is liable to be harmed, and whether the act is necessary and proportionate. It does not depend on whether the motivations of the defender are compatible with our norms of respect. Nor does it depend on whether the harms imposed on the attacker are physical in nature or social in nature. While marginalization and stigmatization constitute very serious harms, so are the harms imposed by defenders in paradigmatic cases of self-defence (eg. death). These serious harms can be justified when causing those harms is an unavoidable side-effect of defending yourself. Thus, my view of public shaming as moral self-defence provides a straightforward response to the kinds of worries raised by Aitchison and Meckled Garcia, and Thomason: publicly shaming a liable wrongdoer is only unjustified if the harms caused by public shaming are unnecessary or

¹⁶⁷ Thomason, *Naked: The Dark Side of Shame and Moral Life* (Oxford: Oxford University Press, 2018), 207.

¹⁶⁸ *Ibid.*, 213.

disproportionate. The fact that public shaming can cause marginalization or stigmatization raises the harmfulness of public shaming, and therefore makes it harder for a given case of public shaming to be proportionate. But by no means does this mean that public shaming will always be unjustified.

VI. Conclusion

In this chapter, we considered the Norm Enforcement View on public shaming, which claims that public shaming is justified when it is used to enforce valuable norms. I raised the authority challenge for the view: that it is unclear why private individuals are permitted to enforce norms via public shaming, when those individuals are typically not permitted to enforce norms via the imposition of serious harm. Among the proposals we have considered, I ultimately defended the claim that private individuals are sometimes permitted to impose harm on one another in order to enforce norms for defensive purposes. In other words, because some norms protect vital interests, we are sometimes permitted to impose serious harms on each other when doing so is necessary to uphold those norms.

One problem with justifying public shaming using the notion of self-protection is that it seems like a limited victory. Paradigmatically, we employ self-defence as a justification in cases where our vital interests are at stake; our lives, our bodily integrity, the integrity of our homes, etc. Thus, if public shaming is just a form of self-defence, it should only be employed against very serious threats, like in cases of sexual assault. This is a conclusion that I somewhat welcome. Cases like that of Justine Sacco (who was shamed for a racist joke on Twitter) or Lindsey Stone (who was shamed for making a rude gesture at the Arlington National Cemetery) are cases where public shamers seem to have overplayed their hands. Public shaming in those sorts of cases, where no one's vital interests were at stake, often seems inappropriate or disproportionate.

Nevertheless, I think that there are some cases where an act may not pose an obvious threat to any vital interests, but where some public shaming is justified. For example, imagine a case of racial discrimination, where a store owner refuses to serve a customer of a certain race. I shall take a particular interest in these sorts of cases in the next chapter. As I shall argue, public shaming can sometimes be justified as a form of moral self-defence: the defence of our moral standing in society.

Chapter IV: Justifying Public Shaming – Public Shaming as Moral Self-Defence

I. Introduction

As we saw in the previous chapter, philosophers who have written on public shaming or other forms of social sanctions have generally characterized these practices as forms of norm enforcement. I agreed that the enforcement of social norms is indeed a valuable function and important goal of public shaming, and that the goal of norm enforcement can justify some contributions to public shaming. In particular, I claimed that because individuals have the right to protect themselves, they have the right to enforce norms which protect their vital interests.

While I have endorsed a version of the Norm Enforcement View, more can be said about the justification of public shaming. The Norm Enforcement View has little to say on the people doing the shaming – on whether specific agents like victims of wrongdoing have special, agent-relative reasons to contribute to public shaming. This strikes me as a blind-spot, because contributions to public shaming often seem to involve significant personal stakes. Contributors are often motivated by a desire to “tell their story” or “be heard”, not just by a desire to enforce social norms for the good of a society at large. Given this, we might wonder if there’s more at stake for these contributors than just the desire to enforce a norm, which plays a role not just in motivating shamers, but in justifying their behaviour.

I believe there is. In this chapter, I will argue that some contributions to public shaming can be justified as a form of *moral self-defence*. Moral self-defence refers to the defence of one’s moral standing – being recognized as an equal in the eyes of oneself and others – rather than the defence

of one's physical body or rights.¹⁶⁹ Wrongdoing can often attack the moral standing of the victim, by implying that the victim lacks the right to be treated better. Certain kinds of wrongdoing, those related to broader social issues like sexism or racism, can attack the moral standing not only of their direct victims, but of all members of the relevant social groups. My claim here is that when these kinds of wrongdoing occur, direct victims, members of relevant groups, and some individuals acting on their behalf, can sometimes be entitled to publicly express anger, contempt, blame, or criticism in order to defend their moral standing. These expressions can together constitute public shaming. I also argue that public shaming is particularly likely to be necessary for moral self-defence in cases which instantiate systemic oppression or historic injustice. In doing so, I hope to draw a stronger link between the justification for public shaming and the cases where public shaming often seems most appropriate – cases related to systemic racism, sexism, failures of public institutions, or other historically salient forms of discrimination. To be clear, the thesis is not that all contributions to public shaming constitute justified moral self-defence, nor that every contribution with the intention of moral self-defence are justified. Rather, the key claim is that moral self-defence is a justification for some contributions to public shaming.

My discussion is structured as follows. In part II, I develop the concept of moral self-defence. In part III, I explain how public shaming can be an important form of moral self-defence. In part IV, I consider some objections, including the potential objection that public shaming as moral self-defence is superfluous, given that norms can protect people's vital interests.

II. Moral Self-Defence

¹⁶⁹ Thomason, Krista, *Naked: The Dark Side of Shame and Moral Life* (Oxford: Oxford University Press, 2018), 187.

In this part, I will develop the concept of moral self-defence, a concept which I owe to Krista Thomason.¹⁷⁰ Moral self-defence, on her view, is the effort to maintain self-respect in response to “offensive behavior of a certain kind.”¹⁷¹ Specifically, moral self-defence refers to an agent’s attempt to protect her own self-respect in the face of arrogant and self-important others. As she writes, “When other people refuse to take us seriously or when they think of themselves as more important than we are, invitations to shame protect our own sense of self-respect”.¹⁷² My first aim in this section is to adapt her idea of moral self-defence as the more general defence of one’s moral standing – a concept which includes an agent’s self-respect, but is not limited to it. Additionally, Thomason does not provide a systematic analysis of why moral self-defence might be a kind of justified self-defence. My second aim to provide this analysis.

Paradigmatically, self-defence involves imposing physical harm on an attacker in order to prevent them from inflicting physical harm on you. *Moral* self-defence refers to the defence of your moral standing, rather than the defence of your physical body. Your moral standing is the recognition of your intrinsic moral status as an equal member of society, by others and yourself. It is thus a combination of two components: recognition of your personhood from others – which philosophers have called “recognition respect”¹⁷³ – and your self-respect. Moral standing is incredibly important, as it is intimately connected with your ability to fulfil most of your other interests. If others do not recognize your status, or if you do not recognize it yourself, both parties are likely to sacrifice your interests for the sake of other goals.¹⁷⁴ So any person who has interests also has a great interest in maintaining their moral standing.

¹⁷⁰ Thomason, Krista, *Naked: The Dark Side of Shame and Moral Life* (Oxford: Oxford University Press, 2018).

¹⁷¹ *Ibid.*, 187.

¹⁷² *Ibid.*, 190.

¹⁷³ Darwall, Stephen, “Two Kinds of Respect,” *Ethics* 88, no. 1 (1977): 38.

¹⁷⁴ John Rawls makes a point like this about self-respect, noting that “Without it nothing may seem worth doing, or if some things have value for us, we lack the will to strive for them. All desire and activity becomes

When does our moral standing require defending? When someone wrongs you, they can implicitly or explicitly attack your standing as an equal. Wrongdoing can express a lack of regard for another's interests. In writing on forgiveness and blame, Pamela Hieronymi notes that intentional wrongdoing can constitute a claim "that you can be treated in this way, and that such treatment is acceptable", due to the way an act reveals certain evaluations made by their author.¹⁷⁵ In principle, even minor, everyday wrongs can attack another's moral standing because they can express a lack of regard.¹⁷⁶ But such threats to one's moral standing are usually small, and are unlikely to justify very harmful defensive action. Sometimes, wrongdoers go further by explicitly asserting their own superiority, like Thomason's example of an arrogant colleague who demands deference.¹⁷⁷ Worse yet are serious and expressive wrongs, like hate speech, which clearly and intentionally communicate that other groups are less important, or less worthy of moral consideration. When wrongdoing goes unaddressed, two things can happen that can impact the moral standing of the victim. First, the moral standing of the victim may diminish in the eyes of the wrongdoer and bystanders. These agents may come to believe, if they don't already, that it is acceptable to treat the victim wrongly, and that the victim has a lower standing. Second and even more pernicious is the possibility that a victim of wrongdoing may fail to recognize her own status as an equal, or lose self-respect. When a victim fails to respond to wrongdoing adequately, she may internalize her wrongdoer's lack of regard, losing her own self-respect. Or she may blame herself for failing to respond, leading to the same result.¹⁷⁸ Not responding adequately to wrongdoing can

empty and vain, and we sink into apathy and cynicism", Rawls, John, *A Theory of Justice* (Cambridge: Harvard University Press, 1971), 440.

¹⁷⁵ Hieronymi, Pamela, "Articulating an Uncompromising Forgiveness," *Philosophy and Phenomenological Research*, 62, no. 3 (2001): 546.

¹⁷⁶ Certain acts, like offering sincere apology or compensation, can often remove a threat to a person's moral standing, thus removing the need for any defensive action. However, a comprehensive discussion on the role of apology and forgiveness is outside the scope of this paper.

¹⁷⁷ Thomason, *Naked The Dark Side of Shame and Moral Life*, 187.

¹⁷⁸ The link between wrongdoing, failure to respond to wrongdoing, and self-respect has been noted in studies on racial and gender discrimination and self-esteem. See for example Harris-Britt, A., Valrie, C.R.,

also constitute or express a lack of self-respect. Scanlon points out that when an agent has been wronged, “to ignore the indifference to her interests that his actions indicate would show a failure to take seriously her own moral status as someone entitled not to be treated in that way. This would be inappropriate, servile, and demeaning”.¹⁷⁹

There are thus two ways in which wrongdoing can threaten the moral standing of the victim – by threatening to lower a victim’s self-respect or their recognition from others. This is where the idea of self-defence comes into play. In theories on paradigmatic self-defence, those who pose threats to the vital interests of others can be liable to be harmed. As we saw in the last chapter, there are different explanations for how such aggressors become liable. Many of these explanations can be adopted for the purposes of moral self-defence. For instance, some philosophers think that an aggressor forfeits their rights against defensive harms by posing a threat to others, which in turn causes them to become liable to defensive harm.¹⁸⁰ Similarly, we might claim that wrongdoers pose a threat to their victim’s moral standing, and are thus liable to be harmed. Other philosophers maintain that an aggressor must be morally responsible for causing a threat.¹⁸¹ We might thus claim that someone is liable to be harmed only if they are morally responsible for threatening another’s moral standing. Yet other philosophers claim that an attacker

Kurtz-Costes, B. and Rowley, S.J., “Perceived Racial Discrimination and Self-Esteem in African American Youth: Racial Socialization as a Protective Factor,” *Journal of Research on Adolescence* 17 (2007): 669-682; Kim, Eunha, and Hansol Park, “Perceived Gender Discrimination, Belief in a Just World, Self-Esteem, and Depression in Korean Working Women: A Moderated Mediation Model,” *Women's Studies International Forum* 69 (2018): 143–150; Bourguignon, David, Eleonore Seron, Vincent Yzerbyt, and Ginette Herman, “Perceived Group and Personal Discrimination: Differential Effects on Personal Self-Esteem,” *European Journal of Social Psychology* 36 (2006): 773–789; Verkuyten, Maykel, “Perceived Discrimination and Self-Esteem Among Ethnic Minority Adolescents,” *The Journal of Social Psychology* 138, no. 4 (1998): 419-493;

¹⁷⁹ Scanlon, T.M., “Giving Desert its Due,” *Philosophical Explorations* 16, no. 2 (2013): 10.

¹⁸⁰ Thomson, Judith, *The Realm of Rights* (Massachusetts, Harvard University Press), 366-371; Renzo, Massimo, “Rights Forfeiture and Liability to Harm,” *The Journal of Political Philosophy* 25, no. 3 (2017): 324.

¹⁸¹ Otsuka, Michael, “The Moral-Responsibility Account of Liability to Defensive Killing”, in *The Ethics of Self-Defense*, eds. Coons and Weber (New York: online edn, Oxford Academic, 2016).

is liable only if they are culpable for creating a threat.¹⁸² In turn, we can claim that only those who are culpable for posing a threat to others' moral standing are liable. The point is that just as attackers in paradigmatic cases of self-defence become liable by posing a threat to others under certain conditions, wrongdoers in cases of moral self-defence can become liable by threatening the moral standing of others, under the right conditions.¹⁸³

Just as there are two ways in which wrongdoing can impact the moral standing of a victim, there are two kinds of interests victims have in relation to moral self-defence. The first interest is an interest *that* her moral standing be protected. The reasons to fulfil this interest are in principle *agent-neutral*, in the sense that the interest can be fulfilled when other people stand up for her and assert her moral standing on her behalf – it does not necessarily matter who takes action. The second interest for an agent is an interest *in* protecting her own moral standing. This interest is connected with self-respect and the need of an agent to assert her own agency or power. It typically exists when an agent's own self-respect is at stake, because of a wrongdoing they may have experienced. It's the kind of interest at play when a victim of wrongdoing thinks "I need to do this", or "I want to face my wrongdoer", or "I need to be heard". The reasons attached to this second interest are *agent-relative*, in the sense that it matters who takes action – it sometimes has to be the agent whose moral standing is at stake. The victim of a wrongdoing may have had other people criticize the wrongdoer on her behalf, but she might still feel a need to stand up for herself. She might feel like unless she does something, she will continue to feel weak or helpless.

¹⁸² Ferzan, Kimberly Kessler, "Culpable Aggression: The Basis for Moral Liability to Defensive Killing," *Ohio State Journal of Criminal Law*, 9, no. 2 (2012): 669–697.

¹⁸³ Because theories of self-defence require an unjust threat, and a liable wrongdoer, in order for justified self-defence to be carried out, agents cannot justify imposing harm as a form of moral self-defence, when they face perceived but non-existent threats, or just threats to their moral standing. For example, a racist cannot appeal to moral self-defence in order to justify harming those who threaten to bring about a more equal society.

This agent-relative interest in defending one's own moral standing sets moral self-defence apart from acts that aim at norm enforcement. Moral self-defence has agent-relative aspects – if a victim has an interest *in* protecting her own moral standing, *she* has to act. Even when other people have already responded to an instance of wrongdoing, the victim of the wrongdoing may still have reasons to undertake some form of defensive action. In contrast, norm enforcement for the most part seems agent-neutral.¹⁸⁴ If the goal of public shaming is just to ensure that people follow social norms, then anyone who is capable of contributing to that goal can potentially justify public shaming on the basis on that goal. In principle, both victims of wrongdoing and other agents can contribute to the goals of norm enforcement via public shaming (in practice, there may be contingent reasons for why victims may be better placed to publicly shame than others). Thus, the goal of norm enforcement does not give victims of wrongdoing special agent-relative justifications to contribute to public shaming.

Nevertheless, it would be a mistake to see moral self-defence as a purely individualistic endeavour which places the focus on agents protecting their own legitimate self-interests. Moral self-defence often has a social aspect to it. The agent-relative aspect of moral self-defence extends beyond an agent acting in order to defend her own interests, but also to agents acting on behalf of friends, family, and members of their community. People often take strong personal interests in the lives and wellbeing of others. Further, people often exist in relationships which entail mutual

¹⁸⁴ To my knowledge, most authors writing on public shaming have not discussed agent-relative permissions to publicly shame. Frye explicitly claims that “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”, “The Problem of Public Shaming,”: 193. An account that is an exception here is Radzik, who argues that an agent's standing to sanction wrongdoers can be defeated by a number of considerations, most notably “the interests victims have in asserting their own authority”, “On Minding Your Own Business: Differentiating Accountability Relations within the Moral Community,” *Social Theory and Practice* 37, no. 4 (2011): 597. Karen Adkins, whose focus is not on the justification of public shaming, notes that shaming can be a way for agents to assert their epistemic and social authority, in “When Shaming Is Shameful: Double Standards in Online Shame Backlashes,” *Hypatia* 34, no. 1 (2019): 82-83. Both claims made by Radzik and Adkins converge nicely with my own view on the notion that individuals can have agent-relative interests based on their interests in asserting their own moral standing.

obligations (like friendships and family). Thus, particular agents may be permitted or even morally required to engage in moral self-defence on behalf of others, even when their own moral status is not under attack, or when the moral status of their social group is not under attack. Such “other-defence” is not always adequate – as I’ve argued, victims of wrongdoing may have an interest in defending their own moral standing, and this interest cannot be fulfilled by others. Nevertheless, other-defence is sometimes very important in cases of moral self-defence. This is because our moral standing and self-respect often depends on the way others treat us. For example, if I am publicly wronged, and no one comes to my defence, I might infer that other members of my community don’t see me as worth defending.

In addition, a person’s moral standing can be threatened indirectly when a wrongdoing picks out a salient social feature.¹⁸⁵ For example, cases of sexual harassment can threaten the moral standing of all women, even those who have not been victims of sexual harassment themselves. This is because some kinds of wrongdoing are a part of broader patterns of injustice and subordination, and can reinforce those patterns. This is why very specific instances of wrongdoing might justify public shaming from many members of the community, not just the person that has been directly wronged.

While paradigmatic cases of self-defence typically involve threats from current or future wrongdoing, a situation may warrant moral self-defence well after wrongdoing has occurred. As I argued in previous paragraphs, after a wrong is carried out, the wrong can threaten a victim’s moral standing by threatening to change the perception of others, or by lowering the self-respect of the victim. This in turn justifies some defensive act. If these ideas are correct, moral self-defence can be “backward-looking” in a way that paradigmatic self-defence is not. You may thus worry that

¹⁸⁵ See footnote 24 for some studies on the relationship between racial, ethnic, gender, or religious discrimination on self-esteem.

there is a disanalogy between moral self-defence and paradigmatic self-defence, which is forward-looking. But there is also a “forward-looking” aspect to moral self-defence, because your moral standing is still on the line. Sufficiently serious wrongdoings can affect the recognition and self-respect of a victim long after the wrongdoing was committed, thus continuing to present a threat to the victim. Hieronymi puts it nicely, noting that sometimes we have reasons to respond to “a past action that persists as a present threat”.¹⁸⁶ Similarly, Margaret Urban Walker observes that unaddressed wrongdoing can have lasting impacts on victims. She writes:

Victims may be uncertain about their own blamelessness and what they deserve. They need to know that others grasp the fact of the violation, its clear wrongfulness, the culpability of the perpetrator, and the reality of the harm and suffering caused them, in order to be validated . . . When these responses are not forthcoming, the victim’s situation is worse than unaddressed, it is aggravated.¹⁸⁷

A threat to a person’s moral standing can persist after the cause of the threat – a wrongdoing – has occurred. And as Walker points out, the threat can sometimes be aggravated by a lack of a response to the threat. Thus, acts of moral self-defence are done in response to past wrongdoing. But in such cases it is still appropriate to characterize moral self-defence as a form of self-defence, because the agent in question protects her interests which are still under threat.¹⁸⁸

III. Public Shaming as Moral Self-Defence

¹⁸⁶ Hieronymi, Pamela, “Articulating an Uncompromising Forgiveness,” 546.

¹⁸⁷ Walker, Margaret Urban, *Moral Repair* (Cambridge: Cambridge University Press, 2006), 19-20.

¹⁸⁸ The notion of moral self-defence can also explain an intuition about punishment that many people have, but is only indirectly implied in some philosophical theories (like Expressivism); that punishment is valuable because it vindicates victims. When an offender is punished, a community broadcasts the claim that it considers the offense wrong and unacceptable. One beneficiary of this broadcast is the victim – she is given clear evidence that her moral community resists the act which wronged her, and that her community denies that she can be treated in such an unjust way.

What does moral self-defence have to do with public shaming? Moral self-defence often requires some expressive act, like protest or criticism. Let us start with protest. In writing on the oppression of African Americans, Bernard Boxill claims that that when a person protests wrongs done to him, “he affirms that his condition is not unavoidable, he insists that what he protests is precisely the illegitimate, and hence avoidable, interference by others in the exercise of his rights”.¹⁸⁹ Protest can do two things. It can influence the behaviour of others by exhorting and pressuring them to treat you as you deserve. Note that this influence can only be achieved when protest is expressed; you cannot change anyone’s behaviour with internal protest “in your own head”. But even when it fails to affect the behaviour of others, your protest is valuable in a more subtle way. As Boxill observes, protesting wrongdoing helps you acknowledge that what has been done is unacceptable, and thus helps you shore up your self-respect. Expressing protest can also help you to restore a sense of agency when you are wronged. We are often powerless to undo wrongs done to us or attain compensation for our losses. But when you express protest, you can tell yourself, “At least I said something”.

But protest does not always amount to shaming, because you can protest a state of affairs without implicating any specific wrongdoer. Public shaming often involves criticism directed towards specific individuals (or sometimes, institutions or groups). Such directed criticism can also be instrumental for moral self-defence, because it expresses not only that a person’s situation is illegitimate, but also that they have been *wronged*. Like protest, directed criticism has several functions. Some of these purposes are oriented towards others. Directed criticism admonishes the wrongdoer, communicating to them that their behaviour is unacceptable. It can also draw the attention of third parties, warning them not to wrong the victim in similar ways, and perhaps encouraging them to intercede on behalf of the victim. Directed criticism can also help a victim

¹⁸⁹ Boxill, Bernard, “Self-Respect and Protest,” *Philosophy & Public Affairs* 6, no. 1 (1976), 61.

shore up their self-respect. When an agent recognizes that she has been wronged, and takes action against the wrongdoer, she can reclaim her sense of agency by “taking the fight” back to her wrongdoer. The #MeToo movement, for example, has been associated with a decreased relationship between sexual harassment and the loss of self-esteem among victims of sexual harassment.¹⁹⁰ Victims can also address a loss of power by exerting some power over their wrongdoer through criticism. Another study focused on those who participated actively in the #MeToo movement (those who publicly criticized sexual harassers), and found that participating in the movement helps victims “find their voice and regain their power”.¹⁹¹

With regard to instances of racial injustice, there is also evidence that various expressive actions (like criticism) can help victims of racial discrimination manage (or “cope” with) incidents of racial discrimination. In a recent systematic review of other empirical studies, Jacob et al. concluded that “Confrontation coping [like direct criticism and expressing anger towards the wrongdoer] has shown to be useful in creating good outcomes such as psychological forgiveness as well as decreased arousal and positive emotions”.¹⁹² Another study found that confrontation coping strategies can help victims of microaggressions overcome feelings of powerlessness.¹⁹³ In addition to confrontation coping strategies, victims of racial discrimination can also use “venting” strategies (like complaining to others and members of their community), which can help victims

¹⁹⁰ Keplinger, Ksenia, Stephanie K. Johnson, Jessica F. Kirk, Lisa Y. Barnes, “Women at work: Changes in sexual harassment between September 2016 and September 2018,” *PloS One* 14, no. 7 (2018): e0218313. <https://doi.org/10.1371/journal.pone.0218313>.

¹⁹¹ Strauss Swanson, C., & Szymanski, D. M. “From pain to power: An exploration of activism, the #Metoo movement, and healing from sexual assault trauma,” *Journal of Counseling Psychology*, 67, no. 6 (2020): 653–668.

¹⁹² Jacob, Grace, Sonya C. Faber, Naomi Faber, Amy Bartlett, Allison J. Ouimet¹, and Monnica T. Williams, “A Systematic Review of Black People Coping With Racism: Approaches, Analysis, and Empowerment,” *Perspectives on Psychological Science* Vol. 18, no.2 (2023): 392–415.

¹⁹³ Sue, D. W., Alsaidi, S., Awad, M. N., Glaeser, E., Calle, C. Z., & Mendez, N., “Disarming racial microaggressions: Microintervention strategies for targets, White allies, and bystanders,” *American Psychologist* 74, no.1 (2019): 128–142.

manage negative emotions and stress when situations cannot be easily changed.¹⁹⁴ In contrast, studies consistently show that “avoidance coping” strategies (eg. doing nothing, exiting, or trying to forget) can lead to further negative mental health effects for victims.¹⁹⁵ Thus, criticism, which can constitute both forms of confrontation coping and venting, can be valuable tools for agents when they’ve been wronged by racial injustice.

Public criticism also often comes with the expression of strong moral attitudes like disgust, resentment, anger, and contempt. Expressing some of these feelings can be valuable for moral self-defence, and dispassionate criticism sometimes doesn’t seem adequate in response to wrongdoing. Why? Firstly, experiencing and expressing strong emotions can sometimes help victims recognize the extent of wrongs done to them, which in turn enables them to address the possible harms associated with the wrongdoing. As Alison Jaggar notes, “outlaw emotions” like anger can often give us indications that we are victims of injustice, and that “Only when we reflect on our initially puzzling irritability, revulsion, anger or fear may we bring to consciousness our 'gut-level' awareness that we are in a situation of coercion, cruelty, injustice or danger”.¹⁹⁶ Secondly, in standing up for ourselves, we don’t just want to communicate that we’ve been wronged. We want to communicate to the wrongdoer the extent of the hurt they have inflicted; we need them to know and recognize the badness of their behaviour. Expressing strong moral attitudes, like anger and resentment, are ways for us to communicate the extent of hurt. In fact, they are sometimes the best ways to communicate serious hurt – we are less likely to treat someone’s complaint as serious or urgent if they do not yell, curse, or cry. In particular, philosophers of

¹⁹⁴ Brown, T. L., Phillips, C. M., Abdullah, T., Vinson, E., & Robertson, J., “Dispositional versus situational coping: Are the coping strategies African Americans use different for general versus racism-related stressors?” *Journal of Black Psychology* 37, no. 3 (2019): 311–335.

¹⁹⁵ Jacob, Grace, Sonya C. Faber, Naomi Faber, Amy Bartlett, Allison J. Ouimet¹, and Monnica T. Williams, “A Systematic Review of Black People Coping With Racism: Approaches, Analysis, and Empowerment”

¹⁹⁶ Jaggar, Alison, “Love and knowledge: Emotion in feminist epistemology,” *Inquiry* 32, no. 2 (1989): 167.

emotion have placed particular emphasis on feelings like resentment,¹⁹⁷ anger,¹⁹⁸ contempt,¹⁹⁹ disgust,²⁰⁰ and ridicule.²⁰¹ As these philosophers have observed, having these feelings is either instrumental for, or constitutive of, possessing self-respect. Expressing those feelings is a means of defending or reinforcing our self-respect.

Why does expressing criticism (and strong moral attitudes) sometimes have to be public for the purposes of moral self-defence? In some cases, the wrongs in question involve public issues. For instance, racist or sexist jokes posted on Twitter. These acts publicly imply that a group of people are inferior. In such cases, what is at stake is not just the lack of recognition from the wrongdoer. What is at stake is the recognition of all members of society, and the public at large. Thus, some acts of moral self-defence need to be addressed not just to the wrongdoer, but to all members of the public. In other cases, the wrongs in question may be private acts connected to public issues. For example, sexual assault. Public criticism may be important in response to these acts, for many of the same reasons stated before – what is at stake is not just the lack of recognition from one or two other people, but from the wider community. Another reason why shaming must sometimes be public is that, as mentioned earlier, our self-respect is tied to the recognition of others. When serious wrongdoing is known by a victim's community but goes uncriticised, the victim may infer that her moral worth within her community is lower than others. This is particularly likely in cases related to systematic injustice or oppression, where a victim may already

¹⁹⁷ See Jeffrie Murphy, who writes that resentment is defends the values of “self-respect, self-defense, and respect for the moral order,” in “Forgiveness, Self-Respect, and the Value of Resentment,” *Handbook of Forgiveness*, eds. Everett L. Worthington Jr. and Nathaniel G Wade (Routledge Handbooks, 2005), 35.

¹⁹⁸ Thomason argues that “anger is part of the recognition that another person has treated me with disrespect or ill will,” in “The Moral Necessity of Anger,” in *The Ethics Of Anger*, eds. Court D. Lewis and Gregory L. Block (Pennsylvania: Lexington Books, 2020), 94.

¹⁹⁹ Macalester Bell claims that contempt “puts those who manifest vices of superiority in their place by presenting its targets as comparatively low in status in virtue of their superbia [the vice of feeling superior to others],” *Hard Feelings: The Moral Psychology of Contempt* (Oxford: Oxford University Press, 2013), 128.

²⁰⁰ Thomason, *Naked: The Dark Side of Shame and Moral Life*, 204-205.

²⁰¹ *Ibid.*

have experienced repeated attacks on her moral standing. On the flip side, a community can affirm a victim's worth through public criticism. Thus, it may be important for members of a community to publicly voice their criticism, for the sake of the victim.

The argument then, is this. Wrongdoing, in particular serious and culpable wrongdoing, can constitute an attack on a victim's moral standing, as well as the moral standing of those indirectly affected. These attacks can harm the victim's direct interests in being treated as an equal, and their indirect interests which depend on their standing in society. Because of these features, wrongdoers can make themselves liable to be harmed. As with paradigmatic self-defence, defending one's interests in the context of moral self-defence is justified (and perhaps morally required), even if doing so causes harm to a liable attacker. Expressing blame, resentment, contempt, and other negative attitudes in public are all ways in which a person can engage in moral self-defence. And when large numbers of people express these attitudes in public, in response to wrongdoing, this constitutes public shaming.

Before moving on to the objections to my view, I want to briefly revisit the issue of proportionality, which we discussed last chapter. We saw in the last chapter that many philosophers worry about the proportionality of public shaming. While I do think that worries about proportionality should give agents pause before contributing to public shaming, this does not mean that public shaming should never be used. As I argued in the last chapter, even when a case of public shaming is narrowly disproportionate, it can still be widely proportionate. When this happens, the case of public shaming can be all-things-considered justified. For example, the harms that Amy Cooper experienced may be disproportionate to what she deserved or what she was liable to suffer. But it is possible that the public shaming of Amy Cooper brought about very important benefits by raising awareness about racist behaviour, enforcing norms against such behaviour, and giving African Americans an opportunity to express their perspectives on the kind

of behaviour carried out by Amy Cooper. If these benefits are weighty enough, they can outweigh the moral disvalue of the narrowly disproportionate harm Amy Cooper experienced. Thus, the public shaming of Amy Cooper might be justified, all-things-considered, even if it is narrowly proportionate. Stated generally, even when a case of public shaming is narrowly proportionate, it may be all-things-considered justified.

The worry about proportionality can be further mitigated once we consider the agent-relative aspect of public shaming. As I argued earlier, public shaming as moral self-defence highlights that contributors to public shaming can have agent-relative reasons to contribute. An agent's interests in contributing to public shaming can sometimes only be fulfilled when the agent herself takes action. Agent-relative reasons mean that *my* contribution to public shaming can be justified, even when I contribute to a case of public shaming that is overall disproportionate (either narrowly or widely). For the sake of illustration, suppose that by making a tiny contribution to a case of public shaming, I will somehow prevent myself from experiencing clinical depression. Even if thousands of people have already contributed to public shaming, resulting in disproportionate harms, it appears that my own contribution is justified. This is because the agent-relative reasons I have to contribute (preventing depression) seem very weighty in comparison to the moral disvalue of my contribution (a tiny contribution to the harm suffered by the target). Of course, real-life contributors to public shaming are unlikely to have such dramatic and stark stakes in a single case of public shaming. But the point is that the agent-relative aspect of public shaming means that there is another way for individual contributions to public shaming to be justified, even when they contribute to overall unjustified cases of public shaming.

IV. Objections

I shall now consider a few objections to my view.

Objection 1: Necessity. In the last chapter, we saw that for an imposition of harm to constitute justified self-defence, it must be necessary.²⁰² With this in mind, an objection to my view is that public shaming is not necessary for moral self-defence. You may wonder whether there are less harmful ways for agents to defend their moral standing in all or most cases. Or you may further think that because of the collective nature of public shaming – where thousands of people contribute in some cases of public shaming – the power imbalance between the shamers and the shamed is so dramatic that they need not use public shaming to achieve their goals.

As noted in the last chapter, an act of harmful self-defence is necessary if and only if it marginally lowers the expectation of the defender (or in cases of other-defence, the potential victim) being harmed by a sufficiently large amount, relative to other courses of action. Frowe and Parry write that “the necessity constraint requires defenders to compare the available means of averting a threat, rank them according to some relevant moral criteria, and use only the means that is favored by that ranking”.²⁰³ The criteria of this ranking system include the likely effectiveness of a course of action at preventing a threat (or the expected harm that a defender will face), as well as the likely harms for the attacker (or expected harm).²⁰⁴ A course of action which imposes moderately greater harms on an attacker in exchange for much better expected outcomes for the defender will rank higher than an alternative. With this in mind, there are some situations in which it is plausible that contributing to public shaming is the best course of action for some agents, for

²⁰² Frowe, Helen and Jonathan Parry, "Self-Defense".

²⁰³ *Ibid.*

²⁰⁴ A more precise proposal, made by Seth Lazar, is that some harmful act of self-defence (H) is necessary to avert a threat (T) if and only if:

There is no other less harmful course of action H*, compared with which the marginal reduction in harm to the prospective victims of T [the threat] achieved by H is insufficient to justify the marginal increase in defensive harm.

Seth Lazar, “Necessity in Self-Defense and War,” 18.

the purposes of moral self-defence. Let us take a final, more detailed look at the case of Amy Cooper:

Amy Cooper: Amy Cooper is a white woman who threatened to call the cops on Christian Cooper, a black man who was watching birds in central park (he had asked her to leash her dog), and falsely claimed he was threatening her life.²⁰⁵ Christian Cooper videoed the event and posted it online. Predictably, she was heavily punished for her actions. According to the New York Times, “internet sleuths” dug into her personal life for more damning information.²⁰⁶ She lost her job, her dog (she appears to treat the dog roughly in Christian Cooper’s video), and was charged in court for lodging a false report.

Amy Cooper’s behaviour constituted a threat in two senses. First, it was a threat to Christian Cooper’s safety and rights. Secondly, it was an attack on the moral and social standing of Christian Cooper, as well as Black people in America (even if she did not intend it as such). For many people, Amy Cooper’s actions represented the threat of racial discrimination and subordination that they experience. As Christian Cooper reflected later, “it’s not about Amy Cooper. What’s important is what her actions revealed: how deeply and widely racial bias runs in the United States”.²⁰⁷ When Christian Cooper chose not to participate in later legal proceedings against Amy Cooper, he notes,

The blowback from some in the Black community for that decision was swift and severe . . . This was personal for many, even within my own family: One relative by marriage was angry, only later sharing with me that her brother had spent years in a South Carolina prison when a white woman had falsely accused him of rape to cover up their affair. Some

²⁰⁵ Walsh, Joan, “Birding While Black: Just the Latest Bad Reason for White People to Call Police,” *The Nation*, 2020. <https://www.thenation.com/article/society/amy-cooper-birding-police/>.

²⁰⁶ Jashinsky, Emily, “Amy Cooper Doesn’t Deserve Sympathy, But Social Media Shaming Is an Unhealthy Norm,” *The Federalist*, 2020. <https://thefederalist.com/2020/05/27/amy-cooper-doesnt-deserve-sympathy-but-social-media-shaming-is-an-unhealthy-norm/>.

²⁰⁷ Cooper, Christian, *Better Living Through Birding* (New York: Random House, 2023).

felt betrayed. Others seemed to want Amy Cooper punished for Emmett Till and every other outrage African Americans have suffered, as if this one individual should bear the brunt of all the wrongs done to us.²⁰⁸

Let us consider the situation of someone who saw Christian Cooper's video, and wants to publicly criticize Amy Cooper for moral self-defence. Let us suppose that they do not coordinate further public shaming, or encourage others to publicly shame Amy Cooper. In publicly shaming her, they individually impose a tiny cost to Amy Cooper, although they contribute to a collective outcome which imposes very large costs on her. For them, publicly criticizing Amy Cooper might be a means to assert their moral worth by protesting the wrongness of her actions, affirming that her behaviour was illegitimate. In doing so, they may achieve a number of goods. First, they can remind members of a community of the discrimination faced by a group of people, and that this discrimination is unjust and intolerable. Second, as noted earlier, public criticism can help agents shore up their self-esteem, and remind themselves that they are entitled to better treatment. Thus, people can use public shaming as a means to avert threats to their moral standing in their community and their self-esteem.

Are there alternative courses of action? Yes, but each has significant disadvantages, which can make them worse options for some prospective contributors. One alternative might be to silently feel resentment, or to express it only in private spaces, to one's friends or relatives. Another alternative would be to express one's resentment privately towards Amy Cooper, perhaps by writing a strongly worded letter to her. I think these two options have serious disadvantages. As I argued earlier, the publicity of some issues, and their nature as public or social issues, means that private criticism is insufficient. This case is a good example. Given the publicity of the original

²⁰⁸ *Ibid.*

wrongdoing, and its connection to the broader public issue of police violence, I doubt this would have been adequate for at least some people most affected by it. A third option is to rely on public institutions, like the law or the news media, to defend one's moral status on one's behalf. While public institutions can indeed be powerful tools for moral defence, they clearly do not cut it in some cases, like when the wrongdoing in question is not illegal. Further, public institutions cannot be relied on when the wrongdoing in question is connected to systemic injustice, like the police mistreatment of Black people, or the routine failure of the criminal justice system in punishing sexual assault. In cases connected to systemic injustice, public institutions are sometimes part of the problem. Because of the downsides to these alternatives, contributing to public shaming may well be the best option for some people.²⁰⁹

In the last chapter, I claimed that public shaming is more likely to be necessary in cases where norms are not well enforced by existing institutions or practices. Here, I can develop this position further and state that public shaming is more likely to be justified as moral self-defence in cases related to oppression, domination, and historical injustice than in other cases. I regard this as a positive feature of my view of public shaming as moral self-defence. One reason for this claim is that when wrongdoing is carried out by members of powerful and dominant groups in society, they are more likely to pose a real threat to the moral standing of others. For example, given the power imbalances across racial groups in America, a White person who publicly makes a racial slur is more likely to pose a threat than a Black person who insults White people. A White person in America usually does not need public shaming to defend herself from a racist attack on her moral standing, because her moral standing may be more secure. A second reason is that cases related to oppression, domination, or historic injustice also involve routine failure of public institutions to

²⁰⁹ Note that this is consistent with claiming that public shaming in her case was all-things-considered unjustified for other reasons. Perhaps, for example, the eventual harms imposed on Amy Cooper were disproportionate in some way, which can potentially give agents decisive reasons not to publicly shame.

administer and deliver justice. When public institutions, like law enforcement, prove unreliable at delivering justice, public shaming is more likely to be the best course of actions for agents seeking to protect their moral standing.

Objection 2: Superfluosity. It might be thought that my view of public shaming as moral self-defence adds little to our existing knowledge. Norm enforcement, to some extent, has a defensive function. This was in fact a key point that I argued for last chapter. We enforce norms against murder, assault, theft, and other crimes partially in order to reduce our own risk of becoming victim to those crimes. We also sometimes enforce norms to protect our moral standing, because norms can play important roles in expressing or protecting the moral standing of people. Even in cases of changing norms, a version of the Norm Enforcement View can cover a lot of what my view has to say. When we establish new norms, our actions can be understood as an attempt to establish norms that better reflect and express the moral standing that we ought to have. So the enforcement of norms and the establishment of norms can potentially be framed as ways of protecting our moral standing. Thus, it could be argued that all I have done here is explore one dimension of the Norm Enforcement View, rather than bring a new way of justifying public shaming to the table.

Public shaming as moral self-defence may indeed overlap significantly with public shaming as a form of norm enforcement. I have not tried to argue that public shaming as moral-defence constitutes a distinct justification for public shaming. What I am interested in is developing an argument that justifies public shaming in some cases, which has not been made before. Furthermore, public shaming as moral self-defence highlights important dimensions of public shaming which have previously been overlooked. Earlier, I pointed out that moral self-defence highlights the agent-relative justification for contributions to public shaming. This can lead to some conclusions that the Norm Enforcement View does not capture. Consider again the position of

someone who wants to contribute to the public shaming of Amy Cooper. Their reasons for wanting to contribute may include the desire to enforce a norm against racist behaviour. This is a somewhat agent-neutral goal, because *they* need not contribute to public shaming for the norm against racist behaviour to be enforced. But their reasons to contribute may also include agent-relative reasons – they may want to speak out against Amy Cooper in order to reclaim their agency, pride, and self-esteem which has been threatened by their own experiences of racist discrimination.

What would the Norm Enforcement View say to this potential contributor? If thousands of people have already shamed Amy Cooper, the Norm Enforcement View might advise this potential contributor not to contribute. With thousands of shamers, it looks like the norm against racist behaviour has already been enforced as much as it can be in that instance. As a result, any additional contribution to the case of public shaming looks like gratuitous piling-on. But public shaming as moral self-defence makes a different claim. As I noted in the last paragraph, one reason in favour of a person's contributing to public shaming is to reclaim their agency, pride, or self-esteem. Even when a norm has already been enforced, a potential contributor may still have something important to defend by contributing. If the contributor has an agent-relative interest in defending their own moral standing, they can continue to have those interests even after many people have already publicly shamed the Amy Cooper. Thus on my view of public shaming as moral self-defence, the potential contributor could be permitted to publicly shame Amy Cooper, long after thousands of people have already done so.

Objection 3: Victimless Wrongdoing. Many people might believe that public shaming is a valuable tool in some cases of victimless wrongdoing. For example, the use of public shaming in response to environmental destruction.²¹⁰ The problem for public shaming as self-defence is that

²¹⁰ Notably, Jennifer Jacquet, in *Is Shame Necessary? New Uses for an Old Tool* (New York: Pantheon Books, 2015), defends the use of public shaming in environmental issues.

it only justifies public shaming when wrongdoing threatens the moral standing of someone. Thus, victimless wrongdoing, which may not threaten the moral standing of anyone, is not something that justifies the use of public shaming.

My main response to this objection is to bite the bullet. I acknowledge that moral self-defence is unlikely to justify all cases of public shaming which we think are justified, but this is not a surprising or devastating conclusion. There may be other moral considerations which may justify public shaming in some way. For instance, the importance of enforcing a norm in some instance may be so great that the consideration outweighs a person's right against public shaming. That said, it is worth pointing out that many cases of apparently victimless wrongdoing can in fact threaten the moral standing of some people. Environmental destruction, for instance, can reduce the wellbeing of future generations and, minorities, and indigenous groups. Each instance of environmental destruction assumes that the interests of the members of those groups can be subordinated to the interests of others. Thus, if left unanswered, environmental destruction can threaten the moral standing of those groups by normalizing the subordination of their interests. So many cases of "victimless" wrongdoing are not in fact so victimless. We might even extend the notion of moral self-defence, or other-defence, to non-human animals. Even if non-humans are not our moral equals, it is likely that at least some non-humans are entitled to some recognition as beings with moral claims. If some wrongdoing threatens the recognition that non-humans are entitled to, we might be in some instances justified in publicly shaming in order to protect this recognition.

V. Conclusion

I have argued in this chapter that public shaming can be justified as a form of moral self-defence. Agents can sometimes engage in moral self-defence by expressing criticism or other negative attitudes publicly, and when many people behave in such a way in response to a single

target, this can constitute public shaming. The claims made here depart from the way existing scholarship has thought about public shaming, as a mostly agent-neutral activity aimed at enforcing valuable social norms. Instead, my arguments focus on the agents carrying out the public shaming, their own reasons for actions, and whether their individual contributions can be justified. By no means are these ways of thinking about public shaming mutually exclusive – I think they are both important facets of the issue. By characterizing public shaming as a form of moral self-defence, in addition to the usual characterization as norm enforcement, I believe that we gain additional insights into many of the compelling cases which motivate our interest in the topic in the first place, like the #Metoo movement, and other cases of public shaming in response to systemic discrimination and historic injustice.

Chapter V: The Moral Responsibility of Individual Contributors

I. Introduction

I now want to focus on the collective nature of public shaming. The goal of this chapter is to examine the relationship between the morality of an individual's contribution to a case of public shaming to the morality of the case of public shaming. Specifically, I am interested in the question: when a case of public shaming causes harm or is unjust, how morally responsible is the average individual contributor?

Many agents contribute to a case of public shaming. Our focus on this chapter is on the average contributor who makes a tiny, perhaps negligible, impact on the relevant case of public shaming. Consider public shaming on social media like Twitter. Most negative tweets in a case of online shaming will never be read by the target and will not directly cause any suffering. Neither will the average tweet be read by the target's boss, loved ones, friends, colleagues, or anyone else important to the target. Instead, it adds to a mass of angry voices, which given enough weight will attract the attention of the target, conventional media, the algorithms used by social media, or people important to the target. Even though a large set of contributions is likely to cause significant harm to a target, by itself, the average contribution to public shaming is unlikely to be a *difference-maker* with respect to the outcome of a case of public shaming. That is to say, it is unlikely, by itself, to cause any noticeable, morally relevant differences in the outcome, even if it may play some causal role in contributing to the outcome.

We can contrast average contributors to other agents within cases of public shaming who do make a significant difference to the outcome of a case of public shaming. Cases of public shaming start with instigators, who kick off a case by publicizing a target's wrongdoing or get the

ball rolling by encouraging others to publicly shame a target. There are also often some outsized contributors, like celebrities, journalists, or people with many online followers, who influence many other people to contribute to public shaming, thus making a significant difference to the case. Those who have a direct relationship with the shamed can also communicate their censure in very direct and impactful ways. They can confront the wrongdoer directly and personally. They can also sever their connections with the shamed. Friends may refuse to maintain contact and communication. Family members may disown. Social groups, such as book clubs, church groups, or improv troupes, can expel the shamed from their networks. Professional organizations can disassociate by firing the shamed or refusing to invite them to conferences. Finally, there are some people who take extreme actions which have come to be called online vigilantism. These range from death and rape threats, using abusive language, tracking down a person's real-life location and harassing them, revenge porn, and doxing, which involves researching personal information about a target (such as their home address or credit card number) and broadcasting it for others to use. This chapter is not about the instigators, friends and family, or online vigilantes. It is about average contributors who are unlikely to make a difference to a case of public shaming.

When an agent makes a difference for the worse and fulfils other prerequisite conditions for moral responsibility (eg. competence, sufficient knowledge, etc.), we take them to be morally responsible. On the other hand, when an agent makes no difference to an outcome, it is unclear if they can be held morally responsible. Public shaming is, in part, what may be called a *collective outcome* case, or a *collective harm* case (or as Julia Nefsky calls it, a collective impact case), where “by acting in a certain sort of way, people collectively cause harm, or fail to prevent it, but no individual act of the relevant sort seems to itself make a difference”.²¹¹ This is a bit of an idealization. As I have noted, some individuals in cases of public shaming do make significant differences to

²¹¹ Nefsky, Julia, “Collective Harm and the Inefficacy Problem,” *Philosophy Compass* 14, no. 4 (2019): 1.

collective outcomes. But the question that puzzles many philosophers is not whether such difference-makers have responsibilities to act in certain ways. The question is whether average contributors have such responsibilities, given that their contributions do not seem to make a direct noticeable impact on the wellbeing of any other individual.

The topic of collective harm has been extensively written on in contemporary philosophy, in relation to a variety of topics including climate change, voting, and consuming animal products. Multiple accounts have been offered to explain how average contributions to collective harm cases are morally wrong. These include Kant-inspired accounts (“What would happen if everyone did that?”)²¹², accounts based on contributions being a part of a relevant set,²¹³ membership, complicity, fairness or doing one’s share,²¹⁴ virtue, expected utility,²¹⁵ and causal involvement.²¹⁶ Many of these accounts are controversial and some philosophers deny that average individuals have any duties in collective harm cases.²¹⁷ I cannot examine each of these accounts in this piece of work. Instead, my goal in this chapter is to identify a feature of public shaming that is somewhat distinctive: that many contributors to cases of public shaming act on *shared intentions* to help bring about harmful collective outcomes. When an agent has a shared intention, they have a goal that some collective outcome (like someone being publicly shamed) be achieved by a group, and they intend to do their part in realizing that outcome. I think that this feature of public shaming can

²¹² Budolfson, Mark, “The Inefficacy Objection to Deontology: What it is, Why it is Important, and How to Respond to it,” *Budolfson.com* (retrieved 17 Nov 2023), <https://budolfson.github.io/files/BudolfsonDeontology.pdf>.

²¹³ Barry, Christian, and Gerhard Øverland, “Individual responsibility for carbon emissions: Is there anything wrong with overdetermining harm?” in *Climate Change and Justice*, ed. Jeremy Moss (Cambridge: Cambridge University Press, 2015).

²¹⁴ Cullity, Garrett, “Climate Harms,” *The Monist* 102, no. 1 (2019): 22–41.

²¹⁵ Kagan, Shelly, “Do I Make a Difference?,” *Philosophy & Public Affairs* 39, no. 2 (2011): 105–141.

²¹⁶ For an overview on some objections to various accounts, see Nefsky, Julia, “Collective harm and the inefficacy problem,” *Philosophy Compass* 14, no. 4(2019): e12587.

²¹⁷ For an example of denial, see Sinnott-Armstrong, Walter, “It’s Not My Fault: Global Warming and Individual Moral Obligation,” *Perspectives on Climate Change: Science, Economics, Politics, Ethics Advances in the Economics of Environmental Research*, Vol 5(2005): 293–315.

help us attribute moral responsibility to those that contribute to public shaming with shared intentions. As a result, when an agent acts on a shared intention to help bring about a case of public shaming, they can be morally responsible in a way that they would not otherwise be, had they not had this intention.

Here is an outline of the chapter. In part II, I make some remarks on moral responsibility – what I take it to be in this chapter, and why it is important for the practice of public shaming and the way we sustain norms regarding public shaming. In part III, I argue that acting on a shared intention to contribute to a harmful collective outcome that is harmful can make one morally responsible for helping to bring about the harm, in a way that one would not otherwise be. Because many contributors to cases of public shaming do indeed act on shared intentions to help bring about harmful collective outcomes, I conclude that we can attribute moral responsibility to those contributors in many cases of public shaming.

II. Moral Responsibility

Why do I focus on moral responsibility in this chapter, rather than the moral duties of contributors and other agents? There is established debate about whether an agent's intentions can affect whether they have duties to act in certain ways.²¹⁸ Consequently, it is unclear if an agent's shared intentions can affect their moral duties. I cannot resolve this debate in this chapter, which is why it focuses on the issue of moral responsibility rather than on the duties of agents. As a result, the themes of this chapter depart from a number of publications in the collective harm literature, as well as many of the themes in the rest of this thesis, where I focus on the rights and duties of people in cases of public shaming. Nevertheless, the analysis here is not inconsequential. Knowing whether an agent is morally responsible in a given situation is valuable because moral responsibility

²¹⁸ See McIntyre, Alison, "Doctrine of Double Effect," *The Stanford Encyclopedia of Philosophy*, eds. Edward N. Zalta & Uri Nodelman, 2003, <https://plato.stanford.edu/archives/win2023/entries/double-effect/>.

plays an important role in the way we hold others to account for their behaviour. In turn, the way we assign moral responsibility in cases of public shaming is consequential in establishing and sustaining norms regarding the practice of public shaming.

When we say that someone is morally responsible for some act or outcome, we can mean a number of things. Moral responsibility could be a function of how much control an agent has over her actions or outcomes²¹⁹ – whether she could have done otherwise, and whether she could have acted in a way which would bring about a different outcome. Call this conception of moral responsibility *responsibility as control*. In collective outcome or collective harm cases, average contributors may have no control over whether an outcome obtains or not. Thus, they may not be morally responsible in the sense of having control over the outcomes of public shaming.

In addition to moral responsibility as control, philosophers have also discussed moral responsibility as *attributability*, *answerability*, and *accountability*.²²⁰ There is significant overlap between these concepts. Still, in this chapter, I am mainly concerned with *moral responsibility as accountability*, which is the responsibility associated with “accountability practices”, most salient of which is the practice of blaming (some philosophers, like Strawson, have also pointed out that we may have other negative reactive attitudes like resentment or contempt. I think this is plausible, but for our purposes I will focus on blame).²²¹ For example, Franklin writes that “An agent S is morally accountable for performing an action φ if S would deserve praise if φ went beyond what could be reasonably expected of S, and S would deserve blame if φ was morally wrong”.²²² Thus, in responsibility as accountability, when we say that a person is responsible for some wrongdoing or

²¹⁹ Talbert, Matthew, "Moral Responsibility", *The Stanford Encyclopedia of Philosophy* (2023), Edward N. Zalta & Uri Nodelman (eds.), <<https://plato.stanford.edu/archives/fall2023/entries/moral-responsibility/>>.

²²⁰ Franklin, Christopher Evan, *A Minimal Libertarianism: Free Will and the Promise of Reduction* (Oxford: Oxford University Press, 2018), 34.

²²¹ Talbert, Matthew, "Moral Responsibility".

²²² Franklin, Christopher Evan, *A Minimal Libertarianism*, 35.

some objectionable outcome, we mean that it is appropriate to blame them. As Angela Smith puts it:

We often use the term 'responsibility,' however, in a way that implies something . . . about our relation to the agent and our (usually negative) assessment of her conduct. When we say that we "hold" someone morally responsible for something she has done, for example, what we usually mean is that we blame her for it. Thus, the question in collective harm cases isn't "could the agent have made a difference to the outcome," but "is the agent blameworthy for contributing in the way that they did?"²²³

Responsibility as control and responsibility as accountability can overlap in some cases. An agent, acting alone, is usually said to be blameworthy when they knowingly and intentionally cause harm, without sufficient considerations in favour of bringing about the harm. In such a case, the agent is taken to have had control over the outcome, and is blameworthy because of their culpable actions which played a decisive role in deciding the outcome.

But the two kinds of responsibility can come apart in other cases. We can regard an agent blameworthy, even when they had no control over a harmful outcome (although we might still require that the agent be in control of the performance of particular actions). This is because the blameworthiness of an agent is also a function of the attitudes, judgments, values, or evaluations made by the agent. For example, if an agent knowingly and maliciously causes harm in a case of overdetermined harm, we would still blame the agent despite the fact that they could not control the outcome.²²⁴ In other cases, having certain attitudes towards another agent can increase an

²²³ Smith, Angela, "On Being Responsible and Holding Responsible," *The Journal of Ethics*, Vol. 11, No. 4 (2007): 468. Also see Shoemaker, David, "Attributability, Answerability, and Accountability: Toward a Wider Theory of Moral Responsibility*," *Ethics* 121, no. 3 (2011): 489-691; Watson, Gary, "Two Faces of Responsibility," *Philosophical Topics* 24, no. 2 (1996): 227-248.

²²⁴ For example, the following case in Parfit, Derek, *Reasons and Persons* (Oxford: Oxford University Press, 1984), 70:

agent's blameworthiness over and above the blame we would normally assign to them for committing the act. For instance, an agent who kills another person out of sheer hatred is normally held to be more blameworthy than someone who kills another as a foreseen but unintended side-effect of pursuing a different goal. P.F. Strawson makes a very similar point with respect to another "reactive attitude" – resentment. He writes that the actions of others "reflect attitudes towards us of goodwill, affection, or esteem on the one hand or contempt, indifference, or malevolence on the other. If someone treads on my hand accidentally, while trying to help me, the pain may be no less acute than if he treads on it in contemptuous disregard of my existence or with a malevolent wish to injure me. But I shall generally feel in the second case a kind and degree of resentment that I shall not feel in the first".²²⁵

In the next section, I will argue that the shared intention to help bring about harmful collective outcomes is blameworthy. But first, what is the point of determining the moral responsibility of agents within the context of public shaming? I think our ability to assign moral responsibility in cases of public shaming is important because moral responsibility plays an important role in the establishment and sustaining of norms regarding the practice of public shaming.

For a norm to be established or sustained, it must be complied with. To ensure compliance, we need to hold each other to account for any violations of the norm. In this case, this means that sanctions need to be placed on non-compliers. These sanctions can take two forms. First, we have "tangible sanctions", which generally involve the "threatened imposition of concrete burdens, or

"X and Y simultaneously shoot and kill me. Either shot, by itself, would have killed."

²²⁵ Strawson, P.F. "Freedom and Resentment," *Proceedings of the British Academy* 48 (1962):187-211. Similar observations have been made by Smith, Angela M. "Control, responsibility, and moral assessment," *Philosophical Studies* volume 138: 368; Scanlon, T.M., *What We Owe to Each Other* (Harvard: Harvard University Press, 1998); Hieronymi, Pamela, "The Force and Fairness of Blame", *Philosophical Perspectives* 18, no. 1 (2004): 115–148; Sher, George, "Out of Control," *Ethics* 116, no. 2 (2006): 285–301.

withholding of concrete benefits, in order to elicit a particular response, or to punish deviation from the response”.²²⁶ Examples of tangible sanctions include legal sanctions, like criminal punishments, but also social punishments like “gossiping about norm violators behind their backs, ostracism, and social exclusion”.²²⁷ And of course, there’s public shaming, which can impose harms like losses in social goods, which I detail in this thesis. Second, we have “intangible sanctions” which come in the form of attitudes like disapproval, “contempt, disgust, disesteem, disapprobation, and so on”.²²⁸ As Brennan et al note, even though intangible sanctions do not impose an obvious cost on targets of the sanctions, they can still influence behaviour because “most people care about what others think of them”.²²⁹ Hence intangible sanctions can help enforce norms.

There is a very close relationship between a person’s moral responsibility and their liability to be sanctioned. Generally, the amount and severity of sanctions which we deem appropriate to impose on someone is proportionate to their moral responsibility. When we discover some fact which makes a wrongdoer less morally responsible for their behaviour, we ought to impose less severe sanctions on them. For example, if we learned that a wrongdoer was not in fact normatively competent at the time when she committed the wrongdoing, we should regard this fact as a mitigating factor that reduces the severity of the sanctions we impose on her. There are a couple of potential explanations for this observation. Firstly, that moral responsibility plays a role in justifying sanctions. While there may be multiple reasons to sanction a wrongdoer, like deterring undesirable behaviour, perhaps imposing any sanction is typically justified only *because* a wrongdoer is morally responsible. Secondly, that there is some underlying set of factors which make a person

²²⁶ Brennan, Geoffrey, Lina Eriksson, Robert E. Goodin, and Nicholas Southwood, *Explaining Norms* (Oxford: Oxford University Press, 2013), 221.

²²⁷ *Ibid.*

²²⁸ *Ibid.*, 224.

²²⁹ *Ibid.*

both morally responsible and liable to sanctions at the same time, or that makes a person's moral responsibility coextensive with their liability to be sanctioned.

For our purposes, it does not matter whether moral responsibility justifies sanctions, or whether there is an underlying common set of factors which explains both moral responsibility and liability. What matters is that a person's moral responsibility reliably indicates when we may impose sanctions on others, and how severe those sanctions should be. As a result, it is important to determine the moral responsibility of agents if we want to sanction them or hold them to account for violating norms.

We can expand on this argument further by concentrating on the specific feature of blame. While philosophers differ on this topic, the majority seem to agree that moral blame involves two components. Firstly, an a) evaluation of another agent as having failed to meet some standard.²³⁰ Secondly, b) a normative element – blame is levelled at agents when they have failed to meet a normative standard.²³¹ We do not blame someone for being short, ugly, or being bald. We blame people for failing to meet standards that they ought to have met.

Again, there seems to be a close relationship between judgments of blameworthiness and judgments that someone is deserving of being sanctioned. It is strange (and sadistic) for someone to say, "I think you're not blameworthy, but I think you deserve to suffer". As I noted before, this could be because moral responsibility as accountability either justifies sanctions, or shares some underlying justifying factor with liability to sanctions. Given the definition of blame, we can now say more about both of these explanations. Blame involves an evaluation that an agent has failed to meet a normative standard (on some accounts, blame *is* such an evaluation). Sanctions only

²³⁰ Coates, D. Justin, and Neal A. Tognazzini (eds), 'The Contours of Blame', in D. Justin Coates, and Neal A. Tognazzini (eds), *Blame: Its Nature and Norms* (New York, 2012; online edn, Oxford Academic, 23 May 2013), <https://doi.org/10.1093/acprof:oso/9780199860821.003.0001>.

²³¹ *Ibid.*

seem appropriate when an agent has failed to meet a normative standard – in fact, sanctions only seem justified when the sanctioner accurately judges the sanctioned to have failed to meet some standard. Thus, a judgment that an agent is blameworthy explains and justifies the judgment that the agent is liable to be sanctioned.

In addition to justifying and explaining sanctions, blame motivates the application of sanctions. As Franklin notes, “blame disposes us to respond with some kind of hostility toward the blameworthy agent, such as verbal rebuke or partial withdrawal of goodwill”.²³² Indeed, blame is often (although not necessarily) accompanied by strong negative feelings directed toward the target of blame. Blamers “characteristically experience a range of core emotions”, including “anger, indignation, resentment, and outrage”.²³³ Given that blame is usually accompanied by such strong emotions, we can hypothesize that blame motivates actions like the imposition of sanctions. This hypothesis explains why we may be willing to impose sanctions on targets even when the imposition of sanctions is very costly, and when it is not in our rational interests (from the standpoint of cost-benefit analysis) to impose those sanctions.²³⁴

At this point, you might point out that blame might not be the force doing the work here, and rather that the emotions accompanying blame (eg. anger) are the ones at work. More precisely, that emotions like anger are the primary motivating forces behind the imposition of sanctions, and that blame just comes along for the ride. I don’t think this view is correct, because feelings like anger cannot always be separated from blame. Blame often causes anger, and a lack of blame can often mitigate feelings of anger. So it is misleading to simply say that anger motivates agents while

²³² Franklin, Christopher Evan, *A Minimal Libertarianism: Free Will and the Promise of Reduction* (Oxford: Oxford University Press, 2018), 35.

²³³ McGeer, Victoria, 'Civilizing Blame', in D. Justin Coates, and Neal A. Tognazzini (eds), *Blame: Its Nature and Norms* (New York, 2012; online edn, Oxford Academic, 23 May 2013), <https://doi.org/10.1093/acprof:oso/9780199860821.003.0009>.

²³⁴ *Ibid.*

blame is inefficacious. When I discover facts about a case that make an alleged wrongdoer more or less blameworthy, my disposition to blame them changes accordingly and with it, my desire to sanction them. Suppose then that someone wrongs me, and I later discover that they did not have the mental capacities of a moral agent, rendering them less blameworthy. This discovery compels me to “recalibrate” the blame that I experience towards the agent,²³⁵ which involves a change in the way I evaluate the agent. It is this change in evaluation which causes a reduction in any anger I feel, which may in turn reduce my motivation to sanction the target.

Furthermore, our motivation to sanction others does not as reliably track the anger we might have towards them, as it does with blame. It seems true that we most desire to impose sanctions on others when we are angry at them. But a person can also desire to impose a sanction on a wrongdoer without feeling any anger at all. A jaded judge who has lost all feeling from the number of cases she has presided over may still feel a desire to sentence a convicted criminal to jail. A victim who has become numb to feelings of anger or sadness may still want to see their wrongdoer punished. Thus, it seems more likely that blame plays a large role, if not the primary role, in driving the motivation to sanction.

Judgments of blameworthiness affect how much we blame targets, which affects our emotional responses to targets, which in turn affects how motivated we are to sanction them. Because sanctions are an important component of the way we establish and sustain norms, judgments of blameworthiness and blame are important elements in creating and regulating norms. A similar point has been observed by McGeer, who claims that the very function of blame is the regulation of norms; “blame is a quasi-autonomous, emotionally mediated response to others that is specifically prompted by, and targeted on, behavior that transgresses personally or socially valued

²³⁵ *Ibid.*

norms”.²³⁶ As a result, moral responsibility as accountability plays roles in indicating when we may impose sanctions on others, and motivates us to impose those sanctions when appropriate. Determining when people are morally responsible can thus help us hold agents to account, and in particular, contributors to public shaming to account.

With this in mind, the key question for us is: “Why can it be appropriate to blame contributors to cases of public shaming?” I shall answer this question in the following part.

III. Public Shaming and Shared Intentions

A somewhat distinctive feature of public shaming is that contributors to public shaming often have *shared intentions* in a way that contributors to many other cases of collective harm don’t. Contributors to cases of public shaming often aim to help bring about certain outcomes, like holding wrongdoers to account,²³⁷ enforcing norms regarding social issues,²³⁸ or imposing retributive justice.²³⁹ These outcomes are collective outcomes, because they are brought about by the acts of many individuals. Contributors often see themselves as playing their part in a bigger group effort. Of course, this may not be the case for all contributors. Some people who contribute to public shaming may be motivated by a desire to express themselves, or by a desire to morally grandstand (acting in order to improve one’s social profile or moral reputation),²⁴⁰ and may thus

²³⁶ *Ibid.*

²³⁷ Tandoc, E. C., Tan Hui Ru, B., Lee Huei, G., Min Qi Charlyn, N., Chua, R. A., & Goh, Z. H. “#CancelCulture: Examining definitions and motivations,” *New Media & Society*, published online, (2022), <https://doi.org/10.1177/14614448221077977>.

²³⁸ *Ibid.*; Also see Skoric, Marko M., Jia Ping Esther Chua, Meiyan Angeline Liew, Keng Hui Wong and Pei Jue Yeo, “Online Shaming in the Asian Context: Community Empowerment or Civic Vigilantism?” *Surveillance & Society* 8, no. 2 (2010): 181-199.

²³⁹ See Barron, Anna C., Lydia Woodyatt, Emma F. Thomas, Jia En Katherine Loh, and Katherine Dunning, who found that participants in a study were more likely to contribute to public shaming when they perceived a wrongdoer as deserving harm, in “Doing good or feeling good? Justice concerns predict online shaming via deservingness and schadenfreude,” *Computers in Human Behavior Reports* 11 (2023): 100317.

²⁴⁰ Tosi, Justin, and Brandon Warmke, “Moral Grandstanding,” *Philosophy & Public Affairs* 44, no. 3 (2016): 167-217.

be indifferent to whether a collective outcome is achieved or not. Nevertheless, a large proportion of contributors to cases of public shaming are interested in the collective outcome and do aim for the collective outcomes to come about. In contrast, things are quite different in many other cases of collective harm. In cases like climate change caused by carbon emissions, and factory farming caused by the consumption of meat, most contributors do not aim to bring about the collective outcome. Quite the opposite, in fact – most of us contribute to climate change *in spite* of the fact that doing so will contribute to collective harm. Or again, when I order fried chicken from a restaurant, I do it because it is a part of a succulent Chinese meal, in spite of the fact that I may contribute to the suffering of some chickens.

Roughly, a shared intention is an intention to do one's part in the realization of a collective goal – a goal that is realized via the contribution of multiple people – on the condition that others do their part in realizing the same goal.²⁴¹ Thus, a person acts on a shared intention to help bring about a case of public shaming if she contributes to it with the goal of helping to bring about a case of public shaming, and if others act on a similar goal. For example, consider the following two agents:

Ali: Ali reads that a celebrity has told a racist joke. Upset, he goes on Twitter and expresses his anger. While he knows that others may be acting similarly, he does not care, and would express his anger regardless of what others do. He also does not care if the celebrity, or anyone else, reads his post. For him, what is important is “putting his views out there”.

Beth: Beth also learns about the celebrity's racist joke. She goes online to express her anger, but she desires that the celebrity be “brought to justice” by the online community. She

²⁴¹ Schweikard, David P. and Hans Bernhard Schmid, "Collective Intentionality", *The Stanford Encyclopedia of Philosophy* (Fall 2021 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/fall2021/entries/collective-intentionality/>.

wants to play her part in a broader effort, and her desires explain her behaviour – if she did not believe that the celebrity would experience any harm from public shaming, she would not bother with contributing.

The claim here is that Beth possesses and acts on a shared intention to help publicly shame the celebrity (on the condition that others also possess similar intentions), while Ali does not. The shared intention to help publicly shame someone involves an intention to contribute in order to (or with a view to) “do one’s part” in causing them to be publicly shamed. Shared intentions are not unique to the issue of public shaming. It is familiar in a very wide range of activities – singing a duet, carrying heavy furniture up the stairs with someone, playing a team sport, voting alongside a bloc of other voters, etc. But as I noted earlier, shared intentions are more common in cases of public shaming than in many other collective harm cases.

Can we be a bit more specific on what shared intentions entail? Theories about shared intentions have been developed by a number of philosophers, which differ in their details. Nevertheless, a common thread between these different accounts is that when a set of agents have a shared intention, they each have the goal that a collective outcome be realized, while also planning to carry about some individual contribution to that collective outcome. Kutz, for example, claims that individual participatory action (his term for acting with shared intention) “aims at two goals: accomplishment of a primary individual task that contributes to a secondary collective achievement”.²⁴² On Tuomela and Miller’s account of shared intentions, agent A “we-intends” to do X only if “(i) A intends to do her part of X”.²⁴³ They further clarify that when an agent A intends to do her part of X, “he intends that X be realized”, and that the agent “wants to do his

²⁴² Kutz, Christopher, *Complicity: Ethics and Law for a Collective Age* (Cambridge: Cambridge University Press, 2000), 82.

²⁴³ Tuomela, Raimo, and Kaarlo Miller, “We-Intentions,” *Philosophical Studies* 53 (1988): 375.

part of X”.²⁴⁴ Some philosophers go a bit further and claim that agents with shared intentions must aim to help bring about the collective outcome *via* making their own contribution. On John Searle’s account, in cases of shared intentions, “each member is trying to achieve the common goal by means of making his or her individual contribution”.²⁴⁵ On Michael Bratman’s account, shared intentions involve “intentions on the part of each in favor of the joint activity . . . by way of relevant mutual responsiveness in sub-intention and action”.²⁴⁶

We may say that an agent acts on a shared intention when the shared intention explains their behaviour. That is, if the agent behaves in a certain way because they aim at a collective outcome, and because they believe that others will also play their part in realizing the collective outcome. To illustrate this, Kutz presents a case of two people respectively buying wine and cheese for a picnic, and notes: “that my conception of the picnic as *our* – that is, in part *your* – project plays an essential role in rationalizing, or explaining teleologically, my particular action. If I did not believe you would also do your part towards our picnic, I would not buy cheese”.²⁴⁷

When an agent acts on a shared intention to help bring about a harmful collective outcome, they can be morally responsible, in that they may be held to account and be considered blameworthy.²⁴⁸ Whether or not they are in fact blameworthy depends on a number of additional

²⁴⁴ *Ibid.*, 377.

²⁴⁵ Searle, John, *Making the Social World: The Structure of Human Civilization* (Oxford: Oxford University Press, 2010), 55.

²⁴⁶ Bratman, Michael E., *Shared Agency: A Planned Theory of Acting Together* (Oxford: Oxford University Press, 2014), 85.

²⁴⁷ Kutz, *Complicity*, 82.

²⁴⁸ It is also possible for an agent to have a shared intention and contribute to a collective outcome in such a way that the shared intention *does not* explain their behaviour. An agent may be motivated to contribute by a number of things, so that even if they did not believe that the collective outcome would be obtained, they would still contribute. For example, I may want to go on a picnic with you, but also want to buy cheese regardless of whether the picnic happens or not. In this kind of case, we may say that I have a shared intention, but that I do not act on a shared intention. When an agent has a shared intention to bring about a harmful outcome, but does not act on the shared intention, I think we may blame them additionally for having the intention. But it is unclear if we should blame them additionally *for their behaviour*, because their behaviour is not explained by their intentions, and is therefore distinguishable from their intentions. Thus,

considerations, like whether the harmful collective outcome is wrongful (with respect to public shaming, whether the target is liable to be publicly shamed). Other considerations that can affect the moral responsibility of an agent include knowledge of the consequences of their behaviour, the underlying reasons for possessing their intentions, whether they possess the capabilities associated with moral agency, etc. An agent's shared intentions to help bring about harmful collective outcomes is thus one property which can justify blaming them, but is neither necessary nor sufficient for blame.

A shared intention to help bring about a harmful collective outcome can be blameworthy for many of the same reasons an intention to harm is often considered blameworthy. Agents who possess these sorts of intentions often possess objectionable attitudes towards the targets of the harm, like ill will or lack of regard. As we saw earlier, blame is partially a function of whether an agent's behaviour "can be seen as indicative or expressive of her judgments, values, or normative commitments".²⁴⁹ Blame is therefore a function of the intention to harm because when you have the goal of bringing about harm, you express certain sets of attitudes towards the target of the harm. Perhaps you think that the interests of the target are morally unimportant, in comparison to other considerations. Or perhaps you actively harbour ill will or animosity towards the target, and desire that they be harmed. Or, perhaps you see the target as means of fulfilling your goals, and are willing to harm them in the process (eg. you may want a wrongdoer to experience harm in order to deter potential wrongdoers from also causing harm), as Warren Quinn puts it, "The agent of direct harm . . . proposes to involve them in some circumstance that will be useful to him precisely because it involves them. He sees them as material to be strategically shaped or framed

any blame that we attribute to them for their behaviour may be attributable on the basis of the behaviour itself, not the intentions.

²⁴⁹ Smith, Angela M. "Control, responsibility, and moral assessment," 368.

by his agency . . . He must treat them as if they were then and there for his purposes”.²⁵⁰ Either way, you have some set of attitudes which the target can potentially object to. If you regard the interests of the target as unimportant, the target might object that you have disregarded her interests. If you bear animosity towards the target, she might object that your animosity is undeserved, inappropriate, or disrespectful. Thus, it would be appropriate for her to blame you under the circumstances.

Intending to bring about a harmful outcome is not necessarily the same as intending to harm. You can intend to cause a harmful outcome without intending to harm, when the harm caused by your intended outcome is a side effect of your actions and when you do not aim for the harm to occur. For the same reason, intending to help bring about collective harm is not the same as intending to harm. A group of contributors to public shaming may aim to bring about a case of public shaming where the harm caused by the public shaming is not an aim but a side effect of the group’s behaviour.

Nevertheless, when an agent acts on a shared intention to bring about collective harm, their behaviour often indicates judgments about the victims of the harm or the status of the harm. Even when the agent does not have an intention to harm to the victims, they may still have attitudes like a lack of regard for their victims. Therefore, they can be held to account and blamed for those judgments. They can also be held to account and blamed for their actions which they carry out because of those judgments. Kutz notes that “We are properly held accountable for the actions of groups (and of individual group members) in which we participate, because those actions represent our own conception of our agency and our projects”.²⁵¹ Consider for instance a voter

²⁵⁰ Quinn, Warren, “Actions, Intentions, and Consequences: The Doctrine of Double Effect,” *Philosophy & Public Affairs* Vol. 18, No. 4 (1989): 348.

²⁵¹ Kutz, *Complicity*, 140-141.

who votes for a racist immigration policy with the goal that the immigration policy gets passed. Such a person may not desire for harm to come to others, but they still have objectionable attitudes and judgments about people of other races, and their attitudes are represented by her vote. And they are blameworthy for acting on those attitudes, even if their actions by themselves are not decisive in bringing about racist policy. Or consider the following examples:

Charles: Charles drives a gas-guzzling sports car up and down the road for no reason other than enjoyment. He is aware that doing so will contribute to climate change, and that climate change will cause serious harm, but this does not stop him from driving for the sake of enjoyment.

Dan: Dan drives a gas-guzzling sports car up and down the road. He enjoys the drive, but his real reason for driving is because he knows that doing so will contribute to climate change. He is aware that climate change will cause serious harm, but for some reason he is motivated to do what he can to help bring climate change about.

Let us hold fixed all other factors other than the intentions of the agents. According to my earlier arguments, while both Charles and Dan may do something wrong here, Dan is blameworthy in a way that Charles is not. This is because Dan possesses some objectionable or sadistic attitudes regarding the potential suffering that climate change causes, and performs his act because of those attitudes.

With respect to public shaming, a person's contributions can also represent their own judgments and attitudes. These contributions may therefore be blameworthy, depending on various independent factors. A contributor may, for instance, bear ill will to her target. When a person has ill will towards a target, and contributes to public shaming, her contribution expresses this ill will. Given that expressing attitudes of ill will are sometimes grounds for assigning blame,

then a person who aims at causing harm via collective activity can be blameworthy. Or again, consider a person who doesn't bear ill will towards a target, but nevertheless has goals which involve harming the target. For example, they want their community to express condemnation for the target's behaviour. Such a person would still have certain judgments towards the target – like that the target's interests are outweighed by the importance of expressing condemnation. In some cases, these judgments can reflect some objectionable underlying attitudes, like an under-valuing of the target's interests, or a willingness to treat the target as a means to an end. This kind of underlying attitude can be blameworthy, depending on the circumstances under which the attitude was formed and the reasons for which the attitude was formed.

Consider the earlier example where both Beth and Ali contribute to a case of public shaming, where Beth acts on a shared intention to bring about public shaming, while Ali does not. Suppose that as it turns out, the celebrity was not liable to suffer harm from the community, and that the case of public shaming was unjustified. In such a case, both Ali and Beth may be blameworthy for contributing to the case of public shaming. But Beth may be blameworthy in a way that Ali is not: for acting on her shared intentions. Because she intends for the celebrity to be subject to public shaming and be “brought to justice”, she intends for an outcome she knows is harmful, and she bears some ill will towards the celebrity. Given that the celebrity is not liable to be harmed, it seems that there is an additional reason to blame Beth – intending an outcome where harm comes to an undeserving person. In contrast, Ali is indifferent to whether the celebrity experiences any harm from public shaming, and therefore does not bear any particular ill will towards the celebrity. So while he may be blameworthy for contributing to harming the celebrity, he is not blameworthy for having any ill will towards them.

IV. Conclusion

Contributors to public shaming typically have shared intentions to bring about collective outcomes, while contributors to other cases of collective harm typically do not. I have made the argument that when a person acts on a shared intention to help bring about collective harm, they are morally responsible in the sense that we may blame them. Even if they do not make a morally relevant difference to the outcome, we may blame them, and this blameworthiness may be the basis for justifying further sanctions on them. Attributing this sort of moral responsibility to public shamers is therefore important in holding those shamers to account, and enforcing the norms that we have regarding public shaming.

Nevertheless, as I stated earlier, there are limits to the claims made in this chapter. A public shamer with the shared intention to bring about harm is not automatically blameworthy. If the imposition of a harm is justified, then an agent may not be blameworthy in intending to bring about the harm. In addition, blameworthiness can also be mitigated by a number of factors, like ignorance, the effect of very strong emotions, or the effect of social influences. I pick up on the specific issue of ignorance in the next and final chapter, which discusses the use of public shaming when we are uncertain if public shaming ought to occur.

Chapter VI: Knowing and Shaming – Public Shaming in Uncertainty

I. Introduction

Public shaming can be a powerful tool in enforcing valuable social norms and punishing immoral behaviour. And among its many uses, it seems particularly instrumental in responding to instances of systemic injustices like racism or sexism, because it provides the means for sanctioning wrongdoers who might otherwise not be held to account for their misconduct.. But there are significant risks that attach to practices of public shaming. We often face significant epistemic uncertainty about whether a prospective target of public shaming has engaged in wrongdoing. And even if we know that someone has committed wrongdoing, we may not know further facts that bear on the gravity of their wrong. Or we might be certain that someone has committed a wrong, but be unsure about whether *that* kind of wrongdoing justifies public shaming.

What should we do when we are considering whether to publicly shame, but face epistemic uncertainties of these sorts? In this chapter, I aim at making some progress in answering this question, by considering the standards of proof (SOP) that should apply to public shaming – standards which pertain to much and how good evidence you must possess to justify contributing to public shaming. There are surely some SOPs regarding public shaming; you are not justified in heaping moral opprobrium on someone on just anyone if you have no evidence whatsoever that they have committed a wrong. But it is less clear how much evidence you need, or how reliable the evidence must be, in order to justify contributing to public shaming.

Broadly, it seems there are two broad camps on this question. One camp, mainly endorsed by some commentators in the media, is that we should apply very high SOPs – akin to the standards used in criminal law – in this case. For example, in response to the #MeToo movement, commentators have worried that the movement erodes important social protections such as *Presumption of Innocence* or requirement to prove guilt *Beyond a Reasonable Doubt*.²⁵² Effectively, those commentators assert that we ought to apply the high standards of presumption of innocence and proof beyond a reasonable doubt to public shaming. Another camp holds that we do not need to employ such high standards. For example, the lawyer Michael Spratt argues that “To insist on the strict application of the presumption of innocence in everyday life is an absurd and insidious act of complicity to the realities exposed by the #MeToo movement. In no other aspect of our daily lives do we employ the presumption of innocence or apply a burden of proof beyond a reasonable doubt”.²⁵³ Kim Ferzan points out that when we deliberate on something like whether to boycott a conference celebrating the work of someone accused of sexual misconduct, “We may want a lower threshold [of evidence], and one that takes into account the stakes for the complainants if the profession continues to celebrate someone who actually has committed sexual misconduct”.²⁵⁴ Within this camp, some further suggest that a standard as low as the *Preponderance of Evidence* should be used. For example, Jane Gilmore argues that “Legal principles do not have to be enforced

²⁵² For example, see Alpher, Roger, “#MeToo Means the End of the Presumption of Innocence,” *Haaretz*, 8 Nov 2021, <https://www.haaretz.com/opinion/.premium-metoo-means-the-end-of-the-presumption-of-innocence-1.10362331>; O’Neill, Brendan, “What happened to ‘innocent until proven guilty?’” *Chicago Tribune*, 28 Nov 2017, <https://www.chicagotribune.com/opinion/commentary/ct-perspec-innocense-presumption-guilty-metoo-sexual-misconduct-1128-20171127-story.html>.

²⁵³ Spratt, Michael, “The Presumption of Innocence is for Courtrooms, not Politics,” *CBC News*, 30 Jan 2018, <https://www.cbc.ca/news/opinion/presumption-of-innocence-1.4509334>.

²⁵⁴ Ferzan, Kimberly Kessler, “#BelieveWomen and the Presumption of Innocence: Clarifying the Questions for Law and Life,” *All Faculty Scholarship*, 2720 (2021). https://scholarship.law.upenn.edu/faculty_scholarship/2720; For a Journalistic treatment, see Gilmore, Jane, “Presumption of innocence and the #MeToo backlash,” *The Sydney Morning Herald*, 23 Oct 2018, <https://www.smh.com.au/lifestyle/life-and-relationships/presumption-of-innocence-and-the-metoo-backlash-20181019-p50aqw.html>.

outside a courtroom but, even if they did, the standard of proof between two individuals is: who do we believe on the balance of probabilities?²⁵⁵

The goal of this chapter is not to defend a single principle like Beyond a Reasonable Doubt for all cases of public shaming. Indeed, as I shall argue, I do not think that a single principle *should* be used for all cases. Rather, I propose that we change the way we think about the SOP in cases of public shaming, conceiving of it as a framework for deliberation instead of a single principle. Such a framework is a set of principles which together helps individuals and communities deliberate on a set of issues, take into account the relevant stakes, and thus determine if the available evidence in a given case of public shaming is sufficient to justify the public shaming. In serving these ends, the framework can fulfil the function of an SOP for public shaming.²⁵⁶ To defend this set of claims, this chapter has three subordinate goals. First, to describe the function of a standard of proof in cases of public shaming and the kinds of considerations relevant to it. Second, to critique some potential standards, like the Beyond a Reasonable Doubt Standard and the Preponderance of Evidence standard. Finally, to propose, describe, and defend the framework for deliberation which I think should be used as a standard of proof in cases of public shaming.

Here is a layout of the chapter. In part II, I analyse the concept of an SOP, which roughly refers to the evidentiary threshold required to treat a claim as proven for the purposes of action. I narrow the scope of this chapter by making a distinction between objective standards, private deliberative standards, and public deliberative standards (with public deliberative standards being the focus of this chapter). In part III, I consider two broad approaches to developing an SOP: an

²⁵⁵ Gilmore, Jane, “Presumption of Innocence and the #MeToo Backlash”.

²⁵⁶ Some may think this is an overly revisionary take on what a standard of proof can be. Even if they are correct, I do not think that the content of this chapter will be seriously affected, because its central claim is that a framework for deliberation can fulfil the function of an SOP. I have chosen to say that a framework for deliberation can be a kind of SOP. But we could also claim that a framework for deliberation ought to replace SOPs in cases of public shaming. Either one of these two possibilities is acceptable for the purposes of this chapter.

approach based on the relevant moral and practical stakes, and another based primarily on epistemic considerations. I defend the use of the former approach based on moral and practical stakes, and then describe the relevant moral and practical stakes related to a practice of public shaming. In part IV, I critique the Beyond a Reasonable Doubt standard and the Preponderance of Evidence standard. Finally, in part V, I sketch out and defend my key proposal, a framework for deliberation for cases of public shaming.

II. Standards of Proof

A standard of proof (SOP) regarding some act, broadly speaking, is an “evidentiary threshold” that is relevant to the justifiability of an act.²⁵⁷ It is a standard which states how much and how good evidence an agent must possess in order to justify belief in some proposition for the purposes of action. There is an initial distinction we can draw between two different kinds of SOPs. An SOP may refer to an *objective standard* regarding when a person acts permissibly, justifiably, or blamelessly. It is the standard that the objective observer would hold an agent to, in assessing whether the agent acted permissibly/justifiably/blamelessly, given the evidence the agent had. Determining the objective SOP for public shaming is not the goal of this chapter. Another standard, which we may call a *deliberative standard*, refers to a principle or set of principles which should play a role in public or private deliberation. For example, the Beyond a Reasonable Doubt (BARD) standard, used in most cases in criminal law, is a deliberative standard. Juries are instructed on the use of the BARD standard, and it is therefore expected that the BARD standard will feature in their deliberations on whether a defendant is guilty. Deliberative standards are subjective – because they are standards that agents should use in their deliberation, they may be relative to contextual features of a deliberative process, like a specific agent’s biases. Thus, the deliberative

²⁵⁷ I borrow the phrase “evidentiary threshold” from Barry, Christian, in “Applying the Contribution Principle,” *Metaphilosophy* 36 no. 2 (2005): 214.

standard and the objective standard can come apart. For example, I may be objectively justified in assigning a distinction to my favourite student if I am more than 50% sure that their essay deserves it. But given my biases, it may be prudent to hold myself to a higher standard in my deliberation.

We can make a further distinction between two kinds of deliberative standards: private and public. I am interested primarily in the latter. A private deliberative standard is a principle that an individual uses to govern their own reasoning. My earlier example of a standard which tells me whether I should assign my student a distinction is a private standard. A private deliberative standard regarding some type of act might take into account the expected outcomes of the act, the agent's biases, and any inherent moral properties of the act itself. In contrast, a public deliberative standard is a social norm or set of norms concerning how people in certain situations ought to deliberate. An example of a public deliberative standard is the BARD standard. A public deliberative standard has a number of functions. First, it provides instructions to members of a community regarding how they should deliberate. Because of this function, public deliberative standards are related to private deliberative standards. Secondly, it tells members of that community roughly what normative expectations they may have of each other (that is, expectations about how others should behave). Thirdly, it gives those members the grounds to hold each other to account and demand justification when others do not perform according to the standard. Like other social norms, public deliberative standards can be formal (like the BARD standard in criminal law) or informal.

A key difference between private and public deliberative standards is the level of analysis. Private deliberative standards are for propositions relevant to specific activities, like "this student deserves a distinction". In contrast, public deliberative standards apply at the practice-wide level, or across a sub-practice (that is, a subset of activities within a practice which can be meaningfully distinguished from other activities not within the subset). For example, the BARD standard is

applied to the practice of criminal trials as a whole, and agents are expected to use the BARD standard whenever they take on the role of judge or jury. Because public deliberative standards are usually applied at the practice level, they must take into account general information about the types of acts in question, and the general tendencies of people within the relevant contexts. For example, if we wanted to determine the right public deliberative standard for public shaming in cases related to sexual assault, we might take into account the typical harms caused by cases of public shaming in cases of sexual assault, and the typical biases held by potential contributors.

One issue that I will set aside, for the purposes of this paper, is the issue of what counts as evidence in the first place. Broadly speaking, “evidence” refers to any information that justifies your belief in the truth of some claim. But there are interesting debates about how we should count certain kinds of evidence. For instance, it is unclear what the role of testimonial evidence is in some cases related to public shaming.²⁵⁸ There are also well-known debates about the use of certain kinds of non-personal statistical evidence in criminal sentencing that could well apply to the issue of public shaming.²⁵⁹ While these topics are important, going into them will take us too far afield.

IIIa. Stakes-Sensitive Standards vs. Stakes-Insensitive Standards

I now want to describe and defend the approach I will take in developing a public deliberative SOP for public shaming. Broadly speaking, we can take two approaches to developing an SOP for a type of activity. The first approach, which I prefer, involves taking into account the moral and practical stakes involved in a given practice, and setting an SOP with the view of

²⁵⁸ As Ferzan points out, there is some debate between reductionists (about testimony), who argue that testimony is just another piece of evidence in favour of some proposition, and non-reductionists, who hold that testimony gives agents reasons to believe the testimony-giver, not just a proposition. Kessler Ferzan, Kimberly, “#BelieveWomen and the Presumption of Innocence: Clarifying Questions for Law and Life”.

²⁵⁹ For an overview of some key arguments in the literature, see Gardiner, Georgi, “Legal Burdens of Proof and Statistical Evidence,” in *The Routledge Handbook of Applied Epistemology*, eds. David Coady and James Chase (London: Routledge, 2018).

achieving an acceptable set of outcomes given the stakes. For instance, we might take into account all the likely benefits and costs associated with a practice, as well as other moral considerations we deem relevant (like deontic considerations). We would then draw up a list of every plausible SOP for that practice and rank them based on their expected costs and benefits when applied to the practice, and how well they perform with respect to other moral considerations. We should then select the SOP which ranks the highest overall given this set of considerations. We may call this approach the “Stakes-Sensitive Approach” to developing an SOP.

Another possible approach is to base an SOP primarily or entirely on some independent epistemic standard. Under such an approach – a “Stakes-Insensitive Approach” – SOPs are mostly not affected by the moral and practical stakes associated with a practice. This approach may be preferred by people who believe the following two claims:

The Strong Knowledge Standard: Iff an agent knows p, she is justified in acting as if p.

And

Evidentialism: Whether or not an agent knows a proposition depends only on the epistemic evidence available to her.

The Strong Knowledge Standard, together with Evidentialism, yields the conclusion that an agent is justified in acting as if a proposition is true if and only if she has sufficient evidence in favour of the proposition. Moral and practical stakes do not matter.

Several philosophers endorse versions of the Knowledge Standard. Bianca Crewe and Jonathan Jenkins Ichikawa seem to endorse the Strong Knowledge Standard, claiming that “to know something is to have it conclusively settled—or at least, conclusively enough for one’s

practical purposes”.²⁶⁰ Other philosophers defend weaker versions of the standard. One version is that *only if* a person knows that a proposition is true, she is justified in acting as if the proposition is true.²⁶¹ Another version is that if and only if a person knows that a proposition is true, then she is justified in treating that proposition as a reason for action.²⁶² Among the philosophers who defend the Knowledge Standard in some form, several link the Knowledge Standard to punitive social practices. For instance, some have argued that legal proof – that is, the standard that must be met to find a defendant guilty – requires knowledge, or that knowledge is sufficient for legal action.²⁶³ Crewe and Ichikawa apply the knowledge standard to other forms of social punishment, like firing someone accused of sexual misconduct, stating that ““I know that he is guilty, but we shouldn’t punish him because of the chance that he is innocent” is a contradiction”.²⁶⁴

Many philosophers who defend some version of the Knowledge Standard do not have stakes-insensitive approaches to standards of proof. Hawthorne and Stanley, for example, claim that an agent is justified in treating a proposition as a reason for action *only if* she knows it to be true.²⁶⁵ This is compatible with there being other reasons for action which can influence whether an agent has met some evidentiary threshold which determines the justifiability of their actions. Other proponents of the Knowledge Standard reject Evidentialism, supporting views on knowledge which take into account the stakes involved. Moss, for example, argues that the stakes

²⁶⁰ Crewe, Bianca, and Jonathan Jenkins Ichikawa, “Rape Culture and Epistemology,” in Jennifer Lackey (ed.) *Applied Epistemology* (New York: Oxford University Press, 2021), 260-261.

²⁶¹ Fantl, Jeremy and Matthew McGrath. "Evidence, Pragmatics, and Justification". *The Philosophical Review* 111, no. 1 (2002): 78.

²⁶² Hawthorne, John, and Jason Stanley, “Knowledge and Action,” *The Journal of Philosophy* 105, no. 10 (2008): 571-590.

²⁶³ Moss, Sarah, “Knowledge and Legal Proof,” in Tamar Szabó Gendler, John Hawthorne, and Julianne Chung (eds), *Oxford Studies in Epistemology Volume 7: Oxford Studies in Epistemology* (Oxford: Oxford University Press, 2022); Blome-Tillmann, Michael, “More Likely Than Not: Knowledge First and the Role of Bare Statistical Evidence in Courts of Law,” in J. Adam Carter, Emma C. Gordon, and Benjamin Jarvis (eds), *Knowledge First: Approaches in Epistemology and Mind* (Oxford: Oxford University Press, 2017); Littlejohn, Clayton, “Truth, knowledge, and the standard of proof in criminal law,” *Synthese* 197 (2020): 5253–5286.

²⁶⁴ *Ibid.*

²⁶⁵ Hawthorne, John, and Jason Stanley, “Knowledge and Action”.

in a given situation can affect how much evidence a person needs to count as knowing.²⁶⁶ Meanwhile, Crewe and Ichikawa endorse a contextualist theory of knowledge, which is the view that “knowing” is context-sensitive. They claim, “one can count as “knowing” that someone has committed a serious offense in a conversation about how the university ought to respond to it, without counting as “knowing” it in a conversation about whether the state ought to incarcerate the perpetrator”.²⁶⁷ Thus, Hawthorne and Stanley, Moss, Crewe and Ichikawa, among others, could agree that the SOP for a given activity should depend, at least partially, on the moral and practical stakes involved. For the purposes of this chapter, I have no quibble with these views.

My target here is a genuinely stakes-insensitive view, which can be obtained from the conjunction of the Strong Knowledge Standard and Evidentialism. Such a view seems implausible. Across a wide variety of cases, the justifiability of our behaviour for a given amount of evidence seems to depend on the stakes involved. For example, if you want to serve a cheesy dish at your party (assuming that your guests will blindly help themselves to your food), you would be justified in serving the dish as long as you are moderately confident that none of your guests are lactose intolerant. But if you want to serve a nutty cake at your party, you will want to be very sure that none of your guests have a potentially fatal nut allergy (again assuming that your guests help themselves blindly to your food). And it would be unjustifiable to serve nuts unless you were very confident that none of your guests have a nut allergy and may accidentally consume the cake. Why would you need to be so much more confident in the second case than in the first? Because the stakes involved with a potentially fatal nut allergy are so much greater than the stakes involved with lactose intolerance.²⁶⁸ Or again, imagine that you are seeking directions to get to a café to meet a friend whom you promised to meet. You would probably be content with getting directions

²⁶⁶ Moss, Sarah, “Pragmatic encroachment and legal proof”.

²⁶⁷ Crewe, Bianca, and Jonathan Jenkins Ichikawa, “Rape Culture and Epistemology,” 15.

²⁶⁸ Moss provides a set of similar examples in “Pragmatic encroachment and legal proof”.

from one or two people. But if you needed directions to get to the hospital to perform life-or-death surgery, you would have to gather more evidence to be very sure that your directions were correct, if you have the time to gather more evidence. We could think of many more examples like these. The point is, the justifiability of such acts – to serve a dish or not; to act on a belief in a direction; etc. – seem to depend on the stakes involved.

So it seems that the objective permissibility of an act ought to depend on the moral and practical stakes involved. Does this mean that the public deliberative SOP for social practices, like public shaming, also needs to be stakes-sensitive? A public deliberative SOP is a set of social norms. In most cases, we want social norms to roughly guide people to do what they are in fact justified in doing. Thus, in most cases, we want a deliberative SOP for an act to roughly converge with the objective SOP for the act (except in exceptional cases, such as in urgent cases where a deliberative process can result in more harm than good). Additionally, as I noted earlier, public deliberative SOPs are constituted by norms which shape our expectations of each other, and with which we hold each other to account. In most cases, we should have normative expectations of each other which roughly match what we are objectively justified in doing, and we should also hold each other to account in ways that reflect what we are objectively justified in doing.

Thus, my view is that the public deliberative SOP for public shaming ought to depend on the moral and practical stakes involved in cases of public shaming. A practice of public shaming which is riskier in terms of the potential harms it can cause than another practice (all else being equal) should have a higher SOP associated with it. Conversely, a practice which is not as risky, or can be expected to bring about more significant relevant benefits, may have a lower SOP. This is not to say that epistemic considerations do not matter – as we shall see later, I do think that SOPs ought to take into account epistemic considerations, like an agent's biases.

IIIb. The Relevant Stakes

So what are the relevant moral and practical stakes for an SOP for public shaming? Roughly, there are three categories of stakes: First, the relevant *interests* of people involved in and affected by a case of public shaming. For organization purposes, I will address these interests in two subsections: interests which generally militate in favour of relatively high SOPs, and interests which militate in favour of low SOPs. Second, the moral stakes pertaining to the *deontic status* of acts of public shaming. Third, *correctives* that ought to be applied for the purposes of practical reasoning.

Interests in Favour of High SOPs

To begin with, let us consider some of the interests which militate in favour of higher SOPs. Because a relatively high SOP makes it harder for people to justify public shaming, it can be expected to result in fewer cases of wrongful public shaming (where targets are not in fact liable to be publicly shamed), and can generally be expected to result in less shaming in total. As a result, the interests that militate in favour of high SOPs are generally interests which are more likely to be infringed with higher rates of wrongful public shaming, and those that are more likely to be infringed with high levels of public shaming in general.

Most or all human beings have interests against being publicly shamed. As covered in previous chapters, public shaming causes risks of serious health effects, loss in valuable social interactions and opportunities, and causes risks of other independent sanctions like being fired from one's job, among other potential harms. Furthermore, being shamed can impose costs on third parties, like the target's family and friends. A study on the families of registered sex offenders found that those family members experienced "stigma, social isolation, harassment, loss of employment, financial difficulties, and loss of familial relationships".²⁶⁹ It is plausible that families

²⁶⁹ Bailey, Daniel S., and Lisa L. Sample, "An Examination of a Cycle Of Coping with Strain Among Registered Citizens' Families," *Criminal Justice Studies* 30, no. 2 (2017): 167.

and friends of targets of other cases of public shaming sometimes face similar difficulties. Given that most members of society can be accused of wrongdoing, everyone faces some risk of bearing the costs of being publicly shamed, although various groups can carry heavier risks, depending on the wrongdoing in question. For example, because members of some racial minorities are more often accused of violent crimes and drug related offences, members of such groups face these heavier risks of being publicly shamed for such crimes.²⁷⁰ Thus, everyone, to varying extents, has some interest in there being a high SOP for public shaming, which minimizes the risk that they will suffer the harm associated with public shaming.

But more needs to be said about how we factor these interests into an account of the moral stakes of public shaming. The moral significance of a person's interests against being publicly shamed can be weighted by some factors. One such factor is a target's liability to being publicly shamed. When a target is liable to be publicly shamed, we can discount their interests (partially or totally, depending on how much they are liable to suffer) against being publicly shamed for the purpose of weighing those interests against other stakes. As we saw in previous chapters, there are different possible explanations for how this discounting works. Many retributivists claim that it is good for deserving wrongdoers to suffer, at least to some extent.²⁷¹ Other philosophers claim that wrongdoers (of certain kinds) forfeit their rights against being harmed in certain ways.²⁷² In either case, we get similar results – that it is not necessarily wrong to impose harm on wrongdoers who

²⁷⁰ Racial minorities may also be more likely to be judged as more morally responsible for wrongdoing, and thus more deserving of sanctions. For example, Goff et al. found that Black teenagers were less likely to be judged as being less innocent, and more responsible for their actions than their White peers. See Goff, P. A., Jackson, M. C., Di Leone, B. A., Culotta, C. M., & DiTomasso, N. A., "The Essence of Innocence: Consequences of Dehumanizing Black Children," *Journal of Personality and Social Psychology* 106, no. 4 (2014): 526–545.

²⁷¹ Walen, Alec, "Retributive Justice," *The Stanford Encyclopedia of Philosophy*, eds. Edward N. Zalta & Uri Nodelman, 2003, <https://plato.stanford.edu/archives/win2023/entries/justice-retributive/>.

²⁷² Wellman, Christopher, "The Rights Forfeiture Theory of Punishment," *Ethics* 122, no. 2 (2012): 371–393.

are liable to be harmed. Thus, we get the result that human beings have morally-weighted interests against being *wrongfully* publicly shamed, but not necessarily against being publicly shamed.

Another factor is the expressive nature of public shaming, which compounds the many costs associated with public shaming. A number of philosophers, writing on criminal punishment, have argued that because punishment comes with condemnation, it imposes an extra cost on the accused in addition to the other ways they are harmed. Barry claims, “Because of its strong expressive function, falsely accusing an agent of criminal offences carries an additional cost . . . Because of these costs, we do not typically want to imprison someone for [for a crime] unless we are confident” that they committed the crime.²⁷³ A similar claim is made by Duff, who writes that “To be convicted of a crime is not just to suffer some loss; it is to be condemned for committing a wrong. This gives us reason to make it quite hard to convict defendants, by defining proof “according to law” as proof beyond reasonable doubt”.²⁷⁴ Barry’s and Duff’s arguments can be applied to public shaming, given that public shaming usually involves some kind of condemnation. Their arguments imply the morally-weighted interests that people have against being wrongfully publicly shamed become greater once we take into account the condemnatory nature of public shaming.

Apart from the interests that people have against being the targets of public shaming, people have interests that can be negatively affected by the practice of public shaming even when they are not themselves targets. Most people have interests in pursuing valuable activities which can be negatively affected by public shaming. When there is too much or too frequent public shaming in society, it has a chilling effect on potentially valuable activities, such as some social

²⁷³ Barry, Christian, “Applying the Contribution Principle,” 217-218.

²⁷⁴ Duff, Antony, Lindsay Farmer, Sandra Marshall, and Victor Tadros, *The Trial On Trial Vol 3: Towards A Normative Theory Of The Criminal Trial* (Oxford: Hart Publishing, 2007), 89.

activities and potentially valuable speech.²⁷⁵ This is particularly the case when people perceive that certain activities carry excessive risk of being wrongfully accused and shamed. For example, imagine again that we set a very low SOP for public shaming in sexual assault cases, which in turn risks a significant number of innocent people getting publicly shamed for sexual assault. Because a disproportionate number of offenders are men, this might disincentivize men from spending time alone with women (and perhaps other men), lest they get accused of sexual assault. This result places risks on various people, as it prevents both men and women from pursuing some valuable social interactions. In societies where men tend to occupy more positions of power, disincentivizing male-female interactions can in fact place additional costs on women, as women get excluded from meetings, networking events, and mentoring.²⁷⁶

Interests in Favour of Low SOPs

On the other hand, people also have interests which are more likely to be fulfilled when it is easier to justify public shaming. So people have interests which militate in favour of lower SOPs. The first of these interests is the interest that all (or most) people have in enforcing valuable norms. As we saw in chapter III, public shaming can enforce norms by deterring future norm violations, expressing society's commitment to important norms, and by communicating disapproval and censure. When an SOP is too high, public shaming becomes harder for agents to justify, which may result in a large number of wrongdoers going unpunished. This can in turn fail to

²⁷⁵ This point has been raised by Guy Aitchison and Saladin Meckled Garcia, who write that "It is plausible that the mere threat of OPS can have a chilling effect on free speech. Faced with potential public, mass ostracism there is the danger that people self-censor and avoid discussing unpopular opinions, for the kinds of reasons famously identified by J. S. Mill." "Against Online Public Shaming: Ethical Problems with Mass Social Media," *Social Theory and Practice* 47, no.1 (2021): 6.

²⁷⁶ Johnson, W. Brad, and David G. Smith, "Men Shouldn't Refuse to Be Alone with Female Colleagues," *Harvard Business Review*, 5 May 2017, <https://hbr.org/2017/05/men-shouldnt-refuse-to-be-alone-with-female-colleagues>. One reported side-effect of the #MeToo movement is precisely that it reduces opportunities for women. Borysenko, Karlyn, "The Dark Side Of #MeToo: What Happens When Men Are Falsely Accused," 12 Feb 2020, <https://www.forbes.com/sites/karlynborysenko/2020/02/12/the-dark-side-of-metoo-what-happens-when-men-are-falsely-accused/?sh=465d6796864d>.

disincentivise wrongdoing of a similar sort in the future. The routine failure to punish wrongdoers with public shaming can also have a negative expressive effect. It may communicate the notion that society condones or tolerates bad behaviour. Or at least, it may communicate that society is somewhat unwilling to punish bad behaviour. Failing to punish wrongdoers is generally bad for most people, but particularly for those most at risk of becoming victims of the types of wrongdoing targeted by public shaming.

Many individual agents can have interests not only in public shaming, but in contributing to public shaming (that is, not only that public shaming happens, but that *they* contribute). First, many people have interests in publicly expressing views and opinions that are important to them. If you believe that someone has done something seriously wrong, it can be important to you to express your condemnation of that wrongdoing via public shaming. Second, a closely related interest is a person's interest in shaping her society. Each of us wants to see our society maintain certain practices and norms and evolve to include desirable practices and norms. We may also have interests in being involved in bringing about those practices and norms. For example, many of us rightly desire that our society become one which is less tolerant of racist or sexist behaviour. We may have an interest in taking part in activities which discourage such behaviour, like publicly shaming people for making racist or sexist jokes. Third, as I argued in chapter IV, public shaming can be used to defend one's moral standing – one's status as an equal in society as perceived by others and oneself. By engaging in public criticism, victims of wrongdoing, or those acting on their behalf, can affirm and defend their moral standing. Thus, victims of wrongdoing have agent-relative interests in contributing to public shaming.

There are further, irrelevant goods that individuals can attain by contributing to public shaming, which would be harder to achieve if the SOP for public shaming was high. These include the satisfaction one might gain from publicly criticizing an unlikeable person, the status one might

gain from “moral grandstanding”,²⁷⁷ or the feeling of vindication one might feel from striking down someone in a powerful position. I call these goods irrelevant goods because I do not think they should count, for the purposes of justifying the imposition of significant harms on others.

Deontic Considerations

In addition to the interests held by people, there are deontic considerations relevant to public shaming. Normally, an act which violates a deontic principle is taken to be more wrongful than an equivalent act (in terms of the consequences caused by both acts) which does not violate any deontic principles. Because agents can violate deontic principles by contributing to wrongful public shaming, deontic considerations generally militate in favour of high SOPs which make wrongful public shaming less likely.

One deontic consideration is the Doing/Allowing distinction, which holds that it is morally worse for an agent to do harm than allow an equivalently costly harm to occur, all else being held constant. This distinction has been used in the criminal justice literature to justify a high SOP. Some legal theorists have argued that because of the Doing/Allowing distinction, a state should prefer a high standard of proof which results in more false acquittals than false convictions, even if such a standard results in more people getting harmed by criminals.²⁷⁸ Most famously, William Blackstone proclaimed that “it is better that ten guilty escape than that one innocent suffer”.²⁷⁹

Similarly, the Doing/Allowing distinction militates in favour of a higher SOP with respect to public shaming. Since publicly shaming someone is clearly an instance of doing harm, we might

²⁷⁷ Justin Tosi and Brandon Warmke define moral grandstanding as the attempt to “get others to make certain desired judgments about oneself, namely, that one is worthy of respect or admiration because one has some particular moral quality”. Tosi, Justin, and Brandon Warmke, “Moral Grandstanding,” *Philosophy and Public Affairs* 44, no. 3 (2016): 199.

²⁷⁸ See Hosein, Adam Omar, “Doing, Allowing, and the State,” *Law and Philosophy* Vol. 33, No. 2 (March 2014), pp. 235-264

²⁷⁹ Blackstone, William, *Commentaries on the Laws of England: In Four Books* (Philadelphia: J. B. Lippincott, 1753).

think that agents should generally avoid contributing to a case of public shaming, unless they are confident that the target of public shaming is liable to be harmed. As a result, agents should use a high SOP to avoid imposing harm on non-labile targets, unless the costs of using such a high SOP are massive enough to justify a higher risk of causing harm to non-labile targets.

Another deontic principle relevant here is the Means Principle, which “forbids actors from achieving otherwise good consequences by using, without their consent, others’ bodies, labor, talents, or the resources that are rightfully theirs”.²⁸⁰ When we punish someone in order to deter other agents from offending, or to express condemnation for a kind of behaviour, we use them as a means to reduce overall crime. According to the Means Principle, this is impermissible. To get around this, it has been suggested that while deterrence (and other instrumental goods) can be the goal of punishment, an offender’s moral desert is a necessary condition to justify punishment. As Alec Walen puts it, “Desert functions like a gate in the normative equivalent of a transistor; until it is switched on, the potentially greater normative force of the instrumental reasons is switched off”.²⁸¹ Because of the “transistor-like function” of desert, the “the instrumental benefits of punishment must be screened off until the prosecution has established guilt”.²⁸² The instrumental benefits of punishment cannot justify punishment unless a person deserves to be punished, thus we have to establish a defendant’s guilt prior to taking into account those benefits. Thus, Walen claims that a high SOP must be used in criminal punishment, because a state is only justified in punishing a person if it is confident that she deserves to be punished. Because we cannot justify punishment on the grounds of instrumental benefits until we establish guilt, we need a high SOP like the BARD standard.

²⁸⁰ Alexander, Lawrence, “The Means Principle,” *Legal, Moral, and Metaphysical Truths: The Philosophy of Michael Moore*, eds. K. K. Ferzan and S.J. Morse (Oxford: Oxford Univ. Press, 2014), San Diego Legal Studies Paper No. 14-138, Available at SSRN: <https://ssrn.com/abstract=2378608>.

²⁸¹ Walen, Alec, “Proof Beyond a Reasonable Doubt,” *Louisiana Law Review* 76 (2015): 427.

²⁸² *Ibid.*

A similar idea can be applied to public shaming. There are certainly instrumental benefits to public shaming. But if Walen is right that desert plays a “transistor-like function”, then a number of these benefits should not count as relevant unless the target of public shaming is liable to be shamed. To justify a case of public shaming, we first need to prove that a target is liable to be shamed. As a result, the Means Principle, or Walen’s application of it, implies that we ought to use a high SOP in order to ensure that the target is liable in the first place. It may be unfortunate that this ends up putting a high burden on individuals who can benefit from public shaming, or those who suffer a cost from the failure to publicly shame. But the Means Principle prohibits us from using possible wrongdoers as a means to benefit those individuals, unless we can establish the guilt of those wrongdoers.

Biases and Correctives

A final set of considerations relevant for an SOP is bias. Generally, human agents tend to have biases in the way they reason about issues. Common biases include confirmation bias,²⁸³ the sunk cost fallacy,²⁸⁴ and the anchoring effect.²⁸⁵ These biases result in patterns. For example, confirmation bias (which involves weighing heavily evidence that supports your beliefs or desires, while discounting evidence against it) reliably results in individuals being more confident in a belief that they held at the beginning of a deliberative process than someone who did not hold the initial belief.²⁸⁶ Because of confirmation bias, when a group of people share an initial belief in some

²⁸³ Dwyer, Christopher, “12 Common Biases That Affect How We Make Everyday Decisions,” *Psychology Today*, 17 Sep 2018, <https://www.psychologytoday.com/us/blog/thoughts-on-thinking/201809/12-common-biases-that-affect-how-we-make-everyday-decisions>.

²⁸⁴ *Ibid.*

²⁸⁵ Furnham, Adrian, and Hua Chu Boo, “A literature review of the anchoring effect,” *The Journal of Socio-Economics* 40, no. 1 (2011): 35-42.

²⁸⁶ Rassin, Eric, Anita Eerland, and Ilse Kuipers, “Let’s Find the Evidence: An Analogue Study of Confirmation Bias in Criminal Investigations,” *Journal of Investigative Psychology and Offender Profiling* 1.7 (2010): 231–246.

proposition, they are more likely to maintain this belief when presented with evidence, even evidence to the contrary.

A corrective principle is a rough principle or guideline which seeks to counteract or neutralize biases.²⁸⁷ Such a principle can give individual agents heuristics for arriving at more accurate or justifiable conclusions when they are aware that they may be susceptible to biases. An example of a corrective, in light of confirmation bias, might be “weigh evidence against your beliefs more heavily than you otherwise would”. Correctives can also play roles in public deliberation or discourse; different members of a community may expect each other to use correctives, use correctives to adjust claims and statements made by one another, or have corrective social norms which simply counterbalance the effect of common biases. In cases of public shaming, it seems that many biases make potential contributors more likely to believe that a target ought to be publicly shamed. Potential contributors to public shaming likely want to contribute to public shaming in the first place, and believe that a target ought to be shamed. Common biases like confirmation bias and the anchoring effect will thus generally make potential contributors more likely to believe that the target is liable to be shamed. A high SOP for cases of public shaming can therefore function as a corrective principle. Because a high SOP is a public deliberative standard, prospective contributors will be expected to muster a lot of evidence in favour of a target’s liability. This makes them less likely to contribute to public shaming, potentially counteracting the effect of confirmation bias.

That said, there are other biases that make it less likely in some cases that a prospective public shamer will contribute. Agents typically have biases against the testimony of historically

²⁸⁷ The idea of a “corrective move” comes from Miranda Fricker, who writes that “The guiding ideal is to neutralize any negative impact of prejudice in one’s credibility judgments by compensating upwards to reach the degree of credibility that would have been given if not for the prejudice”. Fricker, Miranda, *Epistemic Injustice: Power & the Ethics of Knowing* (Oxford: Oxford University Press, 2007), 91-92.

marginalized groups of people, like women, racial minorities, and those who speak with certain accents.²⁸⁸ That is, we tend to have lower credibility in the testimony of members of these groups. As a result, in certain cases of public shaming, prospective shamers are less likely to contribute when the evidence in favour of contributing is supplied by members of those groups. For example, women accusing men of sexual assault are less likely to be believed than they probably should. Correctives can also play a role in such cases where public shaming is less likely to be considered justified because of biases against testimonial evidence. In principle, a relatively low SOP can function as a corrective, because the low SOP makes it easier for a set of evidence to count as sufficient to justify public shaming. This potentially counterbalances the biases held against that set of evidence.

IV. Two SOPs – BARD and POE

(Note: the following section extensively discusses cases of public shaming related to sexual assault and rape.)

I have described my approach to developing an SOP for public shaming, and the kind of moral stakes and considerations I think are relevant. With this in our pocket, we can finally turn our attention to two potential SOPs people have suggested we use in cases of public shaming or other social punishments: the Beyond a Reasonable Doubt (BARD) standard and the Preponderance of Evidence (POE) standard. I will argue that neither of these standards be used as public deliberative standards for the practice of public shaming as a whole - that is, used in every case of potential public shaming.

Beyond a Reasonable Doubt

²⁸⁸ Fricker, *Epistemic Injustice: Power & the Ethics of Knowing*, 30-40.

Let us begin with BARD. As the name suggests, the BARD standard requires that to justify some act, the agent must have enough evidence that it is unreasonable to doubt that the relevant proposition is true. Naturally, this is a very stringent standard – as stringent as a standard can get short of demanding certainty. It is unclear if further precisifications of the BARD standard are possible or desirable.²⁸⁹ Nevertheless, in legal proceedings, the BARD standard is often interpreted as a high standard which requires over >90% confidence in the guilt of a defendant.²⁹⁰

How might a proponent of the BARD standard defend the use of the standard in cases of public shaming? They might claim that the costs associated with cases of wrongful public shaming are so severe that they greatly exceed the positive interests people typically have in public shaming.²⁹¹ A practice of public shaming can be expected to result in some cases of wrongful public shaming. Hence, the expected costs associated with a practice of public shaming over time will be unacceptably high unless the practice is set up to ensure a very low rate of cases of wrongful public shaming. The BARD standard should help ensure this low rate and will therefore be required.

We can think of this argument as having two versions. A first, more precise version is that given the high costs of wrongful public shaming, the total expected harm of a practice of public shaming will exceed its expected benefits, unless a very high SOP is used. For such an argument to succeed, it is not sufficient to show that the costs of a case of wrongful public shaming exceed the benefits of a rightful case of shaming. The expected harms and benefits of a practice of public

²⁸⁹ Diamond, Henry A., "Reasonable Doubt: To Define, Or Not To Define," *Columbia Law Review* 90, no. 6 (1990): 1718.

²⁹⁰ Walen, Alec, "Proof Beyond a Reasonable Doubt: A Balanced Retributive Account".

²⁹¹ Of course, even rightful cases of public shaming can come with costs. But we shall set this aside for now. One reason is for a simpler, more streamlined discussion. Another reason is that of redundancy – if the costs associated with rightful shaming are also so severe that they exceed the positive interests people have in shaming, then public shaming may be unjustified regardless of whether it's rightful or wrongful. This makes the whole discussion on SOPs redundant.

shaming is also a function of the number of wrongful cases of public shaming, in comparison to the number of rightful cases. This number of wrongful cases of public shaming can in turn be affected by several factors, like whether many people are wrongfully accused of wrongdoing in the first place. In the notes section of this chapter, I model the expected value of potential practices of public shaming, given different SOPs. Because I do not have the space to discuss the model here, I shall set aside this particular argument for now.

The other, less precise version of the above argument is just that because the potential costs of wrongful shaming are so large, we cannot justify the risk associated with those costs unless we minimize the chances of wrongful shaming. But this argument oversimplifies things. In determining whether we carry out any act, we must also consider the potential benefits of the act, or the costs of failing to carry it out. While the costs associated with wrongful public shaming are indeed severe, the costs of failing to publicly shame in some cases can also be very severe. Consider public shaming with regard to a particularly egregious type of wrongdoing, like rape. Because rape tends to impose a serious set of harms on its victims, people have very strong interests in deterring rape. When we publicly shame people for rape, we presumably deter some instances of rape from happening. Therefore, public shaming for rape cases protects some very important interests. These interests are comparable to (if not exceed) the competing set of interests – that is, the interests people have against being publicly shamed for rape. Thus, the risks of failing to publicly shame in some cases, when shaming is in fact rightful, seem comparable to the risks associated with wrongful public shaming. When we insist on the BARD standard, we risk putting too much weight on the risks associated with wrongful public shaming, and too little weight on the potential costs of failing to publicly shame.

Another factor we must take into consideration is the marginal value of each case of public shaming. For example, if a case of public shaming can be expected to deter five rapists from

committing rape, then that case of public shaming seems more justified than if it was only expected to deter 0.1 rapists from committing rape. This observation reveals another argument someone might make in favour of the BARD standard. A fan of the BARD standard might claim that the expected marginal value of *any given* case of public shaming is relatively small. Based on this, they might argue that we are required to use a high SOP for public shaming, because unless a target is very likely to be guilty, the expected value associated with a given case of public shaming is insufficient to outweigh the expected costs.

Is this claim about the small marginal value of any case of public shaming correct? Unfortunately, there no empirical evidence, to my knowledge, regarding the expected marginal value of individual cases of public shaming. Thus, I think it is uncertain if the above argument justifies the use of the BARD standard. That said, as I observed in earlier chapters, the marginal value of cases of public shaming depends on contextual factors regarding how society responds to various types of wrongdoing. When a type of wrongdoing is severely and reliably punished, either by formal institutions or by informal methods, the marginal value (in terms of deterrence) of each case of public shaming is likely to be low. When a type of wrongdoing is not reliably punished, or is punished too lightly, the marginal value of each case of public shaming is likely to be higher. Let us consider the #Metoo movement as an example, which includes a set of cases of public shaming in response to a perceived failure to severely and reliably punish cases of sexual assault and harassment. There are no studies, to my knowledge, on the marginal effects of each case of public shaming within the #Metoo movement. Nevertheless, a number of studies have associated the #MeToo movement with increased knowledge and awareness of sexual assault, increased reporting of sexual assault, and some reported decrease in experiences of sexual harassment.²⁹²

²⁹² Jaffe AE, Cero I, and DiLillo D., “The #MeToo Movement and Perceptions of Sexual Assault: College Students' Recognition of Sexual Assault Experiences Over Time,” *Psychol Violence* 11, no. 2 (2021): 209-218; Levy, Ro’ee, and Martin Mattson, “The Effects of Social Movements: Evidence from #MeToo,” *SSRN.com*,

Thus, there is at least some evidence which should make us wary of the claim that the marginal value of any given case of public shaming is too small to justify a lower SOP than BARD.

A proponent of the BARD standard could reply that the role of a public deliberative standard is not to ensure a good outcome in every case of public shaming, but to ensure that a practice of public shaming as a whole delivers good outcomes. Thus, the fact that a very high SOP can result in bad outcomes in some types of cases is not necessarily a mark against the SOP. They might further point out that I have unfairly focused on a small and unrepresentative set of cases which are unusually beneficial. Not every case of public shaming secures very important interests like interests in deterring rape. However, I believe that the cases I have focused on are not dramatic outliers. The use of public shaming to protect very important interests like interests against rape is hardly an exception. Public shaming (more specifically, mass online public shaming) is frequently used to enforce norms against rape, sexual offenses against minors, hate speech, violence against minority groups, and recently, acts which risk public health. Given such cases, it seems plausible that the BARD standard for the practice of public shaming will result in very high costs due to all the harm that we would fail to prevent. It is not obvious if the expected benefits of the BARD standard, in terms of lowering the risks of wrongful public shaming, are sufficient to justify these costs. Thus, it is unclear if a high SOP applied to the practice of public shaming as a whole will deliver a good outcome.

Furthermore, even if it is true that the BARD standard yields good outcomes when applied to the practice of public shaming as a whole, it doesn't follow that the BARD standard should be used for every case of public shaming. In principle, an SOP for a given activity can involve variable factors, such that the evidentiary threshold varies from case to case. For example, consider the

27 Apr 2023, <http://dx.doi.org/10.2139/ssrn.3496903>; Johnson, Stephanie K, Ksenia Keplinger, Jessica F. Kirk, and Liza Barnes, "Has Sexual Harassment at Work Decreased Since #MeToo?", *Harvard Business Review*, 18 Jul 2019, <https://hbr.org/2019/07/has-sexual-harassment-at-work-decreased-since-metoo>.

principle: “Use the BARD standard, except in cases of sexual assault. In such cases, use a lower standard”. Indeed, my view is that an SOP for public shaming should be sensitive to a number of variable factors, and I shall return to this issue in the next section.

A final argument a proponent of the BARD standard might make is one based on deontic considerations. As I noted in the last section, principles like the Doing/Allowing distinction and the Means Principle militate in favour of higher SOP. When a prospective act violates a deontic principle, it is morally worse, or harder to justify. Wrongful public shaming, which harms someone who is not liable to suffer those harms, violates one or more deontic principles. Thus, we should avoid public shaming unless we can be sure that public shaming will not be wrongful.

I am skeptical of this argument. Because of the sheer variety of cases of public shaming, and because of difficulties in weighing deontic considerations against other morally relevant costs and benefits, it is hard to demonstrate that deontic considerations justify the BARD standard for the practice of public shaming as a whole. As I have argued, public shaming is sometimes used to protect very important interests, and failing to public shame can come with serious costs. In light of this, it’s not clear if the risks of wrongful public shaming justify the BARD standard, even taking into account deontic principles. Furthermore, there is a potential dilemma for proponents of this specific argument. To see the dilemma, consider the question: how much worse is the violation of a deontic principle than a similar act which does not violate a deontic principle? One answer to this question is that violating a deontic principle might only make an act a bit worse, or unjustifiable, than it would otherwise be. For example, perhaps causing harm is only a bit worse than allowing an equivalent harm to occur. But if this is the case, then a case of wrongful public shaming is made only a bit worse by the fact that it violates some deontic principle/s. In this scenario, it seems less likely that those deontic considerations will tip the balance of stakes enough to favour the BARD standard in every case of public shaming.

On the other horn of the dilemma, the disvalue attached to violating a deontic principle could be very high – for example, causing harm might be ten times as bad as allowing an equivalent harm. This will justify the use of high SOPs in all or most cases of public shaming, but it can also prove too much and lead to other less plausible results. For example, doctors and emergency workers often inflict some harm in order to reduce overall harm – examples include limb amputations, surgeries, any drug with undesirable side-effects, assisted dying, etc. If causing harm is much worse than allowing harm, then some or all of these activities may be impermissible. Or for another example, consider paradigmatic cases of self-defence, where a defender may kill a liable aggressor in order to save their own life. Such cases are particularly relevant to cases of public shaming, because I previously argued in chapters III and IV that some cases of public shaming are justified because they are forms of self-defence. If causing undeserved harm is much worse than allowing harm, then a defender might be required to use the BARD standard to justify defending themselves. That is, they may need to have no reasonable doubts that their life is at risk and the aggressor is liable. This conclusion strikes me as implausible. Refraining from defensive harm is to potentially surrender one's life, and it seems overly demanding to expect defenders to refrain from defensive harm unless the BARD standard is met.

Preponderance of Evidence

Let us now consider the Preponderance of Evidence (POE) standard, which requires a deliberator to be more certain than not that a course of action is the right one. In tort law, where the POE is usually used, the standard requires a degree of certainty over 50% that the defendant is liable.²⁹³ While the POE standard is largely accepted in tort law, it has more recently and famously applied to cases of campus sexual assault, with the US's Department of Education's Office for

²⁹³ Demougin, Dominique, and Claude Fluet, "Preponderance of Evidence," *European Economic Review* 50, no. 4 (2006): 964.

Civil Rights instructing departments to use the POE standard in such cases in 2011 (those instructions have since been rescinded).²⁹⁴ Some people have also suggested that the POE standard is appropriate for other practices, including practices of social punishment like public shaming. For example, Gilmore, who claims that “Legal principles do not have to be enforced outside a courtroom but, even if they did, the standard of proof between two individuals is: who do we believe on the balance of probabilities?”²⁹⁵

A key claim made by proponents of the POE standard is that it ought to be used in situations where there are roughly symmetrical interests. For example, proponents of the POE standard in cases of campus sexual assault point out that if an assault has occurred and if the school administration fails to take action, the victim of the assault will suffer in a way that is comparable to the suffering of someone punished for sexual assault. Deborah Brake asserts that victims of sexual assault suffer harms when institutions fail to respond adequately to their claims, writing that “Many survivors experience such institutional betrayal that they feel they have no choice but to leave their institutions, either by transferring to another school, often with loss of academic credit and money, or dropping out of college entirely”.²⁹⁶ Similarly, Amy Chmielewski details that female victims of sexual assault are more likely to experience post-traumatic stress disorder, depressive episodes, substance abuse, and drops in grades.²⁹⁷ Victims often drop out of school “primarily because they do not want to encounter the alleged perpetrator who is still on campus or because they feel betrayed by a lack of support from the administration”.²⁹⁸ On the other hand, if the school

²⁹⁴ Office for Civil Rights, Dep't of Education, *Dear Colleague Letter*, 4 Apr 2011, retrieved 27 Dec 2023, <http://www2.ed.gov/about/offices/list/ocr/letters/colleague-201104.pdf>.

²⁹⁵ Gilmore, Jane, “Presumption of Innocence and the #MeToo Backlash”.

²⁹⁶ Brake, Deborah L., “Fighting the Rape Culture Wars through the Preponderance of the Evidence Standard,” *Montana Law Review* 78, no. 1 (2017): 113.

²⁹⁷ Chmielewski, Amy, “Defending the Preponderance of the Evidence Standard in College Adjudications of Sexual Assault,” *Brigham Young University Education and Law Journal* 3, no. 1 (2013): 166-167.

²⁹⁸ *Ibid.*

administration takes action against the accused, he will also suffer greatly. Thus, both parties potentially have something very significant to lose. Lavinia Weizel notes that “Lower evidentiary standards, such as the common civil standard of preponderance of the evidence, are applied in cases in which the relative harm to society of an erroneous finding is approximately equal no matter whether the error favors the plaintiff or the defendant”.²⁹⁹ In contrast, a higher SOP is appropriate when “a false finding of guilt would result in a particularly egregious deprivation of individual rights, such as in criminal trials”.³⁰⁰

According to defenders of the POE standard, alternative SOPs would be unacceptable because they reflect a preference for one side. Brake argues that “Permitting a complainant to prevail upon a POE standard reflects the absence of an empirical judgment favoring either side. Likewise, the POE standard presumes that the stakes of winning and losing are comparable and that both sides have important interests at stake”.³⁰¹ The thought here is that if both sides in a contest have roughly equal interests at stake, we should not prefer one side over the other. We should assign the costs to whomever we judge to be more likely to be the rightful bearer of those costs. Thus, we should adopt an evidentiary standard that allows us to assign costs to the party more likely to be the rightful bearer of costs.

The above argument is an argument from rough symmetry. But just because two sets of competing interests are similar in strength, doesn’t mean that an agent should treat them as symmetrical. There are a number of other moral considerations that could be relevant. Firstly, when we publicly shame a liable target, the costs we impose on them can be morally discounted -

²⁹⁹ Weizel, Lavinia M., "The Process that is Due: Preponderance of the Evidence as the Standard of Proof For University Adjudications of Student-on-Student Sexual Assault Complaints," *Boston College Law Review* 53, no. 4 (2012): 1632.

³⁰⁰ *Ibid.*

³⁰¹ Brake, Deborah L., "Fighting the Rape Culture Wars through the Preponderance of the Evidence Standard," 133.

that is, it may not be morally bad that they suffer some costs. Thus, even in cases where both sides have equally important interests at stake, the morally-weighted *expected costs* for both parties may not be the same if one of those parties is potentially liable. This is because a calculation of the morally-weighted expected costs for a person takes into account the likelihood that they are liable. Given this, the morally-weighted expected costs of the decision for the potentially liable party may be lower (from the perspective of the decision-maker).³⁰² Thus, even when two sets of competing interests are symmetrical, those interests in terms of morally weighted expected value may not be symmetrical. Agents need not be indifferent when this is the case, but may be required to use an SOP which reflects the asymmetry.

Secondly, the moral value of an act can be affected by deontic principles like the Doing/Allowing distinction. Even if the competing sets of interests in a given situation are roughly equal, the choice we face may not in fact be a symmetrical one. If fulfilling one set of interests requires the violation of some deontic principle, then perhaps we should favour one side over the other, all other things being equal. In public shaming, this means that agents should prefer not to publicly shame when the interests in favour and against a case of public shaming are exactly equal.

Finally, it seems unlikely that the morally-weighted interests of both sides in cases of public shaming will be symmetrical in many cases of public shaming. Among them, because the marginal benefits of a case of public shaming can vary wildly based on the case, the morally-weighted interests can also vary wildly. As I argued in a previous chapter, the marginal benefit of a case of public shaming can vary based on whether the norm in question has been successfully enforced.

³⁰² For example, say that public shaming imposes a set of costs C on a target. There is a 50% chance that the target is in fact liable to experience those costs. Assuming that the costs for a liable wrongdoer can be fully discounted, the morally-weighted expected cost of publicly shaming the wrongdoer is $0.5C$. In contrast, let's say that *failing* to publicly shame the target results in a community suffering some cost D . Given that the community is not liable to experience the cost, there is no chance that D will be discounted in the way C is discounted for the target. Thus, the expected cost for the community may not be $0.5D$. As a result, the respective costs for the target and the community are not discounted at the same rate.

A norm that is already (mostly) well-enforced, like a norm against armed robbery, will probably not be enforced to a much greater degree if we publicly shamed someone for armed robbery. It is thus unlikely that the marginal benefit of publicly shaming someone for robbery is sufficient, relative to the relative costs of a case of public shaming, for us to claim that the interests involved are truly symmetrical. A high SOP in such cases seems appropriate because of the relatively small weight of the interests in favour of public shaming. When we then take into account deontic factors which militate in favour of higher standards, it becomes even less likely that the balance of moral considerations favours the POE standard. It thus seems that we should not apply the relatively low POE standard for something like armed robbery. On the other hand, it is more plausible that symmetrical interests exist in cases where norms are not well-enforced, like in cases of sexual assault. Thus, while the POE standard might be appropriate for some cases of public shaming, it does not seem appropriate for the practice as a whole.

V. A Framework for Deliberation

When we consider the diversity of cases of public shaming, and the range of expected costs and benefits associated with them, something becomes clear: it is inappropriate to use a single principle (like BARD or POE) for the SOP for all cases of public shaming. Single principles simply cannot reflect the right balance of costs and benefits in every case of public shaming. For this very reason, it would also be incorrect to seek a single principle which lies somewhere in between the BARD and POE standards, like the Clear and Convincing Evidence (CACE) standard.³⁰³

³⁰³ The CACE standard is typically interpreted to mean that “the party with the burden of proof must convince the trier of fact that it is highly probable that the facts he alleges are correct”, and is seen as an intermediate standard between BARD and POE. Sherwin, Emily, “Clear and Convincing Evidence of Testamentary Intent: The Search for a Compromise Between Formality and Adjudicative Justice,” *Connecticut Law Review* 34, no. 2 (2002): 461.

The role of a public deliberative SOP is to give people instructions for how they should deliberate in a set of circumstances, to tell people what they should expect from one another, and give people the grounds to hold each other to account. To this end, an SOP need not be made up of a single principle, like “You may not act unless you have proof beyond a reasonable doubt” or “you should perform the act favoured by the preponderance of evidence”. An SOP can be constituted by multiple principles which together provide practical instructions for whether agents should carry out certain acts, given their circumstances and available evidence. I shall call such an SOP a *framework for deliberation*. For example, earlier we saw the example: “Use the BARD standard, except in cases of sexual assault. In such cases, use a lower standard”. These two sentences provide a simple framework for deliberation. A framework for deliberation can fulfil the role of a public deliberative SOP. As long as it is understandable, sufficiently accessible, and provides implementable instructions, it can tell people how to deliberate, what they can expect from one another, and how they can hold each other to account.

As we have seen in this chapter, an SOP for public shaming ought to be sensitive to a variety of stakes, including the interests affected by public shaming, deontic considerations, biases, and the correctives we need to counter the biases. The magnitude or weight of some of these various stakes in any case of public shaming can be affected by a number of general, observable factors. These factors include:

- a) The number of likely contributors to a case of public shaming.
 - The more contributors, the more significant the expected costs *and* benefits of a case of public shaming. A larger number of contributors is more likely to impose more significant harms, but it is also more likely to achieve the goals of public shaming.

b) The importance of the norm enforced by a case of public shaming (or the strength of the interests protected by the norm).

- Generally, the more important the norm, the greater the benefits of a case of public shaming.

c) How reliably the norm is enforced by existing institutions or other norms.

- If a norm is reliably enforced by existing mechanisms, the lower the expected marginal benefits of a case of public shaming.

d) The identity of the target of a case of public shaming.

- The more material and social resources available to a target, the more they may be able to insulate themselves from the effects of public shaming (eg. a rich celebrity may simply “lie low” for a few years and live on their existing wealth).
- High profile targets are sometimes also more likely to be targeted by more people, increasing the costs they may incur.
- Women and racial minorities sometimes experience public shaming that is more severe and vicious.

Additionally, there are common biases in reasoning and in assessing evidence which can affect a person’s decision to publicly shame, and these biases can vary from case to case. So ideally, the SOP for a given case of public shaming would take into account factors a-d, as well as the typical or common biases relevant to that case of public shaming.

Can we use this list of factors to generate a framework for deliberation? I think we can. One possible type of framework consists of two types of principles: a *default rule* – a principle that should be followed unless some exceptions attain – and an *adjustment rule* – a principle describing exceptions to the default rule, and what to do when those exceptions occur. These two types of

principles operate as heuristics which underlie many of our day-to-day decisions. For example, when I eat at a restaurant, the following principles may underlie my decisions:

- 1) Assume that food served to you in a restaurant is safe to eat, unless... (default rule)
- 2) Any of the following exceptions obtain, in which case seek further evidence until the safety of the food is beyond reasonable doubt (adjustment rule).
 - i) The food has an unusual colour.
 - ii) The food had an unusual smell.
 - iii) The restaurant has served unsafe food before.
 - iv) etc.

When it comes to the SOP for public shaming, a framework might include a default rule which specifies a single principle, like BARD. It can be supplemented by a number of adjustment rules. In this vein, I think a plausible SOP for public shaming might look similar to the following (I am open to the list of factors in mentioned in this SOP being non-exhaustive):

- a) In cases of public shaming, the public deliberative standard for public shaming should be high (like the BARD standard),³⁰⁴ unless:
 - The norm enforced by the case of public shaming is particularly important.

³⁰⁴ You might wonder why the default principle in this case is for a high standard, rather than a low one. Firstly, a high standard reflects that all cases of public shaming risk violating some deontic principles. In order to minimize the risk of violating a deontic principle, a high SOP should be used as a default. Secondly, I suspect that prospective contributors are predisposed to judge that a target is liable to be shamed because of confirmation bias. I further suspect that this predisposition is strong enough that a high evidentiary standard is justified in order to counteract it. These suspicions are based on my own, slightly speculative, guesses about the strength of the biases held by prospective contributors, which is why I have not allocated space within the main text to defend this particular position.

- The norm enforced by the case of public shaming is currently not well enforced by existing institutions.
- The target of public shaming enjoys a high social status and/or enjoys great wealth, and are likely to insulate themselves from some of the effects.
- There are common biases that are typically held against members of a group of people, which may reduce our credibility in the person/s making allegations against the target.³⁰⁵

b) When any one or more of the above conditions obtain, the public deliberative standard may be revised to something lower (Perhaps something similar to the CACE standard, or as low as the POE standard in some cases).

This SOP, while rough, does at least two things for us. First, it allows us to capture many of the factors that are relevant to the expected costs and benefits associated with cases of public shaming. Given that we want an SOP to bring about an acceptable distribution of costs and benefits, this framework is helpful. Secondly, this SOP is practical. Agents and communities can deliberate on how important a norm is, how well it is enforced, the status of the target, and the biases they might have. These considerations are things that are accessible to us. We can make claims about these considerations, justify those claims to each other, and roughly weigh how much of a role these considerations play. The considerations are not things that only philosophers and decision theorists and data scientists have access to. Those who participate in civic life, like internet

³⁰⁵ In this framework, I have left out the number of likely contributors to a case of public shaming. As I noted above, a larger group of contributors to a case can increase both the costs and the benefits of a case of public shaming. It is hard for normal agents to predict whether the increase in costs or benefits will be more significant in any given case. So it is not clear if this factor can help an agent decide whether to use a higher or lower SOP.

commentators, journalists, and ordinary people debating with their friends can and often do discuss issues like the importance of various norms or the existence of bias.

At this point, we can consider a couple of objections. You might worry that we do not agree on the issues related to the considerations I have mentioned. People do not agree on the importance of norms, whether they are well enforced, or how much bias we have. Thus, we will be unable to reach consensus on what the evidentiary threshold ought to be in any given case of public shaming. However, a lack of consensus is not necessarily a fatal flaw for an SOP. If it were, no SOP would be appropriate for almost any purpose. We do not even agree when it comes to relatively simple principles, like the BARD standard. People often disagree about what constitutes “reasonable doubt”. But this does not mean that the BARD standard is meaningless or useless. The BARD standard requires that agents deliberate about considerations in a specific way: whether any given consideration establishes reasonable doubt. It thus structures agents’ individual deliberations and the way agents engage in collective deliberations. It narrows the scope of any potential disagreement, thus helping a group of deliberators to focus on a set of issues. For example, in justifying their decision to acquit a defendant, a jury must explain why and how a prosecution failed to prove the defendant’s guilt beyond a reasonable doubt. Others may disagree with their decision if they do believe that the prosecution has proven the defendant’s guilt beyond a reasonable doubt. But importantly, by implementing the BARD standard, we’ve taken a broad and vague issue - whether the evidence favours convicting a defendant or not - and condensed it to a somewhat more manageable question - whether there is some consideration which makes it reasonable to doubt the guilt of the defendant.

Similarly, a framework like the one I have proposed may not give us clear and determinate guidelines in every case, but it structures our collective deliberative processes, and narrows the scope of public debates. As a result, it gives individual agents some guidance to think about issues

in certain ways, creates certain expectations we might have of each other, and justifies demands with which we can hold each other to account. For example, imagine I notice on a forum someone contributing to public shaming, despite having relatively weak evidence that public shaming is permissible in this instance. I may point out to them that given the circumstances, she is required to meet a higher evidentiary threshold to justify contributing. A framework like the one I have proposed allows me to mount a sensible challenge to her behaviour, because without some framework, I may have no reasons to back up my challenge - I have “nothing to go on”. And in return, she might argue that the norm in this case is not well enforced, which in turn makes it all the more important to step in and call out behaviour that violates the norm. Even though we do not agree, the framework that we share provides some common ground which establishes the basic expectations, and the considerations that are relevant in affecting those expectations. This common ground provides the basis for any demands I might make of her that she explain her behaviour, which allows me to sensibly challenge her behaviour in the first place. And the common ground provides a structure for any explanation she might give to justify her behaviour.

Another objection is an argument from simplicity. Someone who supports the BARD standard or the POE standard for the practice of public shaming may point out that a single principle is simpler. As a result, it provides a clear set of instructions, which is easier to use both in deliberations and in our practices of holding each other to account, like assessing each other’s behaviour, explaining, and justifying our actions. I agree that simplicity is valuable when we establish a set of standards which constitute standards for public deliberation. However, the virtue of simplicity must be weighed against other desiderata, among them the ability of an SOP to adequately reflect the relevant moral stakes in a situation. While a single principle for all cases of public shaming may be simple, we should reject it if it will result in overly large costs in a significant set of cases. In fact, in many cases in other practices, we often prefer that agents consider a

multitude of factors, rather than follow a single principle or rule, when making important decisions. For example, medical professionals should consider a number of factors in making important decisions – like a patient’s autonomy, wellbeing, the needs of their dependents, competing claims and needs, and so on – and think about how these factors weigh against each other. Sure, we might not want doctors to spend hours each day pondering ethical dilemmas. But we also do not and should not want them to make important decisions based on overly reductive principles or heuristics which do not sufficiently take into account the moral stakes of a situation.

The SOP I suggested earlier may be more complicated than something like the BARD standard. But I do not think it is unacceptably complicated. We can and often do make judgments regarding issues related to elements within the SOP, like the importance of norms, how well norms are enforced, the status and wealth of others, and how biased people are. We use these judgments as considerations in important decisions that many of us do make, like in our votes. These judgments can be made by ordinary people, not just individuals with very specialized skills or instruments of measurements. These judgments may not always be correct, but they may still be roughly correct, particularly in cases of comparative judgments (eg. this norm is more important than that norm). Furthermore, we can communicate these judgments to each other, explain them, challenge and defend those judgments, and thus deliberate about them together. So even if some individuals within a community cannot by themselves understand the SOP and employ it to attain good outcomes, a community of people can still use the SOP as a framework for collective deliberation.

VI. Conclusion

The use of public shaming, like many other tools, comes with potential benefits and risks. In this thesis we have looked at the risks posed by public shaming, but also the potential good that comes from it. While many philosophers consider the risks and the harm caused by public shaming

to be too severe for it to be justified, and while others have decried the use of shame or humiliation more generally, I have taken a slightly sunnier outlook on public shaming. While the risks of public shaming are not to be taken lightly, the potential benefits and interests that can be fulfilled by cases of public shaming can justify some cases of public shaming.

In this particular chapter, we have weighed the risks and the benefits of public shaming under conditions of uncertainty. I argued that our standard of proof for justifying cases of public shaming should include a number of factors, and I structured these factors with a framework for deliberation – a set of principles which gives agents practical instructions for how they should act under a set of circumstances.

Despite the lengthy discussion in this thesis, I do not think I have exhausted the issues. Within the specific issue of SOPs for public shaming, there remain potential further discussions which precisify and flesh out the framework I have sketched out here. Empirical or quantitative research on the effects of public shaming may uncover more details about the costs and benefits of public shaming, which can be used to improve the framework. There are also further issues which pertain to the topic of public shaming as a whole. These include the relationship between public shaming and public justification (whether contributors must justify their actions on the basis of public reasons), as well as social forgiveness (whether communities have duties to targets of rightful public shaming to reintegrate them). These questions and issues present further avenues and ripe ground for more philosophical inquiry. Such a prospect, I think, is an exciting note with which to conclude this project.

Notes for Chapter VI

We can assess the expected value of an SOP when applied to a practice with a rudimentary model.

Let us start with a few idealizing and simplifying assumptions:

- A) A case of public shaming is rightful if and only if the target is liable, and wrongful if the target is not liable.
- B) A case of rightful shaming can be expected to bring about more (morally-weighted) benefits than (morally-weighted) costs, and will on balance bring about net benefit b . A case of wrongful public shaming can be expected to bring about some net cost $-c$ (where c is a positive value). It is of course possible that in some cases, public shaming will result in more harm than good regardless of whether it is wrongful or rightful. In other cases, public shaming will result in more good than harm regardless. In these two types of cases, whether or not we should shame may not depend on the uncertainties about whether the public shaming is wrongful or rightful, rendering the discussion of standards of proof pointless. So we can set these cases aside.
- C) There will be no costs or benefits when public shaming does not occur, because we can model the cost of failing to shame as a benefit of shaming, and vice versa.
- D) Communities comply with their moral judgments about whether an SOP has been met.

Given these assumptions, we can model the expected value associated with a single case of public shaming, given a particular SOP, and then model a practice of public shaming which uses that SOP. For a case of potential public shaming, the possible outcomes and their associated moral values, can be represented as such:

	Is public shaming rightful?	
	Yes	No (it is wrongful)
Community judges that SOP is met (shaming occurs)	Outcome: true positive Value: b	Outcome: false positive Value: -c
Community judges that SOP is not met (shaming does not occur)	Outcome: false negative Value: 0	Outcome: true negative Value: 0

The expected value associated with a potential case of public shaming given an SOP, can be given by:

$$\text{Expected value (EV)} = (\text{probability of true positive}) \times (b) + (\text{probability of false positive}) \times (-c) + 0 + 0$$

If we look at a case of public shaming that has occurred, rather than a potential case of public shaming, we can ignore the bottom row of the above matrix. The EV of a case of public shaming that occurs will still be given by the above equation. However, the case will either be a case of wrongful or rightful public shaming. Thus,

$$(\text{probability of true positive}) + (\text{probability of false positive}) = 1$$

A practice of public shaming is made up of some set of cases of public shaming. The total expected value of applying an SOP to a practice of public shaming is simply the sum of the expected values of every case under the practice. Let the total number of cases of public shaming within the practice of public shaming be n . Thus:

$$\text{Total Expected Value of Practice (EVP)} = EV_1 + EV_2 + \dots EV_n$$

We can now compare the EVPs for different possible practices of public shaming, where we hold fixed all factors except for the SOPs that are used in the different practices. In a very general sense, when we raise the SOP for a practice of public shaming, we might expect two effects. First, we would expect the total number of cases within the practice (n) to decrease. When it is harder to justify public shaming, there will likely be less public shaming. Secondly, for any given case of public shaming that has occurred, the chances of it being a true positive rather than a false positive will increase. Thus, the $[(\text{probability of true positive}) - (\text{probability of false positive})]$ will increase. The reverse is the case when we decrease the SOP for public shaming. n will increase, and $[(\text{probability of true positive}) - (\text{probability of false positive})]$ will decrease.

But there are complicating factors. One is the background rate of false accusations compared to the background rate of true accusations. For instance, imagine that within some society, false accusations of wrongdoing are extremely rare. More than 99% of all people accused of wrongdoing are in fact liable to be publicly shamed. In this society, both a high SOP and an arbitrarily low one will result in roughly the same outcomes: for any given case of public shaming that occurs, it is likely to be a case of rightful public shaming. Therefore, if the background rate of false accusations is sufficiently low, $[(\text{probability of true positive}) - (\text{probability of false positive})]$ will not be very sensitive to the SOP. There is an even further complication when we take into account the fact that an SOP can in turn affect the background rate of false accusations. A low SOP can in some cases incentivize people to make false accusations against people they dislike, because there is a higher probability that they will be believed. Thus, the relationship between an SOP and the $[(\text{probability of true positive}) - (\text{probability of false positive})]$ for any case is not straightforward.

If we set aside these complications, we can use our model to generate a number of claims, which can then be used to generate further conclusions. These conclusions are limited because of the complications we have just set aside. First, when the SOP for public shaming is higher, the EV for any case of public shaming that does occur will be higher. This is because b and c are not affected by the SOP. By increasing the SOP for public shaming, we increase the probability that any case of public shaming will be a true positive rather than a false positive. In turn, by increasing the probability of a true positive, we increase the EV. Second, if we suppose that the EV for an average case of public shaming is positive, we want to maximize the number of cases of public shaming (n). Third, if the EV for an average case of public shaming is negative, we want to minimize n .

With the above model, we can consider a number of scenarios and make conclusions about the SOPs that are best under each one.

Costs are greater than Benefits. Let us begin with a scenario where the expected costs of a case of public shaming are on average much greater than the expected benefits. That is, on average c , is much larger than b . In such a scenario, unless the likelihood of a true positive significantly exceeds the likelihood of a false positive, the EV of an average case of public shaming will be negative. Suppose that c is five times as much as b , we would get:

$$EV = (\text{probability of true positive}) \times (b) + (\text{probability of false positive}) \times (-5b)$$

In this case, EV would be more than 0 only if (probability of true positive) is more than five times as much as (probability of false positive). (probability of true positive) would thus have to be more than 5/6. Because the total EV of a practice is just the sum of every case, the EVP will only be positive if the EV of the average case is also positive. Thus, we would want an SOP that is likely to ensure that the rate of true positives vastly exceeds the rate of false positives. Under such a

scenario, we should prefer a high SOP, like BARD (setting aside the complications that arise from background rates of false accusations). Lower SOPs can be expected to result in negative value over time, and so the benefits will simply not justify the costs of using them.

Benefits are greater than Costs. Now let us consider a reverse scenario where the expected benefits of a case of public shaming are on average much greater than the expected costs. Under such a scenario, any given case of public shaming is likely to come with positive EV, unless the probability of a false positive in that case is particularly high (again, the background rate of false accusations plays a big role in determining how an SOP can affect the probabilities of true positives and false positives). Thus, several considerations become relevant. On one hand, because higher SOPs are associated with higher EVs in each case, there is a still consideration which militates in favour of using a relatively high SOP. On the other hand, a lower SOP can increase the total number of cases of public shaming n . If the average case of public shaming, given a particular SOP, comes with a positive EV, then we can increase EVP by increasing n . Thus, maximizing the EVP of a practice of public shaming in this scenario involves finding the right balance of these opposing considerations. Until we put in some actual figures into the model, we cannot know beforehand what the right SOP should be.

Benefits are similar to Costs. Finally, we consider the scenario where the expected benefits and costs of public shaming in an average case are roughly similar to each other. Neither b nor c are much greater than the other. For illustration, let us take a set of cases where on average $b=c$. In such a set of cases, EV will be positive as long as the number of true positives exceeds the number of false positives. Should we favour a higher or lower SOP with regard to this set of cases? This particular scenario is perhaps more complicated than the other two, because things are particularly sensitive to the ratio of true positives to false positives, or $[(\text{probability of true positive}) - (\text{probability of false positive})]$. If $b=c$ in the average case, then $EV > 0$ if and only if $[(\text{probability$

of true positive) – (probability of false positive)] > 0. As we saw earlier, the respective probabilities of true and false positives can be affected by the SOP *and* the background rates of false accusations. Just to ensure that $EV > 0$ in the average case of public shaming, we need to determine what the minimum SOP is, given the background rate of false accusations, which would result in more true positives than false positives.

After we've determined this minimum SOP, our task is not yet finished. Like the previous scenario, two more considerations come into play. On one hand, a higher SOP will result in a higher likelihood of a case of public shaming being a true positive rather than a false negative. This means that the EV of any given case of public shaming will be higher with a higher SOP, *given that public shaming has occurred*. On the other hand, a higher SOP will result in a lower number of cases of public shaming in total. Because the EVP of the set of cases is a sum of all the cases, the EVP may be lower than what it could otherwise be, if the EV of each case is positive.

The result is striking. As we saw in section IV, proponents of the POE standard argue that when the stakes are roughly symmetrical, we ought to use the POE standard. But if the goal is to maximize the expected value of applying an SOP to a practice, it is not obvious that this is correct.

Bibliography

- “Academic Boycott”. BDS. Retrieved 13 Dec 2023. <https://bdsmovement.net/academic-boycott#guidelines>.
- Adkins, Karen. “When Shaming Is Shameful: Double Standards in Online Shame Backlashes,” *Hypatia* 34, no. 1 (2019): 76-97.
- Aitchison, Guy, and Saladin Meckland-Garcia. “Against Online Public Shaming: Ethical Problems with Mass Social Media”. *Social Theory and Practice* 47, no.1 (2021): 1-31.
- Alexander, Larry, Ferzan, Kimberly, and Stephen Morse. *Crime and Culpability: A Theory of Criminal Punishment*. Cambridge: Cambridge University Press, 2009.
- Alexander, Lawrence. “The Means Principle”. *Legal, Moral, and Metaphysical Truths: The Philosophy of Michael Moore*. Edited by K. K. Ferzan and S.J. Morse. Oxford: Oxford Univ. Press, 2014.
- San Diego Legal Studies Paper No. 14-138, Available at SSRN: <https://ssrn.com/abstract=2378608>.
- Alpher, Roger. “#MeToo Means the End of the Presumption of Innocence”. *Haaretz*, 8 Nov 2021. <https://www.haaretz.com/opinion/.premium-metoo-means-the-end-of-the-presumption-of-innocence-1.10362331>.
- Bailey, Daniel S., and Lisa L. Sample. “An Examination of a Cycle Of Coping with Strain Among Registered Citizens’ Families”. *Criminal Justice Studies* 30, no. 2 (2017): 158-180.
- Barron, Anna C., Lydia Woodyatt, Emma F. Thomas, Jia En Katherine Loh, and Katherine Dunning. “Doing good or feeling good? Justice concerns predict online shaming via deservingness and schadenfreude”. *Computers in Human Behavior Reports* 11 (2023): 100317.

- Barry, Christian. “Applying the Contribution Principle”. *Metaphilosophy* 36 no. 2 (2005): 210-227.
- Barry, Christian, and Gerhard Øverland. “Individual responsibility for carbon emissions: Is there anything wrong with overdetermining harm?” In *Climate Change and Justice*, edited by Jeremy Moss. Cambridge: Cambridge University Press, 2015.
- Bell, Macalester. *Hard Feelings: The Moral Psychology of Contempt*. Oxford: Oxford University Press, 2013.
- Bennett, Christopher. *The Apology Ritual*. Cambridge: Cambridge University Press, 2009.
- Billingham, Paul, and Tom Parr. “Enforcing social norms: The morality of public shaming”. *European Journal of Philosophy* 28, no. 4 (2020): 997-1016.
- Blackburn, Jeremy, Robert W. Gehl, and Ugochukwu Etudo. “Does Deplatforming Actually Quell Hate Speech Online?” *Undark*, 19 Jan 2021, <https://undark.org/2021/01/19/does-deplatforming-actually-quell-hate-speech-online/>.
- Blackford, Russell. “The Shame of Public Shaming”. *The Conversation*. 4 May 2016. <https://theconversation.com/the-shame-of-public-shaming-57584>
- Blackstone, William. *Commentaries on the Laws of England: In Four Books*. Philadelphia: J. B. Lippincott, 1753.
- Blome-Tillmann, Michael. “More Likely Than Not: Knowledge First and the Role of Bare Statistical Evidence in Courts of Law”. In J. Adam Carter, Emma C. Gordon, and Benjamin Jarvis (eds). *Knowledge First: Approaches in Epistemology and Mind*. Oxford: Oxford University Press, 2017.
- Borysenko, Karlyn. “The Dark Side Of #MeToo: What Happens When Men Are Falsely Accused”. *Forbes*, 12 Feb 2020.

<https://www.forbes.com/sites/karlynborysenko/2020/02/12/the-dark-side-of-metoo-what-happens-when-men-are-falsely-accused/?sh=465d6796864d>.

Bourguignon, David, Eleonore Seron, Vincent Yzerbyt, and Ginette Herman. "Perceived Group and Personal Discrimination: Differential Effects on Personal Self-Esteem," *European Journal of Social Psychology* 36 (2006): 773–789.

Boxill, Bernard. "Self-Respect and Protest," *Philosophy & Public Affairs* 6, no. 1 (1976), 58-69.

Braithewaite, John. *Crime, Shame, and Reintegration*. Cambridge: Cambridge University Press, 1989.

Brake, Deborah L. "Fighting the Rape Culture Wars through the Preponderance of the Evidence Standard". *Montana Law Review* 78, no. 1 (2017): 109-153.

Bratman, Michael E. *Shared Agency: A Planned Theory of Acting Together*. Oxford: Oxford University Press, 2014.

Brennan, Geoffrey, and Philip Pettit. *The Economy of Esteem: An Essay on Civil and Political Society*. Oxford: Oxford University Press, 2004.

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

Brooks, Thom. "Shame On You, Shame On Me? Nussbaum on Shame Punishment". *Journal of Applied Philosophy* Vol. 25, No. 4 (2008), 322-334.

Brownlee, Kimberley. "A Human Right against Social Deprivation". *The Philosophical Quarterly* 63, no. 251 (2013): 199-222.

Brownlee, Kimberly. "Ethical Dilemmas of Sociability". *Utilitas* 28, no. 1 (2016): 54-72.

- Budolfson, Mark. "The Inefficacy Objection to Deontology: What it is, Why it is Important, and How to Respond to it". *Budolfson.com* (retrieved 17 Nov 2023), <https://budolfson.github.io/files/BudolfsonDeontology.pdf>
- Cacioppo, John T., and Stephanie Cacioppo. "Social Relationships and Health: The Toxic Effects of Perceived Social Isolation". *Soc Personal Psychol Compass* 8, no. 2 (2014): 58-72.
- Cacioppo, J.T. et al. "Loneliness as a Specific Risk Factor for Depressive Symptoms: Cross-sectional and Longitudinal Analyses". *Psychology and Aging* 21 (2006): 140-151.
- Calhoun, Cheshire. "An apology for moral shame". *The Journal of Political Philosophy* 12, no. 2 (2004): 127-146.
- "Carl Sargeant: Welsh politician who resigned amid sexual misconduct claims is found dead". *ABC News*, 8 Nov 2017, <https://www.abc.net.au/news/2017-11-08/welsh-politician-who-resigned-amid-misconduct-claims-dies/9128596>.
- Carlsen, Audrey, Maya Salam, Claire Cain Miller, Denise Lu, Ash Ngu, Jugal K. Patel, and Zach Wichter. "#MeToo Brought Down 201 Powerful Men. Nearly Half of Their Replacements Are Women". *New York Times*, 29 Oct 2019. <https://www.nytimes.com/interactive/2018/10/23/us/metoo-replacements.html?nl=top-stories&nlid=72995439ries&ref=cta>.
- Chmielewski, Amy. "Defending the Preponderance of the Evidence Standard in College Adjudications of Sexual Assault". *Brigham Young University Education and Law Journal* 3, no. 1 (2013): 143-174.
- Coates, D. Justin, and Neal A. Tognazzini (eds). "The Contours of Blame", in D. Justin Coates, and Neal A. Tognazzini (eds). *Blame: Its Nature and Norms*. New York, 2012; online edn,

<https://doi.org/10.1093/acprof:oso/9780199860821.003.0001>.

Collins, Stephanie, “A Human Right to Relationships”. In *Being Social: the Philosophy of Human Rights*.

Edited by Kimberly Brownlee, David Jenkins, and Adam Neal. Oxford: Oxford University Press, 2022.

Collins, Stephanie. “Duties to Make Friends”. *Ethical Theory and Moral Practice* 16 (2013): 907–921.

Cooper, Christian. *Better Living Through Birding*. New York: Random House, 2023.

Couto, Alexandra. “A Right to Opportunities for Meaningful Relationships”. In *Being Social: The Philosophy of Social Human Rights*. Edited by Kimberley Brownlee, David Jenkins, and Adam Neal. Oxford: Oxford University Press, 2022.

Crewe, Bianca, and Jonathan Jenkins Ichikawa. “Rape Culture and Epistemology”. In Jennifer Lackey (ed.) *Applied Epistemology*. New York: Oxford University Press, 2021.

Cullity, Garrett. “Climate Harms”. *The Monist* 102, no. 1 (2019): 22–41.

Darwall, Stephen. *Morality, Authority, and Law*. Oxford: Oxford University Press, 2013.

Darwall, Stephen. “Two Kinds of Respect,” *Ethics* 88, no. 1 (1977): 36-49.

Delon, Nicolas. “Scarlet Letters: Meat, Normality and the Power of Shaming”. *BooksandIdeas.net*. Sep 2019. <https://booksandideas.net/Scarlet-Letters.html>.

Demougin, Dominique, and Claude Fluet. “Preponderance of Evidence”. *European Economic Review* 50, no. 4 (2006): 963-976.

Diamond, Henry A. "Reasonable Doubt: To Define, Or Not To Define". *Columbia Law Review* 90, no. 6 (1990): 1716-1736.

Donegan, Moira. "Harvey Weinstein went from untouchable to incarcerated. Thank #MeToo".

The Guardian. 25 Feb 2020.

<https://www.theguardian.com/commentisfree/2020/feb/24/harvey-weinstein-untouchable-guilty-thank-me-too>.

Duff, Antony, Lindsay Farmer, Sandra Marshall, and Victor Tadros. *The Trial On Trial Vol 3: Towards A Normative Theory Of The Criminal Trial*. Oxford: Hart Publishing, 2007.

Dumsday, Travis. "On Cheering Charles Bronson: The Ethics of Vigilantism". *The Southern Journal of Philosophy* 47 (2009): 49-67.

Dwyer, Christopher. "12 Common Biases That Affect How We Make Everyday Decisions". *Psychology Today*, 17 Sep 2018. <https://www.psychologytoday.com/us/blog/thoughts-on-thinking/201809/12-common-biases-that-affect-how-we-make-everyday-decisions>.

Every, Daniel, and Richard Perry. "The Relationship between Perceived Religious Discrimination and Self-Esteem for Muslim Australians," *Australian Journal of Psychology* 66, no. 4 (2014): 241-248.

Fantl, Jeremy and Matthew McGrath. "Evidence, Pragmatics, and Justification". *The Philosophical Review* 111, no. 1 (2002): 67-94.

Favarel-Garrigues, Gilles, Samuel Tanner & Daniel Trottier. "Introducing digital vigilantism," *Global Crime* 21, no. 3-4 (2020): 189-195.

Feinberg, Joel. *The Moral Limits of Criminal Law Volume 1: Harm to Others*. New York: Oxford University Press, 1984.

- Ferzan, Kimberly Kessler. "#BelieveWomen and the Presumption of Innocence: Clarifying the Questions for Law and Life". *All Faculty Scholarship*. 2720 (2021). https://scholarship.law.upenn.edu/faculty_scholarship/2720.
- Ferzan, Kimberly Kessler. "Culpable Aggression: The Basis for Moral Liability to Defensive Killing". *Ohio State Journal of Criminal Law*, 9, no. 2 (2012): 669–697.
- Franklin, Christopher Evan. *A Minimal Libertarianism: Free Will and the Promise of Reduction*. Oxford: Oxford University Press, 2018.
- Fricker, Miranda. *Epistemic Injustice: Power & the Ethics of Knowing*. Oxford: Oxford University Press, 2007.
- Fritz, James. "Online Shaming and the Ethics of Public Disapproval". *Journal of Applied Philosophy* 38, no.4 (2021): 686-701.
- Frowe, Helen and Jonathan Parry. "Self-Defense". *The Stanford Encyclopedia of Philosophy* (Summer 2022 Edition), ed. Edward N. Zalta, <https://plato.stanford.edu/archives/sum2022/entries/self-defense/>.
- Frye, Harrison. "The Problem of Public Shaming". *The Journal of Political Philosophy* 30, no. 2 (2022): 188-208.
- Furnham, Adrian, and Hua Chu Boo. "A literature review of the anchoring effect". *The Journal of Socio-Economics* 40, no. 1 (2011): 35-42.
- Gardiner, Georgi. "Legal Burdens of Proof and Statistical Evidence". In *The Routledge Handbook of Applied Epistemology*. Edited by David Coady and James Chase. London: Routledge, 2018.

- Gilmore, Jane. "Presumption of innocence and the #MeToo backlash". *The Sydney Morning Herald*, 23 Oct 2018. <https://www.smh.com.au/lifestyle/life-and-relationships/presumption-of-innocence-and-the-metoo-backlash-20181019-p50aqw.html>.
- Goff, P. A., Jackson, M. C., Di Leone, B. A., Culotta, C. M., & DiTomasso, N. A. "The Essence of Innocence: Consequences of Dehumanizing Black Children". *Journal of Personality and Social Psychology* 106, no. 4 (2014): 526 –545.
- Goodin, Robert E. "Acting in Combination". *Philosophy & Public Affairs* 45, no. 2 (2017): 158- 194.
- Grady, Constance. "Some say the Me Too movement has gone too far. The Harvey Weinstein verdict proves that's false". *Vox*. 24 Feb 2020. <https://www.vox.com/culture/2020/2/24/21150966/harvey-weinstein-rape-conviction-sexual-predatory-assault-me-too-too-far>.
- "Harvey Weinstein timeline: How the scandal has unfolded". *BBC News*. 24 Feb 2023. <https://www.bbc.com/news/entertainment-arts-41594672>.
- Hawkey, L.C. et al. "Loneliness Predicts Increased Blood Pressure: 5-year Cross lagged Analyses in Middle-aged and Older Adults". *Psychology and Aging* 25 (2010): 132-141.
- Hawthorne, John, and Jason Stanley. "Knowledge and Action". *The Journal of Philosophy* 105, no. 10 (2008): 571-590.
- Hieronymi, Pamela. "Articulating an Uncompromising Forgiveness," *Philosophy and Phenomenological Research*, 62, no. 3 (2001): 529-555.
- Hieronymi, Pamela. "The Force and Fairness of Blame". *Philosophical Perspectives* 18, no. 1 (2004): 115–148.

Hohl, Katrin. “New scorecards show under 1% of reported rapes lead to conviction – criminologist explains why England’s justice system continues to fail”. *The Conversation*, 2 April 2022. <https://theconversation.com/new-scorecards-show-under-1-of-reported-rapes-lead-to-conviction-criminologist-explains-why-englands-justice-system-continues-to-fail-180345>.

Hosein, Adam Omar. “Doing, Allowing, and the State”. *Law and Philosophy* Vol. 33, No. 2 (March 2014), pp. 235-264.

Jacquet, Jennifer. *Is Shame Necessary? New Uses for an Old Tool*. New York: Pantheon Books, 2015.

Jaffe AE, Cero I, and DiLillo D. “The #MeToo Movement and Perceptions of Sexual Assault: College Students' Recognition of Sexual Assault Experiences Over Time”. *Psychol Violence* 11, no. 2 (2021): 209-218.

Jaggar, Alison. “Love and knowledge: Emotion in feminist epistemology”. *Inquiry* 32, no. 2 (1989): 151-176.

Jashinsky, Emily. “Amy Cooper Doesn’t Deserve Sympathy, But Social Media Shaming Is An Unhealthy Norm”. *The Federalist*. 27 May 2020. <https://thefederalist.com/2020/05/27/amy-cooper-doesnt-deserve-sympathy-but-social-media-shaming-is-an-unhealthy-norm/>.

Johnson, Rucker, and Stephen Raphael. “How Much Crime Reduction Does the Marginal Prisoner Buy?” *Journal of Law and Economics*, Vol. 55, No. 2 (May 2012), pp. 275-310.

Johnson, Stephanie K, Ksenia Keplinger, Jessica F. Kirk, and Liza Barnes. “Has Sexual Harassment at Work Decreased Since #MeToo?” *Harvard Business Review*, 18 Jul 2019. <https://hbr.org/2019/07/has-sexual-harassment-at-work-decreased-since-metoo>.

- Johnson, W. Brad, and David G. Smith. “Men Shouldn’t Refuse to Be Alone with Female Colleagues”. *Harvard Business Review*, 5 May 2017. <https://hbr.org/2017/05/men-shouldnt-refuse-to-be-alone-with-female-colleagues>.
- Kahan, Dan. “Shame, Stigma, and Crime: Evaluating the Efficacy of Shaming Sanctions in Criminal Law”. *Harvard Law Review* Vol. 116, No. 7 (2003), pp. 2186-2207.
- Kagan, Shelly. “Do I Make a Difference?” *Philosophy & Public Affairs* 39, no. 2 (2011): 105-141.
- Kasra, Mona. “Vigilantism, public shaming, and social media hegemony: The role of digital-networked images in humiliation and sociopolitical control,” *The Communication Review* 20, no. 3 (2017): 172-188.
- Keplinger, Ksenia, Stephanie K. Johnson, Jessica F. Kirk, Lisa Y. Barnes. “Women at work: Changes in sexual harassment between September 2016 and September 2018,” *PloS One* 14, no. 7 (2018): e0218313. <https://doi.org/10.1371/journal.pone.0218313>.
- Kim, Eunha, and Hansol Park. “Perceived Gender Discrimination, Belief in a Just World, Self-Esteem, and Depression in Korean Working Women: A Moderated Mediation Model,” *Women's Studies International Forum* 69 (2018): 143–150.
- Klockseim, Justin. “Harm, Failing to Benefit, and the Counterfactual Comparative Account”. *Utilitas* 34, no. 4 (2022): 428 – 444.
- Kutz, Christopher. *Complicity: Ethics and Law for a Collective Age*. Cambridge: Cambridge University Press, 2000.
- Langone, Alix. “#MeToo and Time’s Up Founders Explain the Difference Between the 2 Movements — And How They’re Alike”. *Time*, 22 Mar 2018.

<https://time.com/5189945/whats-the-difference-between-the-metoo-and-times-up-movements/>.

Lazar, Seth. "Necessity in Self-Defense and War". *Philosophy and Public Affairs* Vol 40, no. 1 (2012): 3-44.

Lee, Kevan. "Infographic: The Optimal Length for Every Social Media Update and More". *Buffer.com*. 21 Oct 2014. <https://buffer.com/library/optimal-length-social-media/>.

Levy, Ro'ee, and Martin Mattson. "The Effects of Social Movements: Evidence from #MeToo". *SSRN.com*, 27 Apr 2023. <http://dx.doi.org/10.2139/ssrn.3496903>.

Lewis, David. *Convention: A Philosophical Study*. Cambridge: Harvard University Press, 1969.

Liao, Matthew S. *The Right to be Loved*. Oxford: Oxford University Press, 2015.

Littlejohn, Clayton. "Truth, knowledge, and the standard of proof in criminal law". *Synthese* 197 (2020): 5253–5286.

Loughlin, M, Dolezal, L, Hutchinson, P, Subramani, S, Milani, R, Lafarge, C. "Philosophy and the clinic: stigma, respect and shame". *J Eval Clin Pract.* 28 (2022): 705- 710.

Loveluck, Benjamin. "The many shades of digital vigilantism. A typology of online self-justice," *Global Crime* 21, no. 3-4 (2019): 213-241.

Marmor, Andrei. "What is the Right to Privacy?" *Philosophy and Public Affairs* 43, no. 1 (2015): 3-26.

McIntyre, Alison. "Doctrine of Double Effect". The Stanford Encyclopedia of Philosophy (Fall 2023 Edition), Edward N. Zalta & Uri Nodelman (eds.), <https://plato.stanford.edu/archives/fall2023/entries/double-effect/>.

- McGeer, Victoria. "Civilizing Blame", in D. Justin Coates, and Neal A. Tognazzini (eds). *Blame: Its Nature and Norms*. New York, 2012; online edn, Oxford Academic, 23 May 2013.
<https://doi.org/10.1093/acprof:oso/9780199860821.003.0009>.
- McMahon, Jeff. "The Limits of Self-Defense". In *The Ethics of Self-Defence*. Edited by Christian Coons and Michael Weber. New York: Oxford University Press, 2016.
- Max, D.T. "The Public Shaming Pandemic". *The New Yorker*. 21 Sep 2020,
<https://www.newyorker.com/magazine/2020/09/28/the-public-shaming-pandemic>
- Mohammad, Farah. "The Danger of Public Shaming in the Internet Age". 25 Jan 2018.
<https://daily.jstor.org/the-danger-of-public-shaming-in-the-internet-age/>.
- Moniuszko, Sara M, and Cara Kelly. "Harvey Weinstein scandal: A complete list of the 87 accusers". *USAToday*. 27 Oct 2017.
<https://www.usatoday.com/story/life/people/2017/10/27/weinstein-scandal-complete-list-accusers/804663001/>.
- Morris, Christopher. "Punishment and Loss of Moral Standing". *Canadian Journal of Philosophy* 21, no. 1 (1991): 53-79.
- Moss, Sarah. "Knowledge and Legal Proof". In Tamar Szabó Gendler, John Hawthorne, and Julianne Chung (eds). *Oxford Studies in Epistemology Volume 7: Oxford Studies in Epistemology*. Oxford: Oxford University Press, 2022.
- Murphy, Jeffrie. "Forgiveness, Self-Respect, and the Value of Resentment". In *Handbook of Forgiveness*, eds. Everett L. Worthington Jr. and Nathaniel G Wade. Routledge Handbooks, 2005.
- Nefsky, Julia. "Collective Harm and the Inefficacy Problem". *Philosophy Compass* 14 (2019): e12587.

- Nussbaum, Martha. *Hiding from Humanity: Disgust, Shame, and the Law*. Princeton: Princeton University Press, 2004.
- Office for Civil Rights, Dep't of Education. *Dear Colleague Letter*, 4 Apr 2011. Retrieved 27 Dec 2023. <http://www2.ed.gov/about/offices/list/ocr/letters/colleague-201104.pdf>.
- O'Neill, Brendan. "What happened to 'innocent until proven guilty'?" *Chicago Tribune*, 28 Nov 2017. <https://www.chicagotribune.com/opinion/commentary/ct-perspec-innocense-presumption-guilty-metoo-sexual-misconduct-1128-20171127-story.html>.
- Otsuka, Michael. "The Moral-Responsibility Account of Liability to Defensive Killing". In *The Ethics of Self-Defense*. Edited by Coons and Weber. New York: online edn, Oxford Academic, 2016.
- Parfit, Derek, *Reasons and Persons*. Oxford: Oxford University Press, 1984.
- Peretz, Evgenia. "How Jill Messick's Suicide Reflects the Tragedies of the #MeToo Era". *Vanity Fair*, 6 Sep 2018, <https://www.vanityfair.com/hollywood/2018/09/how-jill-messick-suicide-reflects-the-tragedies-of-the-metoo-era>.
- Pettit, Philip. "My Three Selves". *Philosophy* 95, no. 3 (2020): 363-389.
- Pridmore S, McArthur M. "Suicide and Reputation Damage". *Australasian Psychiatry* 16, no. 5 (2008):312-316.
- Purves, Duncan. "Harming as making worse off". *Philosophy Studies* 176 (2019): 2629–2656.
- Quinn, Warren. "Actions, Intentions, and Consequences: The Doctrine of Double Effect". *Philosophy & Public Affairs* Vol. 18, No. 4 (1989): 334-351.
- Quinn, Warren. "The Right to Threaten and the Right to Punish". *Philosophy & Public Affairs* 14, no. 4 (1985), 327-373.

- Radzik, Linda. "On Minding Your Own Business: Differentiating Accountability Relations within the Moral Community," *Social Theory and Practice* 37, no. 4 (2011): 574-598.
- Radzik, Linda. *The Ethics of Social Punishment*. Cambridge: Cambridge University Press, 2020.
- Rassin, Eric, Anita Eerland, and Ilse Kuippers. "Let's Find the Evidence: An Analogue Study of Confirmation Bias in Criminal Investigations". *Journal of Investigative Psychology and Offender Profiling* 1.7 (2010): 231–246.
- Rawls, John. *A Theory of Justice*. Cambridge: Harvard University Press, 1971.
- Renzo, Massimo. "Rights Forfeiture and Liability to Harm". *The Journal of Political Philosophy* 25, no. 3 (2017): 324-342.
- Ronson, Jon. *So You've Been Publicly Shamed*. London: Picador, 2015.
- Rumbold, Benedict, and James Wilson. "Privacy Rights and Public Information". *The Journal of Political Philosophy* 27, no. 1 (2018): 3-25.
- Scanlon, T.M. "Giving Desert its Due," *Philosophical Explorations* 16, no. 2 (2013): 101-116.
- Scanlon, T.M. *What We Owe to Each Other*. Harvard: Harvard University Press, 1998.
- Schweikard, David P. and Hans Bernhard Schmid. "Collective Intentionality". *The Stanford Encyclopedia of Philosophy* (Fall 2021 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/fall2021/entries/collective-intentionality/>.
- Searle, John. *Making the Social World: The Structure of Human Civilization*. Oxford: Oxford University Press, 2010.

“Seoul's mayor found dead in presumed suicide after #MeToo allegation”. *France 24*, 9 Jul 2020, <https://www.france24.com/en/20200709-body-of-seoul-s-mayor-found-after-massive-search>.

Sher, George. “Out of Control”. *Ethics* 116, no. 2 (2006): 285–301.

Sherwin, Emily. “Clear and Convincing Evidence of Testamentary Intent: The Search for a Compromise Between Formality and Adjudicative Justice”. *Connecticut Law Review* 34, no. 2 (2002): 453-476.

Shiffrin, Seana. “Harm and its Moral Significance”. *Legal Theory*, Volume 18 , Special Issue 3: Tort Theory (2012): 357 – 398.

Shoemaker, David. “Attributability, Answerability, and Accountability: Toward a Wider Theory of Moral Responsibility*”. *Ethics* 121, no. 3 (2011): 489-691.

Simmons, John. “Locke and the Right to Punish”. *Philosophy & Public Affairs* 20, no. 4 (1991): 311-349.

Sinnott-Armstrong, Walter. “It’s Not My Fault: Global Warming and Individual Moral Obligation,” *Perspectives on Climate Change: Science, Economics, Politics, Ethics Advances in the Economics of Environmental Research*, Vol 5(2005): 293–315.

Skoric, Marko M., Jia Ping Esther Chua, Meiyang Angeline Liew, Keng Hui Wong and Pei Jue Yeo. “Online Shaming in the Asian Context: Community Empowerment or Civic Vigilantism?” *Surveillance & Society* 8, no. 2 (2010): 181-199.

Smith, Angela M. “Control, responsibility, and moral assessment”. *Philosophical Studies volume* 138: 367-392.

- Smith, Angela. "On Being Responsible and Holding Responsible". *The Journal of Ethics*, Vol. 11, No. 4 (2007): 465-484.
- Smith, Richard H., J. Matthew Webster, W. Gerrod Parrott & Heidi L. Eyre. "The role of public exposure in moral and nonmoral shame and guilt". *Journal of Personality and Social Psychology* 83, no. 1 (2002): 138–159.
- Smith, Rohan. "Justine Sacco speaks about her experience one year after the tweet that ruined her life". *News.co.au*. 16 Feb 2016.
<https://www.news.com.au/technology/online/social/justine-sacco-speaks-about-her-experience-one-year-after-the-tweet-that-ruined-her-life/news-story/993b0788fdbda58794abb7ee53be04ed>.
- Solove, Daniel J., *The Future of Reputation*. New Haven and London: Yale University Press, 2007.
- Southwood, Nicholas. 'Does "Ought" Imply "Feasible"?' *Philosophy & Public Affairs* 44, no. 1 (2016): 1-95.
- Spratt, Michael. "The Presumption of Innocence is for Courtrooms, not Politics". *CBC News*, 30 Jan 2018. <https://www.cbc.ca/news/opinion/presumption-of-innocence-1.4509334>.
- Strauss Swanson, C., & Szymanski, D. M. "From pain to power: An exploration of activism, the #Metoo movement, and healing from sexual assault trauma," *Journal of Counseling Psychology*, 67, no. 6 (2020): 653–668.
- "Statistics about Sexual Violence". *nsvrc.org* (2012, 2013, 2015). Retrieved 30 Nov 2023, https://www.nsvrc.org/sites/default/files/publications_nsvrc_factsheet_media_packet_statistics-about-sexual-violence_0.pdf.
- Strawson, P.F. "Freedom and Resentment". *Proceedings of the British Academy* 48 (1962):187-211.

- Stemen, Don. *The Prison Paradox: More Incarceration Will Not Make Us Safer*. New York: Vera Institute of Justice, 2017.
- Stephen Bero. "The audience in shame". *Philosophical Studies* 177 (2020): 1283–1302.
- Strawson, P.F. "Freedom and Resentment". *Proceedings of the British Academy* 48 (1962). Retrieved 30 Nov 2023. <https://www.ucl.ac.uk/~uctytho/dfwstrawson1.htm>.
- "Swedish opera singer Anne Sofie von Otter blames #MeToo for her husband's suicide". *The Sydney Morning Herald*, 31 Jul 2018, <https://www.smh.com.au/entertainment/music/swedish-opera-singer-anne-sofie-von-otter-blames-metoo-for-her-husbands-suicide-20180731-h13de0.html>.
- Svindseth, Marit F., and Paul Crawford. *Humiliation: Mental Health and Public Shame*. Bingley: Emerald Publishing Ltd, 2019.
- Tadros, Victor. *The Ends of Harm: The Moral Foundations of Criminal Law*. Oxford: Oxford University Press, 2011.
- Talbert, Matthew, "Moral Responsibility", *The Stanford Encyclopedia of Philosophy* (2023), Edward N. Zalta & Uri Nodelman (eds.), <https://plato.stanford.edu/archives/fall2023/entries/moral-responsibility/>.
- Tandoc, E. C., Tan Hui Ru, B., Lee Huei, G., Min Qi Charlyn, N., Chua, R. A., & Goh, Z. H. "#CancelCulture: Examining definitions and motivations". *New Media & Society*, published online, (2022), <https://doi.org/10.1177/14614448221077977>.
- Tangney, June P., Rowland S. Miller, Laura Flicker & Deborah Hill Barlow. "Are shame, guilt and embarrassment distinct emotions?" *Journal of Personality and Social Psychology* 70 (1996): 1265–1269.

- Thomason, Krista. "The Moral Necessity of Anger." In *The Ethics Of Anger*, eds. Court D. Lewis and Gregory L. Block. Pennsylvania: Lexington Books, 2020.
- Thomason, Krista. *Naked: The Dark Side of Shame and Moral Life*. New York, NY: Oxford University Press, 2018.
- Thompson, J. D., and Cover, R. "Digital hostility, internet pile-ons and shaming: A case study". *Convergence* 28, no. 6 (2022): 1770-1782.
- Thomson, Judith Jarvis. *The Realm of Rights*. Cambridge, Massachusetts: Harvard University Press, 1990.
- Thomson, Judith Jarvis, "The Right to Privacy," *Philosophy & Public Affairs* 4, no. 4 (1975): 295-314.
- Thomson, Judith Jarvis. "The Trolley Problem". *The Yale Law Journal* 94, no. 6 (1985): 1395-1415.
- Torres, Walter J., and Raymond M. Bergner. "Humiliation: Its Nature and Consequences". *Journal of the American Academy of Psychiatry and the Law* 38, no. 2 (2010): 195-204.
- Tosi, Justin, and Brandon Warmke. "Moral Grandstanding". *Philosophy and Public Affairs* 44, no. 3 (2016): 197-217.
- Tuomela, Raimo, and Kaarlo Miller. "We-Intentions". *Philosophical Studies* 53 (1988): 367-389.
- Verkuyten, Maykel. "Perceived Discrimination and Self-Esteem Among Ethnic Minority Adolescents," *The Journal of Social Psychology* 138, no. 4 (1998): 419-493.
- Walen, Alec. "Proof Beyond a Reasonable Doubt". *Louisiana Law Review* 76 (2015): 355-446.

- Walen, Alec. "Retributive Justice". *The Stanford Encyclopedia of Philosophy*. Edited by Edward N. Zalta & Uri Nodelman, 2003. <https://plato.stanford.edu/archives/win2023/entries/justice-retributive/>.
- Walker, Margaret Urban, *Moral Repair*. Cambridge: Cambridge University Press, 2006.
- Wallbott H.G., & K. R. Scherer, "Cultural determinants in experiencing shame and guilt" in June P. Tangney & Kurt Fischer (eds) *Self Conscious Emotions: Shame: Guilt, Embarrassment, and Pride*. New York: The Guilford Press, 1995. pp. 465–487.
- Walsh, Joan. "Birding While Black: Just the Latest Bad Reason for White People to Call Police". *The Nation*. 26 May 2020. <https://www.thenation.com/article/society/amy-cooper-birding-police/>.
- Watson, Gary. "Two Faces of Responsibility". *Philosophical Topics* 24, no. 2 (1996): 227-248.
- Wenar, Leif. "Rights". *The Stanford Encyclopedia of Philosophy* (2023). Edited by Edward N. Zalta & Uri Nodelman. <<https://plato.stanford.edu/archives/spr2023/entries/rights/>>.
- Wellman, Christopher. "Rights and State Punishment". *The Journal of Philosophy* 106, no. 8 (2009): 419-439.
- Wellman, Christopher. "The Rights Forfeiture Theory of Punishment". *Ethics* 122, no. 2 (2012): 371-393.
- Weizel, Lavinia M. "The Process that is Due: Preponderance of the Evidence as the Standard of Proof For University Adjudications of Student-on-Student Sexual Assault Complaints". *Boston College Law Review* 53, no. 4 (2012): 1613-1655.
- Wittrock, John. "A human right to friendship? Dignity, autonomy, and social deprivation". *The International Journal of Human Rights* 26, no. 9 (2022): 1590-1607.