

Freedom of Expression and Social Coercion

A great deal of ink has been spilled regarding the importance of protecting freedom of expression from legislative curtailment by the state. As worthy and significant as much of that work is, there is a comparative dearth of reflection on a very closely related and similarly important issue pertaining to freedom of expression; indeed, one that has been made more salient than ever by the ways in which societies have changed in recent years. This is the way in which others' responses to our expression can themselves constitute potential penalties that amount to restrictions on our freedom to express. The scope for society to suppress free expression is pervasive, and fears about its potential to do so are longstanding, however they are often seen as secondary compared with trepidation relating to state-centred restrictions. Yet the need for more thoroughgoing theoretical reflection is all the more important given the apparent significance of such suppression in the evolution of western culture. Over the course of 20th and early 21st century, social changes have been ones in which the freedom to express in the face of pre-eminent social norms – for instance in regard to multi-ethnic relations, same-sex relations, religious conviction and dress – are not merely matters of civil and political rights but concern issues of social acceptability and societal pressure. More than this, socially transformative technological changes have made the role of social penalties on speech more salient than ever. Today's world is one in which our social engagement is increasingly taking on an online form. Not only are many more of our interactions with others taking place remotely, our online representation is an ever more important part of how we present ourselves to the world. This evolution in social interaction itself poses new puzzles concerning the character, value and effects of expression, and, indeed, new questions as to the desirability and possibility of the regulation of online expression.

In what follows I stake two related claims. Firstly, I contend that certain social costs that become characteristically consequential on expression can manifest a form of freedom-restricting coercion. This therefore involves contesting the view that the state or, indeed, formal institutions more

generally are unique in their capacity to restrict freedom of expression.¹ Accordingly, the paper denies that there is a fundamental difference, in respect of the damage to freedom of expression, between state imposed restrictions on expression on the one hand (whether in the form of prior restraint,² or by way of subsequent punishment) and societal penalty on the other. There are illuminating parallels to be drawn here between forms of public shame inflicted in the form of societal responses to speech and the more visceral method of shaming through use of restraining stocks in public squares, where the convicted are paraded before the gaze of the community and are represented as guilty and worthy of contempt and condemnation. Just as the stocks had a clear freedom-restricting character as a penalty whose role was, in part, to secure compliance with the law, so too shaming, through online media or otherwise, can have a similar punitive function inasmuch as it can enforce adherence to norms of conduct, belief and expression the deviation from which may invite a comparable pillory. Secondly, and in light of the fact that the freedom-restricting social costs issue are constituted by individuals' behaviour in response to expression, the paper commends two moral duties that pertain to the value of freedom of expression – a *duty of expressive toleration* which requires refraining from acting in ways that imposes expression-stifling costs on others; and a *duty of respect for expressive agency* which requires refraining from stigmatizing others for their expression.

The argument unfolds as follows: **Section I** traces worries about social coercion in the work of Mill and Tocqueville and situates the concerns of the paper with respect to so-called silencing arguments regarding freedom of expression. **Section II** outlines the claim that social coercion can restrict freedom of expression. This involves noting some of the ways in which the internet has augmented its potential to do so and detailing some of the potential costs constitutive of social

¹ This is consistent, of course, with maintaining there are some differences between state and social coercion that are of normative significance.

² Thomas I. Emerson, 'The Doctrine of Prior Restraint', *Law and Contemporary Problems* 20 (1955), pp. 648 – 671.

coercion. **Section III** puts forward two duties related to social coercion – a duty of expressive toleration and a duty of respect for expressive agency. **Section IV** concludes by responding to two objections.

I. MILL, TOCQUEVILLE AND SPEECH THAT SILENCES

Despite recent comparative scholarly inattention, the *social* threat to freedom of expression, as opposed to that embodied by the coercive mechanisms of the state, was an animating concern for two of the foremost thinkers of the 19th century. Both Mill and Tocqueville were preoccupied with a concern for societal pressure over behaviour and saw its role as one of the most important questions concerning the character of a democratic culture and the scope it allows for citizens to express themselves freely. Mill, for instance, is explicit that his harm principle is supposed to govern both ‘legal penalties’ and ‘the moral coercion of public opinion’ and is at pains to point out the particularly grave threat to expression embodied by forms of social coercion outside of the state. He writes:

*“when society is itself the tyrant... its means of tyrannising are not restricted to the acts which it may do by its political functionaries... and if it issues wrong mandates instead of right, or any mandates at all in things with which it ought not to meddle, it practises a social tyranny **more formidable** than many kinds of political oppression, since, though not usually upheld by such extreme penalties, it leaves fewer means of escape, penetrating much more deeply into the details of life, and enslaving the soul itself... therefore... there needs protection also against the tyranny of prevailing opinion and feeling; against the tendency of society to impose, by other means than civil penalties, its own ideas and practices as rules of conduct.”³*

³ J. S. Mill, *On Liberty* in *The Collected Works of John Stuart Mill, Volume XVIII – Essays on Politics and Society Part I*, John M. Robson (ed.) (University of Toronto Press, 1977), p 220 (my emphasis).

Here, as elsewhere in *On Liberty*, there is a worry about the threat to freedom of thought and expression in particular. In the passages in which Mill reflects on ‘vituperative and intemperate expression’ he is quick to note the particular vulnerability of those with less popular views to such potentially hostile forms of expression. He remarks how “*In general, opinions contrary to those commonly received can only obtain a hearing by studied moderation of language, and the most cautious avoidance of unnecessary offence... while unmeasured vituperation employed on the side of prevailing opinion, **really does deter people from professing contrary opinions**, and from listening to those who profess them.*”⁴ We find very similar concerns present in Tocqueville’s reflections on the state of America, and there seems little doubt that his own observations influenced Mill in this regard.⁵ It is with reference to the social restrictions on expression that Tocqueville says that “*I know of no other country in which there is so little true independence of mind and freedom of discussion as in America.*”⁶ Though they were to a great extent free of legal penalty on the expression of opinion, the people of America remained subject to the moral coercion of public opinion. He writes that “*In America the majority raises very formidable barriers to the liberty of opinion: within these barriers an author may write whatever he pleases, but he will repent if he ever step beyond them.*”⁷ If such fears about the repressive potential of societal opinion are well-grounded, the relative lack of scholarly attention on socially imposed restrictions on freedom of expression is both surprising and regrettable.⁸

⁴ Ibid. p 209 (my emphasis).

⁵ J. S. Mill, ‘De Tocqueville On Democracy in America’, in *The Collected Works of John Stuart Mill, Volume XVIII – Essays on Politics and Society Part I*, John M. Robson (ed.) (University of Toronto Press, 1977), at p 81

⁶ A. de Tocqueville, *Democracy in America Volume 1*, Henry Reeve (trans.) Bruce Frohnen (ed.) (Regnery, 2002), at p 209

⁷ Ibid.

⁸ For a notable exception see Frank Michelman, ‘Conceptions of Democracy in American Constitutional Argument: The Case of Pornography Regulation’, *Tennessee Law Review* 56 (1989), pp 291 – 319. See also Charles Lawrence, ‘Crossburning and the Sound of Silence: Antisubordination Theory and the First Amendment’, *Villanova Law Review* 37 (1992), pp. 787 – 804, at p 803. For an insightful discussion of how fear of social sanction such as ostracism can result in expressive reticence, albeit one which frames the issue in terms of self-censorship rather than societal restrictions on freedom, see Glenn Loury, ‘Self-Censorship in Public Discourse: A Theory of “Political Correctness” and Related Phenomena’, *Rationality and Society* 6 (1994), pp. 428 – 461.

There is a very notable exception to this relative lacuna in attention on the expression-damaging potential of expression itself. This is in the form of work on the so-called silencing effect of speech. Though I will not directly assess the merits of the arguments offered in that work, the argument I pursue here has an obvious kinship with them. However, as I will explain, the argument of the paper is importantly distinctive from them. Much of the work which argues that speech can have a silencing effect involves an appeal to one or more of the following four main dynamics. Firstly, that speech can crowd-out the speech opportunities for others.⁹ If speech opportunities are scarce then some person's speaking can deprive others of the opportunity to do the same. Secondly, that speech can damage others' sense of confidence in the worth of their speech.¹⁰ This could be because of damage to the self-respect or self-esteem of would-be speakers, such that they lack a self-conception of being a worthy speaker, or it could be because they simply have no confidence in their speech being recognised as worthy by others (if I don't believe anyone will listen, what is the point in speaking in the first place?). Thirdly, that speech can damage the credibility of others as relevantly worthy speakers in the eyes of others.¹¹ In that sense persons aren't listened to or believed by others. Fourthly, and very closely relatedly, that speech can disable the meaning of others' speech, such that a person's speech either fails to qualify as the intended speech-act, or that it fails to communicate the intended idea to the hearer.¹²

⁹ See for example Owen Fiss, *The Irony of Free Speech* (Harvard University Press, 1996)

¹⁰ Caroline West, 'The Free Speech Argument Against Pornography', *Canadian Journal of Philosophy* 33 (2003), pp. 391 – 422, at p 398; Caroline West, 'Words that Silence? Freedom of Expression and Racist Hate Speech', in *Speech and Harm: Controversies Over Free Speech*, I. Maitra and M. K. McGowan (eds.), (Oxford University Press, 2012), pp. 222 – 247, at p 236; Rae Langton, 'Speech Acts and Unspeakable Acts', *Philosophy and Public Affairs* 22 (1993), pp. 293 – 330, at p 315.

¹¹ Most prominently, in the literature, through the publication of pornography. See Catherine MacKinnon *Feminism Unmodified: Discourses on Life and Law*, (Harvard University Press, 1987), at p 193. See also Miranda Fricker, *Epistemic Injustice: Power and Ethics of Knowing*, (Oxford University Press, 2007). In particular, Fricker discusses (at p 131) what she terms 'pre-emptive testimonial injustice', under which persons are not only accorded an unjustly lower credibility as speakers, they are, on that basis, not solicited to speak.

¹² Rae Langton, 'Speech Acts and Unspeakable Acts'; Ishani Maitra, 'Silencing Speech', *Canadian Journal of Philosophy* 39 (2009), pp. 309 – 338.

I am not here going to offer an extensive exploration of the ways in which speech might be thought to silence, nor will I seek to defend the idea that speech can indeed silence in those senses. Nevertheless, it is worth clarifying that work's distinctness from, and connection to, the concerns of this paper. These silencing arguments all concentrate on what might broadly be understood as the *difficulty* associated with expressing something. Whether because a person's speech isn't efficacious, or they aren't offered the opportunity to speak, or because they lack the internal resources of confidence and self-esteem to speak up, expression is rendered more *difficult* in the sense that the possibility of successful expression (whatever that consists in) is lowered.¹³ In contrast, I am concerned first and foremost with the *costs* associated with expressing something by way of the *burdens* one might incur as a result of expressing.¹⁴ So I share with the silencing arguments a concern for the *social environment of speech*,¹⁵ but my focus is specifically on the ways in which people's expression renders them vulnerable to certain social costs in a way which can stifle their freedom of expression. In that sense one can read the following as, in part, the elaboration of a dynamic of 'silencing' that is insufficiently explored in existing work.

II. SOCIAL COERCION AS UNFREEDOM

The social threat to free expression is far from a new phenomenon, however the expansion of the role played by the internet in our social presence makes the matter especially acute.¹⁶ It is probably not hyperbole to suggest that, for many, the internet has transformed the number of others we are able to interact with and the character of many of our interactions. In that sense the internet has,

¹³ In this way they more closely resemble a social equivalent of prior restraint.

¹⁴ As Cohen illustrates the distinction, "*At the far end of the difficulty continuum lies the impossible, but it is the unbearable which occupies that position in the case of costliness.*" G. A. Cohen, 'On The Currency of Egalitarian Justice', *Ethics*, 99 (1989), pp. 906 – 944, at 919.

¹⁵ Susan J. Brison, 'The Autonomy Defence of Free Speech', *Ethics* 108 (1998), pp. 312 – 339, at 314.

¹⁶ For a valuable collection of such work analysing some of its lamentable dimensions see Saul Levmore and Martha C. Nussbaum (eds.), *The Offensive Internet* (Harvard University Press, 2012)

in a manner of speaking, democratised the means of widespread publication.¹⁷ The accessibility of online communication and the audience reach of what we say have resulted in the vast expansion in the bounds of expression, and we can now communicate on an ongoing basis, openly or anonymously, with others with whom we would never in the past been able to converse. We only need switch on a device and type away. Nowhere is this more noticeable than respect of our far greater vulnerability to defamation. Whereas slander of the past tended to reach only a limited audience, seldom extending beyond the bounds of a more local community, now it can spread virulently far afield.¹⁸ And the persons who make up that audience are more able than ever to add to the chorus of defamation; proliferating the verbal assault. The internet makes our world both larger and smaller all at the same time. Not only are there considerable potential real-world consequences that result from one's reputation coming under attack, say in the form of one's employment prospects, but our online presence itself can be sullied and stained. Insofar as our online interactions are of increasing salience to us, losing friends on Facebook, followers on Twitter, likes on Instagram, views on Youtube, journal article downloads, not to mention the ability to engage in respectful, abuse-free dialogue with others, are potential costs resulting from a damaged reputation that shouldn't be discounted. Any such reputational damage is only heightened by apparent semi-permanence of the stains on reputation recorded in the online domain and their barnacle-like resistance to being cleansed away over time.¹⁹ The vulnerability we have to others' speech, especially speech with the potential to inflict specifically expression-stifling costs, has been magnified beyond all recognition.

¹⁷ Timothy Garton Ash, *Freedom of Speech: Ten Principles for a Connected World* (Atlantic Books, 2016)

¹⁸ Brian Leiter, 'Cleaning Cyber-Cesspools: Google and Free Speech', in Levmore and Nussbaum, *The Offensive Internet*, pp. 155 – 173.

¹⁹ As Danielle Keats Citron eloquently puts it, "[t]he Internet guarantees a Nietzschean "eternal return" of damaging disclosures." See her 'Mainstreaming Privacy Torts', *California Law Review* 98 (2010), pp. 1805 – 1852, at p 1813. See also Daniel Solove, *The Future of Reputation: Gossip, Rumor, and Privacy on the Internet* (Yale University Press, 2008); Daniel Solove, 'Speech, Privacy, and Reputation on the Internet', in Levmore and Nussbaum, *The Offensive Internet*, pp. 15 – 30; Anupam Chander, 'Youthful Indiscretion in an Internet Age', in Levmore and Nussbaum, *The Offensive Internet*, pp. 124 – 139;

One incident which has to some extent become a touchstone concerning matters of reputational assault of this kind, and, indeed, a stark and poignant instance of it, is the case of Justine Sacco, who tweeted what might be best described as an ill-considered joke that very quickly spread globally and led to her becoming widely maligned and vilified.²⁰ Shortly before getting on a flight abroad, Sacco sent the following tweet “Going to Africa. Hope I don't get AIDS. Just kidding. I'm white!” By the time she had landed, her comment had spread online and incurred thousands of aspersing, often denigrating, reactions from the hastily gathered virtual lynch mob. Her public image was devastated and the vociferously negative online response near totalising.²¹ Another infamous case involved Lindsey Stone, who posted a photo of herself, with a friend, mocking a sign outside Arlington National Cemetery calling for ‘silence and respect’. The image spread virally online and resulted in a 30,000 member Facebook page, and a petition signed by thousands, both demanding she lose her job. Stone received several death and rape threats, and she did lose her job. These are just two of a plethora of such instances of public shaming, increasingly evident from online exchanges; notable because they are extreme but not because they are otherwise unique. They are instances of an increasingly familiar and widespread tendency for instances of expression, deemed in some way objectionable, to invite a powerfully withering response.²²

Though the harms of such shaming are increasingly under the normative-theoretical microscope, one of the chief mischiefs of the phenomena implicated in that shaming – its propensity to stifle

²⁰ For a thorough discussion see Jon Ronson, *So You've Been Publicly Shamed* (Picador, 2015)

²¹ *Ibid.* pp. 63 – 86.

²² For some of the work that explores such cases see Danielle Keats Citron, *Hate Crimes in Cyberspace* (Harvard University Press, 2014); Danielle Keats Citron, ‘Civil Rights in Our Information Age’, in Levmore and Nussbaum, *The Offensive Internet*, pp 31 – 49; Mary Anne Franks, ‘Sexual Harassment 2.0’, *Maryland Law Review* 71 (2012), pp. 655 – 704; Kristine Gallardo, ‘Taming the Internet Pitchfork Mob: Online Public Shaming, the Viral Media Age, and the Communications Decency Act’, *Vand.J. of Entertainment and Tech Law* 19 (2017), pp. 712 – 746; Winhkong Hua, ‘Cyber mobs, Civil Conspiracy, and Tort Liability’, *Fordham Urban Law Journal* 44 (2017), pp. 1217 – 66; Kate Klonick, ‘Re-Shaming the Debate: Social Norms, Shame, and Regulation in an Internet Age’, *Maryland Law Review* 75, pp. 1029 – 1065.

free expression – remains largely neglected, at least in the context of academic scholarship.²³ And yet there are clear parallels between the ways in which the state-imposed costs restrict freedom of expression and the manner that societal costs can. The state ensures adherence to the law through the threat of legal punishment just persons can be disciplined into expressive conformity by the spectre of social penalty. Of course, one of the ways in which the state restricts freedom is through incarceration. Incarceration restricts individual freedom by depriving ability. However, the *threat* of incarceration – like many other forms of state-imposed penalties – has freedom restricting force as a cost threatened as punishment for the transgression of law. Now, just as legal sanctions that restrict freedom may be justified, so too societal restrictions might be justified. Although this is an important point – and one to which I will return – it does not diminish the importance of recognising the freedom-restricting force of social costs and appreciating that they may stand in need of justification in the same way. Those sceptical of the parallel I draw between state and social coercion may well point to features of state-imposition of penalties that differ from societal imposition and which might be considered essential to explain its capacity to restrict free expression. I will respond to some such doubts later in the paper but at this stage it's important to first lend credence to the claim that some social costs are of the relevant character and gravity to constitute costs that have the potential to restrict expression.

Whilst my concern with social coercion isn't exclusively tied to the online domain, it doesn't stretch credulity to conclude that the digital era has amplified its danger. Online participation is an increasingly important part of individuals' social existence and an ever-more central part of how social identity is constructed and formed. One of the defining features of the internet era that enhances the reach and power of social coercion as an expression-stifling phenomenon is the fact that the internet is not a self-contained bubble. To sharpen this thought it's worth appreciating

²³ For a recent exception see Danielle Keats Citron, 'Restricting Speech to Protect it', in *Free Speech in the Digital Age*, Susan J. Brison and Katherine Gelber (eds.), (Oxford University Press, 2019), pp. 122 – 136, at pp. 129 – 130.

that there is a *location* question both with respect to the expression that is shamed, and in respect of the costs which manifest the shaming. The locus of the target of the shaming can be *online expression*, say an offensive blog post, or *offline expression*, say a speech at a political rally. The locus of the shaming penalties can also involve *online costs*, say in the form of a cascade of tweeted insults, or *offline costs*, say in the form of losing employment, or being shunned by neighbours who no longer greet you as you walk down the street. Although clearly rough and permeable, such a schematic can clarify how the online world has changed the extent and character of societal shaming. One of the important things to draw attention to here is that the persons who are subjected to online shaming are not only those who are participants in online publication. Rather, we can be dragged into the online conversation against our will, so to speak, when our offline behaviour becomes the subject of online shaming.²⁴ And if real-life spills into the online domain in this way, the same is true of the overspill in the other direction. The costs inflicted by online shaming go beyond the harsh words and vitriol published online but can include profound effects on one's broader, offline existence. The reciprocal relation between the 'real' world and the 'virtual' one is ongoing and fluid in a way that, in certain respects, heightens the threat to free expression.²⁵

Now, a thoroughgoing analysis of the necessary weight or features that might make certain forms of online shaming expression-restricting is beyond what can be argued here, but extant work exploring the badness of hate speech offers some illumination. Work on hate speech is voluminous and the varieties of different speech claimed from different perspectives as instances of hate speech

²⁴ To give one example, at a conference in 2013 Adria Richards tweeted a photo of two male audience members with the caption "Not cool. Jokes about forking repo's in a sexual way and "big" dongles. Right behind me." The tweet led to the organiser of the conference reprimanding the two individuals and reportedly culminated in one of the two being released from his job. In a cruel twist to the case the shaming rebounded on Richards as she was personally subjected to online vitriol and threats and lost her own job. <https://www.theverge.com/2013/3/21/4132752/thug-mentality-how-two-dick-jokes-exploded-into-ddos-and-death-threats>. See also Klonick, 'Re-Shaming', at pp. 1030 – 1031

²⁵ Although I distinguish between online and offline costs the distinction is not a deep one, and they are seldom neatly segregated from one another in actual instances of social coercion.

indicate the elusiveness of an uncontested category.²⁶ For the purposes here it suffices to note certain ways in which speech has been said to inflict harm on, or entail a serious wrong against, others. Quite apart from the silencing potential of hate speech touched on in the above, hate speech has been argued to involve, amongst other things, dignitary harm,²⁷ community libel,²⁸ subordination,²⁹ psycho-emotional harm³⁰ and damage to self-respect.³¹ Though I cannot hope to explore and adjudicate this variety of ways of understanding the badness or wrongness of hate speech, insofar as the arguments in their favour have strength they offer a supportive backdrop for the argument pursued here. For they suggest that others' expression can inflict real and significant harms (or wrongs) on others. Insofar as individuals might become subject to harms of this kind *as a result of expression*,³² this itself suggest that the costs to certain expression can have the right character to restrict the freedom to express, either because they are harms that persons would

²⁶ For an account that emphasises the importance of disaggregating 'hate speech' see Alexander Brown, 'What is Hate Speech? Part 1: The Myth of Hate', *Law and Philosophy* 36 (2017), pp. 419 – 468. See also his *Hate Speech Law: A Philosophical Examination* (Routledge, 2015).

²⁷ Jeremy Waldron, *The Harm in Hate Speech*, (Harvard University Press, 2014)

²⁸ Bhikhu Parekh, 'The Rushdie Affair: Research Agenda for Political Philosophy', *Political Studies* 38 (1990), pp. 695 – 709.

²⁹ Mari Matsuda, 'Legal Storytelling: Public Response to Racist Speech: Considering the Victim's Story', *Michigan Law Review* 87 (1989), pp. 2329 – 2334; Andrew Altman, 'Liberalism and Campus Hate Speech: A Philosophical Examination', *Ethics* 103 (1993), pp. 302 – 317.

³⁰ Richard Delgado and Jean Stefancic, *Must We Defend Nazis?* (New York University Press, 1997)

³¹ Jonathan Seglow, 'Hate Speech, Dignity and Self-Respect', *Ethical Theory and Moral Practice* 19 (2016), pp. 1103 – 1116.

³² Now, there might be suspicion that this argument rests on something of an analytical sleight of hand. It might be said, for instance, that the harms of hate speech are precisely marked by a verbal assault on the *status* or *nature* of the person, as opposed to constituting an attack on what persons express. Hate speech is strongly associated with racist and sexist speech for good reason; because it characteristically targets the identity of the victim and not their words or views. If so, the harms of hate speech cannot be included among the costs that are invited by *expression* and hence their support cannot be recruited for the argument pursued thus far. Although, there is some truth to this thought, it involves too crude an understanding of identity, as well as resting on a fallacy. Firstly, even if the harms of hate speech are inflicted on, and targeted towards, the status or identity of the harmed agent, it is wrong to suppose that this necessarily means that they do not pertain to the expression of the agent. Our social identity is often intimately bound up with the beliefs we hold and express and we are understood by others, and ascribed a certain status, often partly because of how we express ourselves. I believe it is reasonably clear, for instance, that Muslim women who express their religious identity through the wearing of a Burka could be subject to dignitary harm, libel, or subordination precisely *as Burka-wearing Muslim Women*. That is, that they are denigrated precisely for the expression of belief. Furthermore, even if it were true that the harms of hate speech involve damage to a status that is necessarily conceptualised independently of the expression of the affected agent, the *prompt* for a verbal attack is not necessarily the same as its *target*. It may have been engaging in protest against the injustice of US segregation that made some civil rights activists the targets of certain instances of racist abuse, for instance.

rationaly prefer to avoid, or because they are wrongs that persons would rationaly prefer not to be subject to.³³

Even insults that don't necessarily inflict the harms that hate speech is claimed to involve can be potentially devastating when conjoined together. In the cases of Stone and Sacco, any one of the many insults and associated other berating comments may have been comparatively easy to cope with when taken in isolation. However, it is the cacophony of aspersion that made the responses so formidable and overwhelming.³⁴ Alongside the threats of harm and death that are also all too often integrated into such hostilities, it is the conglomeration of denigrating verbal barbs that makes the *collective* assault such a profound threat to free expression.³⁵ Faced with such a prospect, I hazard that it is far from a stretch to believe that people can find themselves shrinking into dialogical submission.

Within the range of possible *offline* social costs which might be thought to fall into the category of social coercion, the loss of employment may be among the least controversial as a potential limit

³³ See Matsuda, 'Legal Storytelling' p 2337 where she connects hate speech with restrictions on freedom, saying "[i]n order to avoid receiving hate messages, victims... curtail their own exercise of speech rights." Now, quite plausibly, it is often thought that hate speech supervenes on pre-existing relations of subjugation and oppression. Matsuda suggests, for instance, that hate speech harms because it reinforces "*historically vertical relations*," and is directed against "*a historically oppressed group*." Ibid. pp. 2357 – 2358. This might well imply that some groups are more vulnerable to the harms of hate speech than others, and, in turn, more vulnerable to its expression-stifling force. In appealing to the harms of hate speech I do not intend to downplay the role that context and extant injustice plays in underpinning and aggravating the harm. However, it seems quite plausible to maintain that even if some of the associated verbal harm depends on this, there are further dimensions which remain relevantly injurious.

³⁴ It has become familiar to refer to attacks of this kind as 'cyber mobs', or, because of their platform deployment of choice, 'Twitter mobs'. 'Mob' is an apt descriptor. Being confronted by a mob is an unquestionably unpleasant experience. The mob can be unpredictable, self-emboldening, its hostile purposes can build towards frenzy; it can feel powerful and inescapable. The multiple insults unleashed by the cyber mob conjoin to encourage a collective tone of rejection and disparagement.

³⁵ The recent history of the internet is replete with cases where expression has invited a flurry of such threats. One such case is that of Caroline Criado-Perez, whose campaign in 2017 to feature women on UK banknotes resulted in a torrent of threats and abuse, to the point where she claimed to be receiving up to 50 rape threats an hour. https://www.huffingtonpost.co.uk/2013/07/27/twitter-rape-abuse_n_3663904.html. For the contention that such abuse tends to be disproportionately targeted towards, or especially impactful on, women see Citron, 'Civil Rights', p 32; Chander 'Youthful Indiscretion', pp. 126 – 17; Franks, 'Sexual Harassment', p 658.

on free expression. Indeed, it is a longstanding rationale for the practice of academic tenure that it fosters a culture of free enquiry, insulated from the pressures of conforming to accepted points of view.³⁶ Having a job can have great value for a person. Amongst many other things, a job can offer income, status, self-respect and structure to one's life. Quite apart from those generic reasons why a job might be of value, for many people the particular job they have is something which they are personally invested in as a self-defining vocation that forms a part of what makes the life they lead of distinctive significance for them. Persons are not merely employees, or income earners, but artists, teachers, bankers, police officers and so on.

Now, loss of employment may be relatively widely accepted as, at least in some cases, a potential freedom-limiting cost, so flagging its expression-restricting potential is not revelatory. What I want to draw attention to, though, is the way in which social hostility or a certain degree of social consensus can result in pressures that have the potential to result in either employment loss, or an inability to gain employment. One case that clearly illustrates the danger in question – albeit one without a genesis in online speech specifically – is the now prominent saga of Colin Kaepernick and the NFL. In broad strokes, Colin Kaepernick is an American Football quarterback who was drafted in the second round of the 2011 NFL draft by the San Francisco 49ers. The controversy stoked by Kaepernick related to his unwillingness to stand for the singing of the national anthem; something which had become a tradition for NFL players since 2009. When Kaepernick was asked why he chose to sit during the national anthem he alluded to police brutality and racial injustice, and answered, “I am not going to stand up to show pride in a flag for a country that oppresses black people and people of color,” and that “[t]o me, this is bigger than football, and it would be selfish on my part to look the other way. There are bodies in the street and people getting paid

³⁶ Alluding to the silence that can attend to the non-tenured, Matsuda rhetorically asks “[w]hat don't we hear when some young scholar chooses tenure over controversial speech?” Mari Matsuda, “Who Owns Speech,” *Address at Hofstra School of Law*, (Nov. 13 (1991))

leave and getting away with murder.”³⁷ Kaepernick’s protest quickly gained media traction and he and those who joined him in kneeling³⁸ were roundly condemned by others, including even President Trump who called for the NFL players involved (including, by implication, Kaepernick) to be fired.³⁹ Shortly after the end of the season in which the anthem protests took place, Kaepernick’s contract ended. Though Kaepernick himself opted out of the final year of his contract, the 49ers general manager John Lynch made it clear that the contract was due to be voided by the team in the advent that Kaepernick hadn’t pre-empted this.⁴⁰ At the time of writing, Kaepernick does not have a job as an NFL quarterback. The situation echoed the blacklisting emblematic of the height of McCarthyism. Although his exclusion was not public or explicit, the fact that Kaepernick was out of the NFL in the season that followed his role in the anthem protests was an evident example to the players league-wide who could reasonably anticipate a similar fate. The example of Kaepernick is a significant one in part because of the sense in which the anthem protests were as much a matter of conscientious objection as they were proactive protestations. In the first instance, this was just as much a matter of what Kaepernick *did not want to express*, as much as it was a mode of positively communicating to others.⁴¹ The explicitly stated motivation from

³⁷ Steve Wyche, ‘Colin Kaepernick explains why he sat during national anthem’, NFL.com, August 27th 2016

³⁸ Kaepernick initially simply sat during the playing of the anthem but then began kneeling, reportedly after discussing the issues with the US Army veteran and former NFL player Nate Boyer. Matthew Impelli, ‘A Timeline of Colin Kaepernick’s Kneeling In Protest To His Upcoming Workout for NFL Teams’ <https://www.newsweek.com/timeline-colin-kaepernicks-kneeling-protest-his-upcoming-workout-nfl-teams-1471819>

³⁹ In a speech, Trump said about the protests, "That's a total disrespect of our heritage. That's a total disrespect of everything that we stand for," and "Wouldn't you love to see one of these NFL owners, when somebody disrespects our flag, to say, 'Get that son of a b**** off the field right now. Out! He's fired. He's fired!'".

⁴⁰ Nick Wagoner, ‘If Colin Kaepernick didn’t opt out, 49ers would have released QB’, ESPN.com March 4th 2017

⁴¹ Joshua Cohen says about such cases that where an agent’s concern is “*to ‘bear witness’ [and in her] view places her under an obligation to speak out... expressive liberty is on a footing with liberty of conscience, and regulations are similarly burdensome.*” See Joshua Cohen, ‘Freedom of Expression’, *Philosophy and Public Affairs* 22 (1993), pp. 207 – 263, at p 225 (his emphasis).

Kaepernick was to refrain from endorsing a symbol of a society that, in his view, at that time embodied profound forms of injustice.

Loss of employment resulting from expression frequently manifests itself in ways that are still more direct than the Kaepernick saga. In a famous case, pre-dating the prominence of social media, the then England football team manager, Glenn Hoddle, was fired from his role after giving an interview in which he implied that persons suffering from disabilities were, by way of those disabilities, being punished for the sins of their prior lives.⁴² More recently, William Sitwell, a prominent UK food critic, was forced to resign his post as chief editor of the Waitrose Food Magazine following the dissemination of an email response he sent to a freelance journalist in which he joked about the possibility of producing a series on “killing vegans.”⁴³ A similar case is that of Mary Bono, then interim boss of US Gymnastics, who resigned after criticism, including from prominent US athletes, concerning Bono’s own negative Tweets about Nike. Bono had posted an image of her redacting the company logo on some Nike trainers along with the message “...unfortunately I had these shoes in my bag. Luckily I had a marker in my bag too...”⁴⁴

Such are a select few instances of the now very familiar phenomenon of losing employment as a consequence of expression, united not by the substance or character of the speech they involve but by the social penalties they invited. Now, to be clear, the claim here is not that employment loss is never a justified consequence of expression. There are multiple and varied considerations which bear on the justification for losing employment as a result of expression and there seem to

⁴² John Arlidge and Patrick Wintour, ‘Hoddle’s future in doubt after disabled slur’, *theguardian.com*, January 30th 1999, <https://www.theguardian.com/football/1999/jan/30/newsstory.sport7>.

⁴³ Olivia Petter, ‘Waitrose Food Magazine Editor Resigns After Sending Email About ‘Killing Vegans’’, *https://www.independent.co.uk/life-style/food-and-drink/waitrose-food-magazine-editor-william-sitwell-resign-killing-vegans-a8610106.html*

⁴⁴ Interestingly, Bono’s criticism of Nike was itself a response to the brand’s support for Kaepernick, and the decision of the company to adopt his image to front a new marketing campaign.

be clear instances in which this might be justified.⁴⁵ What I am concerned with here specifically, though, is what might be called the *social acceptability* of speech; whether, for instance, the fact that Kaepernick's expression was disagreeable, either to NFL owners, or to parts of society at large, is a compelling basis for thinking that losing employment was an unproblematic consequence. I suggest that the prospect of employment loss has clear expression-restricting potential, and that the considerations that justify the manifestation of those freedom-restrictions must themselves be weighty ones. Very few argue for censorship on the basis of *mere* social acceptability. That is, if speech considered socially unacceptable is legitimately subject to regulation it is not *qua* its social unacceptability but in view of some other feature. Why should the expression-restricting pressures imposed by social pressures on employment be any different? Inasmuch as freedom of expression can be stifled by the spectre of employment loss, we need similarly weighty reasons for thinking that pressures on employment on the basis of social acceptability can be vindicated.⁴⁶

Though I have singled out employment for special consideration it is really a particular, albeit often significant, instance of the more general category of association with others, and the suggestion here is that disassociation from others, understood as the damage done to relationships with others, broadly construed, can constitute a freedom-restricting social cost. As such, I have in mind a category of 'association' which encompasses a range of things from the formal ejection or exclusion of persons from institutionalised groups, to the corrosion of friendships or even acquaintances. The claim here, then, is that damage to relationships across a wide range of contexts can be of real significance and can thereby reflect, in principle, expression-restricting considerations. Losing association with others can have real costs to us, and can involve the loss of group membership and associated privileges and status that affords; friendships and closeness

⁴⁵ Indeed, there are vocations where restrictions on expression seem integral to the successful performance of the role itself, such as patient confidentiality in the case of doctors.

⁴⁶ The interest persons have in *refraining* from expressing something with which they deeply disagree, as in the Kaepernick case, is a particularly weighty reason counting against such a loss.

with others; valuable time and activities; and identification and affiliation, among other things. We are sociable beings, who enjoy, and sometimes yearn for, others' company and community. On the one hand this is the impetus behind so much of what is enriching about our lives, on the other hand it is precisely this sociability which, as Jeremy Waldron aptly puts it, "*constitutes a dreadful vulnerability on the part of the individual to social pressure and social coercion.*"⁴⁷ At the risk of putting it too bluntly, legal pressure and coercion limits freedom as we fear imprisonment; social pressure and coercion limits freedom as we fear ostracism and exclusion.

III. DEFUSING THE 'ENGINES OF MORAL REPRESSION'

Among the most significant features of social coercion, and a reason for thinking it is importantly different from state regulation, is that it is itself constituted by the free choices and responses of others. It is therefore the very freedom of others, including their free expression, that gives rise to the costs that can be expression-restricting. In one of the few explicit examinations of the role of social coercion in Mill's thought, Waldron again captures this very well when he says that "*the "engines of moral repression" and "the influences hostile to individuality" are not alien forces: they are nothing but the upshot of people's actions and inclinations in a social context. They comprise things that people want to do, things they feel like doing... people may have a right—or it may for other reasons be desirable for them to have the freedom—to do the things which, at least when taken en masse, constitute the social coercion which Mill is attempting to stop.*"⁴⁸ This sage observation might be thought to lend plausibility to a line of resistance to my concerns. I have been suggesting that it is important to attend to the freedom-restricting character of social coercion because doing so highlights the fact that such coercion stands in need of

⁴⁷ Jeremy Waldron, 'Mill as a Critic of Culture and Society', in David Bromwich and George Kateb (eds.) J. S. Mill *On Liberty* (Yale University Press, 2003), pp. 224 – 245, at p 227.

⁴⁸ Waldron, 'Mill as a Critic of Culture and Society', p 231.

justification and, in principle, some form of redress. However, it might be argued that precisely because social coercion is constituted by individuals' free behaviour and expression it is straightforwardly justified by whatever values commend leaving people free to express themselves.

For instance, a critic might observe that what I have referred to as 'social acceptability' often comes to have an influence on employment decisions as a result of members of the public collectively exercising their own rights to object to certain forms of expression. Indeed, this is exemplified in the Kaepernick case, where NFL owners were plausibly under commercial pressures to sustain the popularity of the sport and avoid potential boycotts from fans who take exception to players protesting in a way many such fans deem disrespectful. It might be suggested, then, that the loss of employment is justified in virtue of simply being a prudential commercial response to the manifestation of others' free association and consumer choices. I maintain, however, that so crude a conclusion ignores important ethically salient dimensions of the situation. To see this, it's worth noting the clear parallels between cases such as Kaepernick's and instances of discrimination on the basis of so-called reaction qualifications: where a person's qualifications for a role depends on how others are likely to react to them.⁴⁹ An example of a reaction qualification is politeness, where politeness is a relevant qualification precisely in virtue of the fact that customers would react positively towards a polite employee and, perhaps, patronise the establishment with greater regularity. That there are some legitimate reaction qualifications does not show that reaction qualifications in general (still less in all cases) are legitimate bases for discriminating between candidates or, indeed, bases for dismissal. If a sports team rejects a player because its fans won't pay to watch a player of from that racialised group it would be a profound injustice. In view of

⁴⁹ Alan Wertheimer, 'Jobs, Qualifications, and Preferences', *Ethics* 94 (1983), pp. 99 – 112; Kasper-Lippert Rasmussen, *Born Free and Equal? A Philosophical Inquiry into the Nature of Discrimination* (Oxford University Press, 2013), chapter 9 pp. 235 – 260; Andrew Mason, 'Appearance, Discrimination, and Reaction Qualifications', *The Journal of Political Philosophy* 25 (2017), pp. 48 – 71.

this, the simple fact that employment pressures are the result of others' free behaviour does not tell us decisively that the costs to freedom of expression are unproblematic.

Given this, I propose the first of the two duties we have as individuals to obviate the damage wrought on freedom of expression by social coercion. This is what I will term the *duty of expressive toleration*. This is a duty to refrain from acting in ways that one could reasonably avoid and one could reasonably foresee would, if generalised, risk the imposition of otherwise unjustified expression-stifling costs on others. The duty of expressive toleration thereby reflects a moral reason against being party to forms of behaviour that collectively risk manifesting social coercion. In this sense it involves a recognition of the expressive interests of others and the significance of one's own actions as contributors to an environment of speech that can be more or less hostile to others' expression.⁵⁰

Applying this to the Kaepernick case, suppose that Kaepernick's failing to retain employment in the NFL is explained by the fact that NFL teams feared revenue loss from fans boycotting any team that employed him. The duty of expressive toleration holds that, from the perspective of the individual fan, participation in the boycott is itself *pro tanto* unjust if i. the fan could reasonably foresee the expression-stifling effects of such a boycott (restricting the freedom to kneel during the playing of the anthem); ii. they did not themselves have interests at stake sufficient to warrant engaging in the boycott, and iii. the interests served by the boycott in respect of the its expression-stifling impact are not themselves significant enough to justify expression-restrictions along with the concomitant damage to whatever values are served by persons enjoying the freedom to expression. Let me briefly unpack these elements.

⁵⁰ For two accounts which explore the interests persons have in expression see Joshua Cohen, 'Freedom of Expression' and Tim Scanlon, 'Freedom of Expression and Categories of Expression', *University of Pittsburgh Law Review* (1978 – 79), pp.519 – 550. 223 – 225.

Under i. the duty holds that the individual's responsibility for their conduct depends on whether it is reasonable to expect them to foresee the expressive-stifling effects of boycotting. Under ii. the duty of expressive toleration holds that the fan is morally required to act from reasons which take account of the generalised effects of their responses to the expression. It's important to note here, though, that the duty of expressive toleration also takes account of the interests the fan has in supporting the team (or the NFL in general) and conditionalizing their support on the extent to which those interests are served. Suppose the fan has such distaste for Kaepernick's expression that they simply lose the appetite for watching him play anymore and would lose interest in following the team if he were an active team member. Although this might be considered a kind of unvirtuous intolerance, given the apparent irrelevance of Kaepernick's expression to his appealingness as a football player, it does not necessarily violate the duty of ethical toleration. This is because the fan has a relevant interest in making their support of a team conditional on their enjoyment of watching them play. The nature and strength of the interest plausibly permits them to do so even when it has foreseeable consequences for the freedom to express for others.⁵¹ Consider a radio talk show host whose expressed views increasingly misalign with those of the audience such that listenership drops considerably. Although the departing listeners might be cognisant of the generalised impact of their decision (and its implications as a potential freedom-restricting discouragement to express), their interest in arbitrating for themselves which views they find it enriching to listen to licenses them to make their listening decisions on that basis.

Notice, though, that neither the fan whose enjoyment of football is spoiled, nor the listener who loses interest in the radio show meaningfully *boycott* the expression. Boycotting connotes a more targeted objection to the expression – an intentional withdrawing of support specifically for the reason that something is wrong. Moreover, boycotting implies something beyond collateral loss in

⁵¹ This is consistent with allowing that there is a moral failing in the person and that there are other duties they have in respect of that.

popularity but involves a more concerted effort to refrain from support.⁵² When the fan or the listener makes the decision to cease to support or listen on the basis of such interests of their own, rather than on the basis of a boycotting of expression they deem objectionable, the generalised impact of their decision, suitably described, is also less likely, in general, to risk expression-stifling cost. To put it very roughly, they are following the maxim ‘watch football you enjoy’ or ‘listen to radio you find enriching’, rather than the maxim ‘watch/listen only the football/radio which is not permitting objectionable expression’. The former maxims lack a necessary connection to the expression in question making it more contingent whether following that course will manifest itself in collective behaviour which stifles it.

Given that boycotts are defined, in part, by behaviour motivated by an objection to the boycotted target, as opposed to behaviour engaged in for other reasons, the duty of expressive toleration thereby tilts against boycotts undergirded by an objection to something expressed by (or expression protected by) the boycotted entity. That is not to say as individuals we don’t have interests, including expressive ones, that are themselves served by participation in boycotts. Boycotting can straightforwardly reflect an expression of one’s own identity, for instance as one who disapproves of another person’s expression.⁵³ However, the mere fact that such an interest exists isn’t enough to unseat the idea that we have may have sometimes a duty to refrain from serving it in ways that run counter to the expressive interests of others.⁵⁴ Indeed, this is why it is appropriate to label this a duty of expressive *toleration* – it involves a requirement to bear some

⁵² Note that a person can boycott an entity that they did not, in any case, previously support, if they embrace a commitment to refraining from support such that would pre-emptively exclude other considerations that may have led them to support. It is important to register, that is, that the full social costs to expression go beyond the immediately obvious *inflicted* costs (a barrage of insults) and the *withdrawn* benefits (loss of employment) but also the future opportunities a person is deprived of (their employability; the possibility of association with others)

⁵³ What’s more, persons may also have an expressive interest in ceasing to associate with others in order to *refrain from expressing approval* of something incongruent with their own beliefs. In relation to boycotts this might obtain where it becomes socially understood, in the relevant sense, that failing to participate in the boycott expresses endorsement for the boycotted entity.

⁵⁴ This is very different from claiming that the duty of expressive toleration requires a *general* sacrifice of expressive interests in order to protect others.

reasonable sacrifice in order to avoid complicity in a coercive imposition on others' freedom to express.⁵⁵

From the perspective of the NFL teams who refrain from signing Kaepernick (or, indeed, any other player whose speech is deemed socially unacceptable), the burdens associated with the duty of expressive toleration are potentially greater; including, for instance, the loss of revenue from fan ticket and merchandise sales, and sponsorship consequential on any possible fan boycott resulting from their signing him. If it transpires that the damage to such interests is sufficiently great, it may be the case that the duty of expressive toleration does not require teams to refrain from ruling out signing him, given that the burdens may make doing so overdemanding.⁵⁶ Even in this case, though, it is worth noting that the permissibility of any given team refraining from signing him does not extinguish all moral badness from the state of affairs if the burdens exist in virtue of fan boycotts in violation of those fans' duties of expressive toleration.⁵⁷

⁵⁵ The account I am offering admittedly leaves open the question of what counts as too great a sacrifice. In this way my account offers an outline of the general shape of the duty to refrain from foreseeably imposing expression-stifling costs, without necessarily giving more concrete guidance as to the cases in which that duty obtains.

⁵⁶ Of course, the burdensomeness for any given team does not itself support any *coordinated efforts* on the part of teams to ensure that Kaepernick was not signed (I make no claim as to whether this took place). The justification of such conduct at the bar of the duty of expressive toleration would further depend, for instance, on whether Kaepernick's presence in the NFL as a member of any given team imposed league-wide burdens on even the teams that did not sign him, say by damaging overall league revenues by lessening its popularity, and that, crucially, *such burdens are so great so as to render refraining from participation in such collective action unreasonably demanding*. As with the individual fan case, then, the duty of expressive toleration is still more likely to condemn coordinated attempts to impose costs – through boycotts, for instance.

⁵⁷ Again, parallels between this case and unjust reaction qualifications are illuminating. Suppose racial animus in a community makes some would-be employees less successful in customer-facing roles. It may be overdemanding to require a small business to refrain from taking racially offensive reaction qualifications into account when doing so will predictably lead to collapse of the business (and, suppose, the loss of subsistence income for the business owner). It is consistent to hold that it is permissible for the business owner to take into account racially offensive reaction qualifications and at the same time to hold that this involves an unjust deprivation of equal opportunity, in virtue of injustice pertaining to the reaction qualification themselves. For the same conclusion about such a case see Mason, 'Appearance, Discrimination', pp. 67 – 68.

Under iii. the duty of expressive toleration depends on a prior account of the value of freedom of expression and, more specifically, how far other values justify restrictions on freedom of expression. Suppose that instead of conscientiously refraining from standing for the anthem Kaepernick was screaming racial epithets against teammates. If, because of the harm it causes, people's freedom to express in this way *should* be restricted, then individual may lack the duty to refrain from participating in a boycott manifesting expression-stifling costs.⁵⁸ This indicates that boycotting is not necessarily incongruent with the duty of expressive toleration, but its congruence depends importantly on the broader question of whether the freedom restriction is a justified one. If on the other hand, for instance, Kaepernick, or any other NFL player for that matter, made racially sensitive public comments contending that genetic factors explain intelligence disparities between different racialised groups, then the duty of expressive toleration applies in just the same way as it does in respect of claims, like Kaepernick's, concerning racial injustice. Insofar as people should not be restricted in the expression of such views, the duty of expressive toleration holds that we have a duty to refrain from reasonably foreseeably and reasonably avoidably contributing to the manifestation of costs liable to stifle that expression.⁵⁹

The duty of expressive toleration helps us make sense of Mill's own attempt at navigating a distinction between oppressive moral coercion on the one hand and innocent, educative social discipline on the other. Mill says about a person for whom we have distaste that though "*we are not bound to seek his society [and] we have a right to avoid it,*" we ought not to "*parade the avoidance,*" and that though a person who is judged at fault might legitimately "*suffer very severe penalties at the hands of*

⁵⁸ There would still be a question, of course, of whether participating in a collective imposition of cost is unjust for other reasons.

⁵⁹ A very different, but still illustrative, case is that of the NBA player Jonathan Isaac who, in something of an inversion of the Kaepernick case, *declined* to kneel during the playing of the national anthem before a game, in a context where all other players involved in that game, and in all other games since the league had restarted following a hiatus, knelt in symbolic support for the Black Lives Matter movement. See Nick Friedell, 'Orlando's Jonathan Isaac first in NBA bubble to stand during anthem', https://www.espn.co.uk/nba/story/_/id/29573762/orlando-jonathan-isaac-first-nba-bubble-stand-anthem. Just as in the Kaepernick case, the duty of expressive toleration commends refraining from imposing costs liable to stifle expression of the kind that Isaac engaged in.

*others... [he should] suffer those penalties only in so far as they are the natural, and, as it were, spontaneous consequences of the faults themselves, not because that are purposely inflicted on him for the sake of punishment.”*⁶⁰

Whilst Mill offers these passages as a guide to when social coercion restricts freedom, rather than as an account of the civic duties individuals have to refrain from being party to it, as I do, the thrust of his distinction between ‘natural’ and ‘spontaneous’ consequences rather than those that are ‘purposely inflicted’ can be illuminated by the duty of expressive toleration. Mill’s distinction trades on comparing the significance of the interests we have in not seeking others’ society and avoiding others’ company in virtue of the interests we have in forging our own existence in the social world, and the lesser interest at stake in the purposeful infliction of social punishment. So when says that “*we may stand aloof from a person as well as from a thing that displeases us; but we shall not therefore be called on to make his life uncomfortable,*”⁶¹ it could be that he is saying that though we all have an interest in avoiding the displeasing in life, the same interest isn’t at stake in the proactive engagement in infliction of misery on others to whom we object. Because of this, and in view of the significance of the expressive costs on others, we have a duty to refrain from behaviour that ‘punishes’ in this way.

The second, and related, duty in this context is what I will call the *duty of respect for expressive agency*. Such a duty is particularly pertinent to the forms of online verbal abuse and ‘public shaming’ noted in the above. To more fully explain the character of that duty let me begin with a distinction between shame and guilt. ‘Shame’, as I understand it here, is, broadly, an unpleasant or uneasy self-reflection prompted by cognisance of one’s transgression of a moral or social norm.⁶² Whilst it is commonplace to distinguish shame from *guilt*, there are different conclusions as to how to draw that distinction. One way of trying to draw it is to see guilt as a reflection on a specific undesirable aspect of oneself – perhaps a particular wrong or harm – and, in contrast, to see shame as a more

⁶⁰ Mill, *On Liberty* p 278, my emphasis.

⁶¹ Ibid. p 279.

⁶² This is close to the conception of shame offered by Klonick in her ‘Re-Shaming’, at p 1033.

totalising negative reflection on the self as a whole.⁶³ Though there is something to this, I suspect that it still makes some sense to feel shame with respect to a particular aspect of oneself (even if part of what's implicated here is a lessened sense of self-worth as a whole). I would emphasise an alternative difference. That is that although both shame and guilt imply awareness of a negative judgment on the self, the relevant perspective from which the negative judgment is rendered is *external* in the case of shame and *internal* in the case of guilt. When I feel guilty, it is against a standard of judgment that *I* take to be authoritative (and *because I* take it to be authoritative) that I feel guilt. When I feel shame, it is against the spectre of standards of judgment of others.⁶⁴ To suggest to someone that they ought to feel guilt, as opposed to shame, is therefore to commend *their own appreciation* of the standards by which they have putatively failed, whereas shame has no such implication. In that sense shame, as opposed to guilt, is very much about the judgments of others. Stigma, as I will understand it, has an internal relation to shame and involves, as Nussbaum puts it, the “[effacing of] the human individuality of the other, whether by a literal brand or simply by classifying the person as a member of a shamed class, rather than an individual person. By classifying a person as “a cripple,” “a mongoloid idiot,” “a homosexual,” we deny both the humanity we share with the person and the person’s individuality.”⁶⁵ In that way it is a matter of “casting shame outwards.”⁶⁶

Where the duty of expressive toleration is concerned with the expression-stifling *consequences* of conduct, the duty of respect for expressive agency is a more immediate matter of *respectful treatment*. To better understand the latter, I propose a further distinction between stigmatization on the one hand, which involves the *thrusting of shame*, and what might be termed ‘moral criticism’ on the other, which involves an *invitation* to feel *guilt*. As I explained, guilt is felt as a failure against the backdrop

⁶³ See June Price Tangney and Ronda L. Dearing, *Shame and Guilt* (Guildford Press, 2002) p 25. See also Martha Nussbaum, *Hiding from Humanity: Disgust, Shame, and the Law*, (Princeton University Press, 2004), at p 11.

⁶⁴ Note that this is different from saying that the judgments originate from factors outside the self. It also different from saying that there are *actual* others who render the judgments; the others could be imagined or hypothetical.

⁶⁵ Nussbaum, *Hiding from Humanity*, pp. 39 – 40.

⁶⁶ *Ibid.* p 38.

of normative standards of which one recognises the authority, whereas shame involves a sense of failing against the standards (normative or non-normative) of others. For shame, it matters not whether the shamed recognise the authority, force or value of the ideal to which they putatively fail to measure up.⁶⁷ In contrast, those who feel guilt need to internalise and affirm the basis for their imperfections. With this in view, I think it's sensible to see moral criticism, as opposed to stigmatization, as engaging the agent as someone capable of understanding and appreciating their own would-be failings. It is not necessarily *reason-giving* but the reproach is presented in a way that in some sense respects the other agent as one who can, in principle, agree with the assessment. For shame, all of that is irrelevant. I suspect this is partly what explains why it is appealing to understand shame in terms of the wholesale inadequacy of the person, rather than a specific feature. It is because, in contrast to the invitation to feel guilt, the shamed see themselves, and are seen by others, as the *object* of failure, rather than the subject capable of understanding that (and even why) they have failed.⁶⁸ 'The Look' in Sartre's *Being and Nothingness* is a telling illustration of this.⁶⁹ Sartre describes an instance of a person looking through a keyhole, lost in the process of observing others, when they become aware of someone behind them, observing them. The shame that the person feels comes into being by way of the person feeling themselves as the object of another's gaze. It is the other's judgment, and our consciousness of it, in which our shame inheres. We are objectified in the mind of the other.⁷⁰ With *guilt* induced by moral criticism, on the other hand, the indictment to which we are subject isn't wholesale because it is a communication to us as an agent who is, at the very least, capable of the virtue of recognising their own moral failings. We are not entirely irredeemable precisely because we can see and respond.

⁶⁷ Although the shamed may need to respect, or in some sense care about, the view of the other, against which they feel shame

⁶⁸ For an account of objectification and its disaggregated incidents see Martha Nussbaum, 'Objectification', *Philosophy and Public Affairs* 24 (1995), pp 249 – 291.

⁶⁹ Jean-Paul Sartre, *Being and Nothingness*, (Routledge, 2003), at pp. 282 – 283.

⁷⁰ As Sartre puts it, "Shame is shame *of oneself before the Other*," p 246. See also Bernard Williams, *Shame and Necessity* (University of California Press, 1993)

Stigmatization is therefore an anathema to the duty of respect for expressive agency. Not only is it doubtful whether there is a weighty expressive interest involved in stigmatizing others, stigmatizing involves the thrusting of shame in a way that brands a person as worthy of contempt. The aim of stigmatization is thereby to direct the eyes of the *audience* and to represent its target as an appropriate object of scorn and repudiation. Although the agency of the object of stigma is present by way of their expression making them liable to the stigma, the stigmatization reflects a lack of regard for the agency of the stigmatized, who is represented as the *object* of others' negative gaze. Where we confront expression to which we in some way object, our duty of respect for expressive agency thereby commends moral criticism in place of stigmatization. Insofar as moral criticism invites guilt, it recognises as its appropriate object an *agent* who is capable of recognising the authority of the standards against which it is judged that they ought to feel that they have failed. It is expression whose intended audience *includes* the criticised person and, in that way, is more consistent with a respect for their own agency.

The duty of respect for expressive agency therefore requires framing the response to another person's expression in terms that the expressor can, in principle, affirm in a way that is consistent with their having an authentic sense of self-respect. Now, it's worth emphasising that any concrete application of such a duty is going to be highly context-sensitive – the very same words uttered on a different occasion can have profoundly different meaning. Still, let me give an illustration of what this requires of our discourse with or about others. In the case of Sacco, for instance, the duty is clearly inconsistent with many of the insults hurled at her, such as the one labelling Sacco a 'stupid white b*tch'. To characterise Sacco as 'stupid' is to devalue her general capacity to think such that a self-understanding internalising this conclusion would be incongruent with a sense of self-respect. To frame the insult in terms of her whiteness is to ascribe negative moral valence to her

race in a way that encourages an objectification of her failings.⁷¹ To call her a ‘b*tch’ is to asperse her character in a way that brands her as an appropriate object of collective contempt. Contrast this with a reproach claiming that Sacco’s comments are ‘unthinking and inconsiderate, as well as being hurtful and disrespectful’. To see that such a reproach constitutes moral criticism, rather than stigmatisation, notice that Sacco is in a position to internalise those same conclusions and retain self-respect. She can agree, and verbally affirm, that ‘my words were unthinking, inconsiderate, hurtful and disrespectful’ with her sense of self-worth in-tact. Conversely, were she to state ‘I am a stupid, white b*tch’ it would suggest an unhealthy form of self-loathing.

The difference between moral criticism and stigmatisation doesn’t rest on a contrast between the use of moderate, mild language and harsh and severe words on the other. Moral criticism can involve strong condemnation – ‘that was an abhorrent thing to do’ – and stigmatisation can involve relatively mild insults – ‘you idiot’. Alternatively, it might seem that moral criticism and stigmatisation are divided on the basis that the former targets to conduct and the latter targets the person. This is closer to the mark but fails to capture the important normative difference. Suppose it was charged of Lindsey Stone, who mocked the cemetery signs calling for ‘silence and respect’, that ‘her behaviour spoke of an ignorant mindset’. The criticism of the conduct itself constitutes an indictment of something about her character. Even so, the comment is framed such that Stone herself could, on reflection, and whilst retaining a sense of self-worth, adopt that same perspective. Suppose, instead, that someone calls the behaviour ‘imbecilic’. The target of the criticism remains the conduct, but its implication is to convict her of a failing (being an imbecile) which she couldn’t with self-respect share the judgment about.

⁷¹ This is not to deny that it is possible to conceptualise ‘whiteness’ in a way that refers to status, perspective or behaviour. How far invocation of race involves stigmatisation rather than criticism will importantly depend on shared understanding of the meaning of ‘race’ in what is expressed.

There may be a worry here that my account necessarily classifies criticism of deeply held beliefs as stigmatisation. It might be thought, that is, that my account suggests that if the Atheist says to the Christian ‘your belief that God exists is wrong’ that the self-respecting *Christian* could not in principle share the same judgment. The general form of the worry is that the duty of respect for expressive agency seems to prohibit any comments that are critical of another’s behaviour or belief where they are constitutive of that person’s identity. Such a duty would be quite implausible.⁷² To clarify, then, when I say that we have a duty to frame our responses to others such that they could *in principle* share in the judgement, this is intended to include instances in which the person to whom, or about whom, we are expressing could share that judgment only when they cease to affirm certain beliefs, about the world or about themselves, *when their own reflections on whether to affirm the beliefs is consistent with an internalised conception of their own self-worth*. In other words, it would be consistent with their sense of self-worth to critically reflect on their own commitments by putting them into question. So, whilst not denying that commitments we have are constitutive of our identity, and that respect for identity, and its expression, may be profoundly important in all sorts of ways, such respect does not involve a general prohibition on criticising the deeply held beliefs of others. Indeed, sometimes criticising others’ beliefs itself embodies respect for a persons’ capacity as an agent capable of critically reflecting on their own views such that, in principle, they could share the judgment of the one who criticises.

Now a critic might offer a concern in the other direction – that the duty of respect for expressive agency is too permissive of profoundly disrespectful slurs against identity. Suppose in response to a Christian blog post someone publishes the comment ‘Christians are pigs!’ My account seems to suggest that given that the Christian could in principle share that judgment, subject to undergoing a (reverse) damascene conversion, there is no transgression of the duty of respect for expressive agency. Contrary to this impression, the duty of respect for expressive agency, as I conceive it,

⁷² One need only imagine a proudly self-defined Nazi to see this.

is inconsistent with such identity slurs. It is inconsistent with them because even if a person subjected to slurs of this kind could in principle divorce themselves from the commitment – Christianity – that is associated with the degradation, and so untether themselves from the aspersion on that identity, the Christian could not critically reflect on their commitment – one that they currently hold at the point of critical reflection – whilst retaining an authentic self-respect (countenancing that they are a ‘pig’). In contrast, although the Christian whose most fundamental belief is challenged by the sceptical atheist can’t share the judgment that God does not exist, it is consistent with their sense of self-worth, as an agent whose worth is not exhausted by their commitments, however constitutive, to put into question whether this is a judgment that they ought to share.⁷³

It is worth adding that it is too much to expect of the duty of respect for expressive agency that it could provide anything approaching an instruction manual, rather than a relatively abstract normative guide. Language and meaning are contextual and so, too, therefore are the normative requirements pertaining to their use. Nor do I propose that these duties exhaust what it takes to engage in virtuous discursive activity and exchange (such as might include shared reason-giving, efforts at mutual understanding, sympathetic rendering of one’s opponent in argument etc.). Rather, what I offer here are defeasible moral constraints on what Mill referred to as “*the real morality of public discussion*.”⁷⁴

IV. CONCLUSION

⁷³ This is true even when the identity that is targeted by the slur is an abhorrent one. If someone posts a message ‘all Nazis are scum’ this also transgresses the duty because the Nazi could not critically reflect on their Nazi commitments and countenance sharing the judgment whilst retaining an undamaged sense of self-worth.

⁷⁴ Ibid. p 259.

I have argued that social coercion – whether in the form of shaming penalties, or through loss of employment and association – restricts freedom of expression. In closing, I briefly address two kinds of reservation about the line of argument I have pursued; one contesting the claim that costs imposed from societal sources, as opposed to state institutions, can be freedom-restricting, the other contending that, contrary to the picture I have offered, social coercion is in fact an important and productive force.

In response to all that I have said about the *freedom-restricting* power of social coercion, it might be argued that the phenomena to which I have pointed lack the requisite character to be freedom-restricting. To return to the parallel I drew at the outset between the use of restraining stock shaming and contemporary forms of public shaming, it might be argued that there are important differences between the two cases that cast doubt on whether the latter is freedom-restricting in the sense I suggest. For one thing the former issues from a specific and intentional policy whose design is to secure compliance in a way that the latter characteristically does not. It could be said, then, that this kind of *intentionality* present in the stocks public shaming is relevant for its freedom-restricting character, and that, shorn of this intentional dimension, shaming imposed through a dispersed set of negative responses from others isn't *freedom-restricting* in the relevant sense. Of course, this does not cast doubt on the freedom-restricting character of employment loss – where there is standardly a specific employer making an intentional decision to impose some cost, even if the pressures through which such a decision is incentivised have a more disparate social origin. However, it may apply to cases of employability (such as Kaepernick's), where it is collective but not necessarily coordinated behaviour that manifests the cost in question, and it certainly applies to many instances of public shaming, where the aggregative dimension of the aspersive response is a crucial part of makes it so fearful.

This move, of course, echoes a familiar fault-line in debates over what should count as the appropriate conception of constraints on *freedom* – and the importance, or otherwise, of the

intentional character of the constraint.⁷⁵ I can't hope here to offer a satisfactory defence of a move in that debate. Still, let me cast down on the significance of intentionality here. The invocation of intentionality as a prerequisite for something counting as a constraint on freedom seems to trade, in part, on our sense that restrictions on *freedom* must have a social source – they cannot involve merely natural inabilities. Clearly, though, this cannot be a basis for rejecting these forms of social coercion as freedom restricting, given that the would-be freedom restrictions in question are *constituted by* social phenomena. Alternatively, it might be claimed that *genuine* freedom restrictions generally involve directed intentional agency because restrictions on *freedom* imply an agent who can be held *accountable for* the restrictions.⁷⁶ In that sense the very language of freedom involves an implicit call on some particular agent to justify the existence of a would-be restriction. Although this is a plausible thought, the foregoing has precisely argued that social coercion is something that persons can be accountable for, even if the actions for which people are accountable are not alone sufficient to manifest the freedom-restricting cost.⁷⁷

It's also worth noting that existing discourse surrounding the permissibility of state regulation of expression ought, in fact, to be sympathetic to the notion that social coercion can restrict free expression. One of the most widely cited reasons against some forms of state and institutional regulation of expression is the potential that regulation has to 'chill' freedom of expression. Expression is said to be 'chilled' by such regulation in instances where expression that isn't strictly subject to the regulative measures is *de facto* discouraged, often because of mistaken beliefs, or uncertainty, about which speech is liable to penalty.⁷⁸ However, to appeal to the chilling effect of regulation is itself to go beyond the notion that institutional regulation alone restricts speech.

⁷⁵ For a parallel discussion see Serena Olsaretti, 'Rescuing Justice and Equality from Libertarianism', *Economics and Philosophy*, 29 (2013), pp. 43 – 63, at pp. 51 – 53.

⁷⁶ David Miller, 'Constraints on Freedom', *Ethics* 94 (1983), pp. 66 – 86.

⁷⁷ For related discussion, and similar conclusions, see Jeremy Waldron, 'Homelessness and the Issue of Freedom', *UCLA Law Review* 39 (1991), pp. 295 – 324, at pp. 315 – 318.

⁷⁸ See, 'The Chilling Effect in Constitutional Law', *Columbia Law Review* 68 (1969), pp. 808 – 842.

Although the chilling effect may be the result of the regulation, it isn't merely in virtue of being a consequence of speech regulation that it is expression-damaging, but *simply because of the fact of discouraged expression*. Inasmuch as the chilling effect is accepted as a reason why regulating speech has expression damaging potential, this implies an acceptance that the chilling is *itself* expression-damaging and, in principle, may be inflicted by things other than state or institutional regulation.

Finally, even if it is conceded that social coercion restricts freedom, it might be argued that, far from being a lamentable expression-stifling set of phenomena, social coercion is in fact a healthy and productive force for social good. If anything, the argument goes, social coercion has an important place in a civilised society as a way of disciplining persons towards more socially or morally acceptable behaviour. What's more, to take another lesson from Mill, vigorous expressive exchanges are of crucial importance in fostering self-understanding and giving persons a live and vivid apprehension of the views that they hold.⁷⁹ There are several things to say here. I think it is quite right that our critical responses to others have a key role to play in society in steering, advising and, in a manner of speaking, educating. But it matters here whether the norms and principles to which people are guided are themselves justified ones. What's more, the manner in which people are so disciplined itself needs justification. There is a world of difference between discouraging the expression of a view by way of intimidation and overbearingness as against uncomfortably received but reasoned criticism. Parallels between legal and social coercion are again worth attending to here. There are straightforward and good reasons for legal constraints and a penal system that stands behind them, including, of course, their facility for condemning the morally unacceptable and incentivising good conduct. But this doesn't in any way suggest that there is no danger of their being put to misuse, still less that they shouldn't be considered freedom-restricting. The same is true with respect to social coercion – there may indeed be justifications for restrictions on

⁷⁹ Mill, *On Liberty* p 258.

expression manifested by social coercion, but part of what it is to identify some of the phenomena under discussion as *freedom-restricting* is to say that they *stand in need of such justification*.

Still, it might be argued, social coercion is valuable as a way of disciplining deviant expression for the simple reason that the world would be better without such expression, but we have special reasons to avoid licensing the state to do so. Civil society therefore has an important role in filling the void that the state is ill-suited to occupy. I can't hope to address the many important and sophisticated arguments against state regulation that would help support such a conclusion. It is nevertheless worth mentioning one of the most important such arguments that, tellingly, fails to support such a view. One such popular argument against state regulation of expression has its roots in suspicions about state misuse of power. To paraphrase such a line of thinking, granting power to the state to censor expression gives a license to the state to control expression and thought that will be liable to abuse; supplying a mechanism which it will be tempting to harness for its own political ends. The result is a dangerous expansion of state control in all sorts of ways. Indeed, Mill himself emphasises this worry, when he says that the “*most cogent reason for restricting the interference of government, is the great evil of adding unnecessarily to its power. Every function superadded to those already exercised by the government, causes its influence over hopes and fears to be more widely diffused.*”⁸⁰ I have sympathy with this, now quite familiar, thought, and I quite accept that there are such associated dangers concerning licensing state regulation of expression. However, it is far from clear that these worries pertain to the state specifically and do not also relate to society in similar ways. State regulation of speech often takes the form of tightly legislated statutes that target specific types of speech. To be sure, the categories of speech regulated under law, such as ‘incitement to racial hatred’ or ‘fighting words’, are themselves open to interpretation and contestation, and one might have the worry that the contours of state defined categories may begin to expand to cover more

⁸⁰ Mill, *On Liberty* p 306.

and more forms of expression.⁸¹ However, contestation over what the categories cover and which speech falls foul of them is often pursued through the courtroom, with open, public scrutiny and the need for explicit reasoning justifying classification and restriction of speech. In this way individuals whose speech is regulated have a direct and clear basis on which to publicly challenge the permissibility of such regulation and call for its justification.⁸² Contrast this with the way in which social coercion can limit expression through non-legal mechanisms and I believe worries about the spread of influence and potentially nefarious concentration of power seem similarly salient. Social norms and attitudes that have a role in regulating expression are often changeable, not always clear in advance; they are intangible and sometimes confused. It's far from clear we should be confident in society's aptitude for sensibly judging which speech ought to be disciplined through social cost. Moreover, the freedom-limiting power of moral coercion is often diffused and more difficult to identify but not for that reason less powerful and potentially pervasive. Here, Mill's anxiety over the 'formidable' character of social tyranny and its leaving "*fewer means of escape [and] penetrating much more deeply into the details of life*"⁸³ is particularly prescient. Similarly, as regards the chilling of speech, if it is accepted that social coercion itself imposes restrictions on freedom of expression, then lesser clarity and publicity involved in socially coercive restrictions are likely to have a still more significant chilling effect.⁸⁴ If the penalties are obvious but the which expression is deemed transgressive is obscure, then we won't know what not to say but only to be sure not to say it. If anything, then, we might be even more wary of societal suppression of expression that may veer towards an oppressive culture of conformity and against which individuals may lack a clear public standard by which to hold the restriction to account.

⁸¹ For an account which suggests examples of such a danger see James Weinstein, 'Hate Speech Bans, Democracy, and Political Legitimacy', *Constitutional Comment* 32 (2017), pp. 527 – 583.

⁸² That said, I don't suppose that bringing such claims comes without financial costs; sometimes prohibitive ones.

⁸³ Mill, *On Liberty* p 20.

⁸⁴ Here I have in mind speech which will not in fact receive societal censure but it is erroneously thought that it will.

Acknowledgements:

Versions of this paper were presented at a MANCEPT, University of Manchester workshop panel and a Centre for the Study of Social Justice workshop, University of Oxford. I'm grateful to the participants for the many helpful comments. I would also like to thank Zofia Stemplowska for offering some perceptive guidance, as well as two anonymous reviewers for the journal, whose insightful comments provoked substantial improvements to the paper.

Works cited:

Altman, Andrew, 'Liberalism and Campus Hate Speech: A Philosophical Examination', *Ethics* 103 (1993), pp. 302 – 317

Brisson, Susan J., 'The Autonomy Defence of Free Speech', *Ethics* 108 (1998), pp. 312 – 339

Brown, Alexander, 'What is Hate Speech? Part 1: The Myth of Hate', *Law and Philosophy* 36 (2017), pp. 419 – 468

Brown, Alexander, *Hate Speech Law: A Philosophical Examination* (Routledge, 2015)

Chander, Anupam, 'Youthful Indiscretion in an Internet Age', in Levmore and Nussbaum, *The Offensive Internet*, pp. 124 – 139

Citron, Danielle Keats, 'Mainstreaming Privacy Torts', *California Law Review* 98 (2010), pp. 1805 – 1852

Cintron Danielle Keats, 'Civil Rights in Our Information Age', in Levmore and Nussbaum, *The Offensive Internet*, pp 31 – 49

Citron, Danielle Keats, *Hate Crimes in Cyberspace* (Harvard University Press, 2014)

Citron, Danielle Keats, 'Restricting Speech to Protect it', in *Free Speech in the Digital Age*, Susan J. Brison and Katherine Gelber (eds.), (Oxford University Press, 2019), pp. 122 – 136

Cohen, Gerald. A., 'On The Currency of Egalitarian Justice', *Ethics*, 99 (1989), pp. 906 – 944

Cohen, Joshua, 'Freedom of Expression', *Philosophy and Public Affairs* 22 (1993), pp. 207 – 263

Delgado, Richard and Stefancic, Jean, *Must We Defend Nazis?* (New York University Press, 1997)

de Tocqueville, Alexis, *Democracy in America Volume 1*, Henry Reeve (trans.) Bruce Frohnen (ed.) (Regnery, 2002)

Emerson, Thomas I. 'The Doctrine of Prior Restraint', *Law and Contemporary Problems* 20 (1955), pp. 648 – 671.

Fiss, Owen, *The Irony of Free Speech* (Harvard University Press, 1996)

Franks, Mary Anne, 'Sexual Harassment 2.0', *Maryland Law Review* 71 (2012), pp. 655 – 704

Gallardo, Kristine, 'Taming the Internet Pitchfork Mob: Online Public Shaming, the Viral Media Age, and the Communications Decency Act', *Vand.J. of Entertainment and Tech Law* 19 (2017), pp. 712 – 746

Garton Ash, Timothy, *Freedom of Speech: Ten Principles for a Connected World* (Atlantic Books, 2016)

Hua, Winhkong, 'Cyber mobs, Civil Conspiracy, and Tort Liability', *Fordham Urban Law Journal* 44 (2017), pp. 1217 – 66

Klonick, Kate, 'Re-Shaming the Debate: Social Norms, Shame, and Regulation in an Internet Age', *Maryland Law Review* 75, pp. 1029 – 1065

Langton, Rae, 'Speech Acts and Unspeakable Acts', *Philosophy and Public Affairs* 22 (1993), pp. 293 – 330

Lawrence, Charles, 'Crossburning and the Sound of Silence: Antisubordination Theory and the First Amendment', *Villanova Law Review* 37 (1992), pp. 787 – 804, at p 803.

Leiter, Brian, 'Cleaning Cyber-Cesspools: Google and Free Speech', in Saul Levmore and Martha C. Nussbaum (eds.), *The Offensive Internet* (Harvard University Press, 2012)

Levmore, Saul and Nussbaum, Martha C. (eds.), *The Offensive Internet* (Harvard University Press, 2012)

Lippert-Rasmussen, Kasper *Born Free and Equal? A Philosophical Inquiry into the Nature of Discrimination* (Oxford University Press, 2013), chapter 9 pp. 235 – 260

MacKinnon, Catherine *Feminism Unmodified: Discourses on Life and Law*, (Harvard University Press, 1987)

Mason, Andrew, 'Appearance, Discrimination, and Reaction Qualifications', *The Journal of Political Philosophy* 25 (2017), pp. 48 – 71

Matsuda, Mari, 'Legal Storytelling: Public Response to Racist Speech: Considering the Victim's Story', *Michigan Law Review* 87 (1989), pp. 2329 – 2334

Matsuda, Mari "Who Owns Speech," *Address at Hofstra School of Law*, (Nov. 13 (1991))

Michelman, Frank, 'Conceptions of Democracy in American Constitutional Argument: The Case of Pornography Regulation', *Tennessee Law Review* 56 (1989), pp 291 – 319

Mill, John Stuart., *On Liberty* in *The Collected Works of John Stuart Mill, Volume XVIII – Essays on Politics and Society Part I*, John M. Robson (ed.) (University of Toronto Press, 1977)

Mill, John Stuart., 'De Tocqueville On Democracy in America', in *The Collected Works of John Stuart Mill, Volume XVIII – Essays on Politics and Society Part I*, John M. Robson (ed.) (University of Toronto Press, 1977)

Miller, David, 'Constraints on Freedom', *Ethics* 94 (1983), pp. 66 – 86

Nussbaum, Martha, 'Objectification', *Philosophy and Public Affairs* 24 (1995), pp 249 – 291

Nussbaum, Martha, *Hiding from Humanity: Disgust, Shame, and the Law*, (Princeton University Press, 2004)

Olsaretti, Serena, 'Rescuing Justice and Equality from Libertarianism', *Economics and Philosophy*, 29 (2013), pp. 43 – 63

Bhikhu Parekh, 'The Rushdie Affair: Research Agenda for Political Philosophy', *Political Studies* 38 (1990), pp. 695 – 709

Ronson, Jon, *So You've Been Publicly Shamed* (Picador, 2015)

Sartre, Jean-Paul, *Being and Nothingness*, (Routledge, 2003)

Scanlon, Tim, 'Freedom of Expression and Categories of Expression', *University of Pittsburgh Law Review* (1978 – 79), pp.519 – 550

Seglow, Jonathan, 'Hate Speech, Dignity and Self-Respect', *Ethical Theory and Moral Practice* 19 (2016), pp. 1103 – 1116

Solove, Daniel, *The Future of Reputation: Gossip, Rumor, and Privacy on the Internet* (Yale University Press, 2008)

Solove, Daniel, 'Speech, Privacy, and Reputation on the Internet', in Levmore and Nussbaum, *The Offensive Internet*, pp. 15 – 30

Tangney, June Price and Dearing, Ronda L., *Shame and Guilt* (Guildford Press, 2002)

Waldron, Jeremy, 'Homelessness and the Issue of Freedom', *UCLA Law Review* 39 (1991), pp. 295 – 324

Waldron, Jeremy, 'Mill as a Critic of Culture and Society', in David Bromwich and George Kateb (eds.) *J. S. Mill On Liberty* (Yale University Press, 2003)

Waldron, Jeremy, *The Harm in Hate Speech*, (Harvard University Press, 2014)

Weinstein, James, 'Hate Speech Bans, Democracy, and Political Legitimacy', *Constitutional Comment* 32 (2017), pp. 527 – 583

Wertheimer, Alan, 'Jobs, Qualifications, and Preferences', *Ethics* 94 (1983), pp. 99 – 112

West, Caroline, 'The Free Speech Argument Against Pornography', *Canadian Journal of Philosophy* 33 (2003), pp. 391 – 422

West, Caroline, 'Words that Silence? Freedom of Expression and Racist Hate Speech', in *Speech and Harm: Controversies Over Free Speech*, I. Maitra and M. K. McGowan (eds.), (Oxford University Press, 2012), pp. 222 – 247

Williams, Bernard, *Shame and Necessity* (University of California Press, 1993)