In the wake of the *Charlie Hebdo* attacks in Paris and the chanting of racist slurs by members of the Sigma Alpha Epsilon fraternity at the University of Oklahoma, freedom of expression remains the object of live debate in Western democracies. As David Brink explains, the controversy over hate speech regulation opposes two different strands in the tradition of political liberalism (2001). Egalitarians claim that concerns for equality might justify limits on individual liberties. Those who adhere to such a view believe that the effects of hate speech, as the public expression of discriminatory attitudes, justify such limitations. In the other camp, libertarians hold that freedom of expression is one of our most fundamental liberties and therefore cannot be curtailed. They argue that the cost of regulating hate speech is greater than whatever desirable results doing so might yield. In this paper I take issue with the latter of those two views. *Pace* libertarian philosophers, I argue considerations of individual liberty ground arguments for the legal regulation of hate speech.

The prevalent publication of hate-speech in a society undermines the dignity of its targets, as well as their sense of assurance that others will respect their rights and their personhood. I begin this paper by outlining the kind of speech I am concerned with, and examining its effects as demonstrated in empirical studies. I propose three examples of hate speech, and outline the different agents in these scenarios. This framework allows me to consider the autonomy interests of these different agents in light of libertarian arguments, thereby giving concrete examples in every case. The evidence I consider suggests that hate speech deals psychological damage to its victims by undermining their sense of self-worth.
There have been many defenses of freedom of speech in the philosophical literature. One important line of thought has been the ‘autonomy defense’ of free speech. Those who propose such arguments tend to hold that, of all our liberties, freedom of expression stands out as special due to its intimate link with personal autonomy. In the second part, I argue that considerations of self-governance not only fail to ground a right to free speech, but provide a rationale for suppressing certain types of discourse. Autonomy defenses of free speech fail to ground themselves in satisfactory conceptions of autonomy. I consider one such account, and show how, in light of evidence from section one, hate speech has the effect of undermining the agential autonomy of its target.

A separate class of ‘autonomy’ arguments for freedom of expression considers the interests of autonomous agents. In those arguments, agents value autonomy for certain reasons, and those are undermined by hate speech regulation. In the third part, I consider a recent version of this view, that of Seanna Shiffrin’s (2011). In light of evidence from section one, I argue that regulating (and not tolerating) hate speech best preserves the interests of autonomous agents, described as such. Thus, I conclude that the concept of autonomy lends itself to arguments for regulating hate speech.

1. Hate Speech

Consider three examples of hate speech: (i) a black college student finds herself at a party, and overhears a chant in the room next door, ‘there’ll never be a ****** at ASE, you can hang him from a tree’. (ii) A black child walks down the street and notices a series of adverts that
suggest becoming ‘whiter’ is desirable. Following this, she sees a series of signs outside shops that forbid entrance to African Americans. (iii) A Muslim individual is browsing content on the anonymous social networking site ‘Reddit.’ There, he comes across a trending discussion wherein people of his religion are claimed to be perverts and mass murderers, and those claims prove unchallenged. This type of discourse constitutes a part of those individuals’ environment. Jeremy Waldron describes this phenomenon in saying hate speech can be a ‘world defining’ activity (2012). In its most virulent and uncontrolled forms, hate speech permeates social fabric to the effect of being unavoidable; individual agents cannot help but bear witness.

Commenting on the phenomenon of hate speech, Charles Lawrence explains that “separate incidents of racial stigmatization … are part of a mutually reinforcing and pervasive pattern of stigmatizing actions that cumulate to compose an injurious whole that is greater than the sum of its parts” (1987, 351). This is to say that, in its prevalent forms, hate speech comes to form a general, self-supporting structure, the effects of which trump those of its individual manifestations. As Jeremy Waldron sees it, such structures amount to the establishment of a ‘rival public good’ to the one a fair society aims to establish (2012). The idea is that the absence and suppression of hateful, publicly displayed discourse conveys an assurance to vulnerable minorities that their personhood and rights are recognized and preserves their dignity. When hate speech systemically pollutes the environment of a society, the opposite is achieved. Assurance is delivered to those that seek to undermine a group of individuals based on physical attributes that their view is prevalent and widely accepted, and that minorities are not welcome and will not be
treated as equals. Such is the form of hate speech I am concerned with in this paper – discourse that has the effect of creating such an environment.

Consider three different parties in events of hate speech. First, there are the authors of this speech. Second, there are the receptive listeners of the discourse. Third, there are the bystanders. Such individuals are the people who pursue their daily routines and tacitly accept the reality of hate speech and discrimination that the authors and receivers are all too aware of. They are not a receptive audience to such discourse, yet cannot help but hear it. What follows is a discussion of the autonomy interests of each of these categories of people: the bystanders, the authors, and the listeners.

The effects of living in a society that is permeated with hate speech are well documented by social psychologists, and have been openly recognized by some judiciaries. In the 1990 Canadian Supreme Court case, *R v. Keegstra*, the court ruled against the authors of hate speech and in favor of the appellant, claiming “(t)he derision, hostility and abuse encouraged by hate propaganda … have a severely negative impact on the individual’s sense of self-worth and acceptance” (1990, CR697). There, the court recognized a link between hate speech and decrease in agential self-worth.¹ C.R. Lawrence presents evidence of this in his paper, ‘*The Id, The Ego, and Equal Protection: Reckoning with Unconscious Racism*’ (1987), where he reproaches

---

¹ It is interesting that, in the celebrated 1954 U.S. Supreme Court case ‘*Brown v. Board of Education*’, the NAACP legal team developed an argument based on the research of the sociologist, Kenneth Clark. Clark had shown that African American children in segregated schools suffer from self-rejection at higher rates than those who were not segregated. It is out of the scope of this paper to establish the role of *speech* in this case, but I would like to suggest such authoritative discourse plays a prominent part in this. For discussions of authoritative speech and its effects, see Langton (1993), and Maitra (2009).
American jurisprudence for failing to take contemporary developments in psychology into account. Stigmatizing actions such as hate speech, he writes, “harm the individual in two ways: They inflict psychological injury by assaulting a person’s self-respect and human dignity, and they brand the individual with a sign that signals her inferior status and designates her as an outcast” (1987, 351).

Jurists have not only recognized the damaging effects of hate speech on targets’ sense of personhood, they also have hinted at the notion that self-worth is crucial for personal contribution to society. As Chief Justice Dickson of the Canadian Supreme Court continued in his ruling on *R v. Keegstra* case, lacking a sense of self-worth “may cause target group members to take drastic measures in reaction, perhaps avoiding activities which bring them in contact with non-group members or adopting attitudes and postures directed towards blending in with the majority” (1990, CR687). Lawrence reinforces this, saying that “stigma theory recognizes the importance of both self-esteem and the respect of others for participating in society’s benefits and responsibilities” (1987, 351). In other words, being the victim of hate speech does, in some cases, lead to a decrease in one’s sense of self-worth and a consequent incapacity to participate actively in society. This latter implication points to a double dividend effect of hate speech: by creating a hostile and unwelcoming environment that encourages already marginalized members of society to recede from the active fabric of the social structure, hate speech is in effect further

---

2 Lawrence cites two sources of evidence. Freudian theory, which has now been largely discredited, but also 1980s social psychology. Among the latter, Lawrence draws from the work of Kenneth Clark, whose work on selfhood was influential in the ‘Brown V. Board of Education’ ruling of 1954.
pushing these groups to the outer realms of a society and, as I will argue, reducing their autonomy and self-regulation.

2. An autonomy defense of hate speech regulation

According to T.M. Scanlon, the “doctrine of freedom of expression is generally thought to single out a class of ‘protected acts’ which it holds to be immune from restrictions to which other acts are subject” (1972, 204). This opinion is found throughout the anti-regulation literature, where philosophers often argue that expressive acts are somehow ‘special’ in a way that other harmful acts are not. Thus, much of this literature aims to clearly define the category of protected acts and explain the grounds of their privilege. Of the different methods used by commentators to achieve this result, the ‘autonomy defense’ of free speech has been the most popular (Brison 1998). While separate versions of this argument differ substantively, they all propose that exercising free speech and exercising one’s autonomy are connected in such a way that barring the exercise of the first entails barring the exercise of the second.

As Susan Brison points out, defenders of the autonomy argument tend to appeal directly to the concept of autonomy without fully explaining their conception of it (1998, 324). While

---

some talk of moral autonomy (Dworkin 1985), others talk of ‘rational self-legislation’ (Scanlon 1979), and yet others talk of self-realization (Baker 1978). In her paper ‘The Autonomy Defense of Free Speech’ (1998), Brison identifies six different notions of autonomy that are used in the philosophical literature. Of the six, she argues, five fail as accounts of personal autonomy. The sixth one, which she deems promising, fails to ground a doctrine of freedom of expression. Thus, she concludes, “(i)t may be that hate speech should be protected, but, if so, … such protection must be grounded in a different defense of free speech” (1998, 317).

While there exists a vast literature on the subject of autonomy, one of its sections is under-represented in the debate over freedom of expression – the relational tradition. Relational autonomy grew out of feminist critiques of traditional concepts, such as those employed by some of the defenders of the autonomy argument for free speech. Crucially, philosophers who develop such accounts tend to interest themselves in how social oppression limits personal autonomy (Stoljar 2013). Given this dimension, the under-representation of the relational literature in the debate is regrettable. In what follows, I consider Paul Benson’s relational account of personal autonomy (1994; 2005). Following this, I show how this fully-developed theory grounds a rationale for suppressing certain kinds of discourse. Hate speech, when successful, bars its targets from exercising their autonomy.

Autonomy is typically thought of in terms of ‘self-governance.’ In this way, agential autonomy is thought to be realized when agents act on motives that are, in the relevant manner, ‘their own.’ Since Harry Frankfurt’s ‘Freedom of the Will and the Concept of a Person’ (1971), philosophers have explored how the idea that ownership of one’s action is realized
psychologically. For example, Frankfurt famously proposes that an agent ‘owns’ her action in the relevant sense when she identifies with its motive (1988). Reacting to the literature that stems from this influential proposal, Paul Benson observes that philosophers who pursue Frankfurt’s line of inquiry “typically set about constructing an account of free action by first noting various kinds of impediment or restraint that prevent” it (1994, 650). They then build their theories against the background of those cases, attempting to understand what unites them and abstracting necessary and sufficient conditions for personal autonomy. Those same philosophers neglect one important mode of unfreedom – social oppression. Thus, Benson remarks, their accounts are ill-equipped to explain how oppressive conditions can limit personal autonomy.

One aspect of traditional theories of autonomy is that they impose no constraints on agents’ values, beliefs, plans, etc. (Benson 1994, 653). The problem with this is that agents who appear unfree can satisfy the conditions posited by traditional theories. Illustrating this, Benson considers three different characters. The first is a woman who has been falsely convinced that she suffers from a mental illness. She does not trust her own judgment and willfully does what her husband prescribes. The second is an individual who is gravely ashamed of himself. Third, Benson considers slaves in the pre-1865 United States who do not see themselves as ‘persons’. While those individuals need not lose their capacity to ‘own’ their actions as described by traditional conceptions of personal autonomy, they appear to lack freedom in some relevant way. Thus, traditional conceptions of personal autonomy fail to capture this phenomenon.

4 The term ‘willfully’ requires qualification. Benson points out that such an individual can satisfy the conditions for personal autonomy as they are described by traditional accounts. Recall Frankfurt’s notion of ‘identification’. This individual’s capacity to ‘identify’ is not hampered by her social conditions. And yet, she lacks significant freedom.
The Autonomy Case for Regulating Hate Speech – Louis Ramirez, Keenanprize 2015 - p. 9

The common ground between Benson’s three agents is that they lack ‘self-worth’ in a relevant sense. What unites their psychological condition is that this sense of self-worth is sensitive to others’ attitudes. That is, in all of Benson’s examples, agents have difficulty engaging in relationships with others. The woman who sees herself as insane does not believe she can relate to other people appropriately. The shamed character does not believe he is respectable enough to enter into relationships. The slaves do not see themselves as persons and hence do not believe they can partake in the kinds of relationships that are appropriate to persons. Because what undermines such agents’ freedom is their self-conception, agential autonomy requires them to have a sense of their own self-worth. That is, they need to see themselves as competent to answer for their conduct in the face of the demands of others. Barring this, agents cannot be considered ‘free’ in the relevant sense, as they fail to engage with reasons and motives in the kind of way that makes them ‘their own’.

Recalling section one, the Canadian Supreme Court recognized in the landmark Keegstra Case that hate propaganda led to diminished self-worth in its targets. Moreover, social psychologists suggested hate speech constituted an assault on its target’s self-respect and sense of dignity. Thus, an argument for regulation in the name of autonomy requires equivocation between the ‘self-worth’ that the courts refer to and Benson’s concept.

In light of the U.S. Supreme Court’s accepting an argument that segregating schools damaged black Americans’ sense of personhood in the 1954 case, ‘Brown v. Board of Education’, the writer, Ralph Ellison, claimed “[t]he court has found in favor and recognized our human psychological complexity and citizenship” (quoted in Patterson 2001, xiv). Following this
decision, ”Martin Luther King Jr., often rhapsodized about a ‘New Negro’ emerging out of the black freedom movement who replaced self-pity with self-respect and self-doubt with dignity” (McGuire 2011, 93). It is no coincidence that, in a later paper, Paul Benson uses a character of Ellison’s novel ‘Invisible Man’ for an example of how social conditions and lack of self-worth undermine personal autonomy (2005). This sense of ‘psychological complexity’ that Ellison’s character talks of is precisely what agents need to achieve personal autonomy.

The courts and psychologists may not make use of developed notions in moral psychology, but they do consider the same kinds of cases that Benson does. Indeed, in both cases, those commentators consider the effects of social oppression on an individual’s mental life. In both cases, the same commentators talk of a loss of ‘self-worth’. While the courts do not talk of a subjective sense of normative competence, it is implausible that an agent who qualifies as ‘lacking in self-respect’ according to their notion would satisfy Benson’s self-worth condition. Thus, I suggest (i) that, in recognizing that hate speech undermines its target’s sense of self-worth, the courts imply they fail to satisfy Benson’s condition; (ii) that if this might be otherwise, the burden of proof is on those who argue so.

In a society permeated by hate speech, the victims of this speech may be led to internalize the message of their tormentors. In such cases, individuals fail to fully see themselves as persons and suffer restrictions on the exercise of their agency. Therefore, considerations for individuals’ capacity to self-govern yield the following argument when it comes to freedom of expression:

P1: Hate speech undermines its victims’ relevant sense of self-worth.
P2: Self-worth is a necessary condition for personal autonomy.

C1: Therefore, hate speech undermines the personal autonomy of its targets.

P3: Upholding individual liberty is a desirable goal for legislation in a liberal democracy.

C2: Therefore, restrictions on hate speech are desirable in a liberal democracy.

3. A thinker-based defense of hate speech regulation

In her paper ‘A Thinker-Based Approach to Freedom of Speech’, Seanna Shiffrin writes, “I do not believe that the autonomy case for protecting free speech hinges upon whether we have … the full panoply of executive skills and control over our actions that the broader ideal of self-authorship and self-governance involves” (2011, 9). Rather than our actual capacity for executive self-governance, Shiffrin argues freedom of expression protects a set of interests that are derived from our capacities as free thinkers. The exercise of those capacities, she assumes, constitute the core of what we value about ourselves as agents.5 Thus, Shiffrin aims to establish the connection between valuing ourselves as rational agents, valuing speech, valuing freedom of speech, and regarding speech as deserving of special protection (2011, 3).

Our valuing ourselves as rational agents amounts to recognizing the following interests. While Shiffrin does not claim this list is exhaustive, she argues it captures most of what is relevant to being a rational agent:

5 I take this assumption to be uncontroversial. For further discussion of why this is so, see Shiffrin (2011, 3).
(i) Developing capacities for practical and theoretical thought: Human agents have an interest in developing their capacity to be “aware of and appropriately responsive to the true, the false, and the unknown” in matters both practical and theoretical (Shiffrin 2011, 4).

(ii) Apprehending the true: Human agents are interested in understanding and believing true things about themselves. That is, the contents of their own minds and the environment in which they live.

(iii) Agents have an interest in ‘exercising the imagination’. Such activities, Shiffrin claims, allow agents to develop projections of what the future might be. Furthermore, they also provide an opportunity for philosophical activities (amongst other valuable exercises).

(iv) Agents have an interest in becoming distinct individuals. That is, engaging in a mental life that underpins their distinctiveness as persons.

(v) Rational agents have an interest in being moral agents, acquiring the kind of knowledge and character that are required for complying with the requirements of morality.

(vi) Agents have an interest in ‘responding authentically’. By this, Shiffrin roughly means that agents have an interest in pursuing all the above interests in a manner that is “independent of distortive influences” (2011, 4).

6 I find this passage somewhat unclear. Shiffrin explains it in saying thoughts that are free some such influences “do not follow a trajectory fully or largely scripted by forces external to the person that are distinct from the reasons and other features of the world to which she is responding” (2011, 4). She does not, however, flesh this claim out any more than this, and thus leaves the question as to what exactly might count as such forces entirely open.
(vii) Living among others: Agents have an interest in living among other autonomous agents who share in the capacities described above. Shiffrin points out that satisfying this interest is not a mere matter of companionship. Rather, it is a crucial condition for developing interests i, ii, iv, and v.

(viii) Rational thinkers have an interest in being treated in a morally correct manner by others.

According to Shiffrin, “(s)peech, and free speech in particular, are necessary conditions of the realization of these interests” (2011, 5). Her first argument seems to suggest that communication through speech is a crucial condition of being well known to others. Because this is a necessary condition for the development of complex forms of social life, it is necessary for the satisfaction of the seventh interest: living with others. And because, as was said, this is a pre-condition for satisfying the first six interests, free speech is required to preserve what we most value about ourselves. Moreover, Shiffrin thinks being known by others is a pre-condition for moral agency, the fifth interest. Moral agency, she argues, requires one to know other people’s capacities, knowledge, and perspective, so as to make correct moral judgments (Shiffrin 2011, 5). Finally, communicating the content of one’s mind is a pre-condition for being treated in a morally adequate manner by others. In what follows I call this trio of interests (being with others, moral agency, respect) the social interest in free speech.

Shiffrin also argues that being able to externalize some aspects of one’s mental life is crucial to the first interest – developing good capacities for rational and practical thought. In support of this, she points to the manner in which prisoners in solitary confinement tend to lose their minds:
they “suffer terrible depression, despair and anxiety; moreover, their emotional control and stability wane and their abilities to interact with others atrophy” (2011, 7). Because what seems to cause this is the absence of communication with others, Shiffrin proposes that substantially restricting freedom of expression is on a spectrum with instituting solitary confinement. In what follows, I call avoiding this the self-development interest in free speech. On this view, speech is crucial to securing interests that are grounded in our valued capacity to be rational agents. Thus, Shiffrin concludes, valuing ourselves as thinkers includes valuing free speech and seeing it as deserving of special protection due to its intimate tie to this core capacity.

At this point it is interesting to re-examine the case of the Muslim blogger. This individual finds a platform where the group he belongs to is portrayed as perverse, cruel, and violent. Furthermore, this portrayal reflects a consensus among those who post on this platform. If we apply the ‘thinker-based’ defense of freedom of expression to this case, we find that restricting the online publication of such material undermines the Muslim blogger’s moral agency. Indeed, in order to make the correct decisions, he needs full information about what others are thinking and are capable of. Also, the authors of this material will have difficulty engaging in complex forms of social life, as the content of their minds are prevented from being externalized. In a society where hate speech is prohibited, their development interests are likewise hampered, they cannot discuss their ideas and sharpen their rational capacities. In such a society, the argument goes, the interests of individuals, as rational thinkers, are undermined by government regulation. Therefore, the argument goes, hate speech legislation is hurtful to the autonomy of free thinkers and should be avoided.
In response to this, it is interesting to consider the effects of hate speech alongside the interests of free thinkers as Shiffrin describes them. Rational agents have an interest in engaging in the kind of social interactions that facilitate their development as thinkers (interest vii). While Shiffrin believes this kind of milieu is best developed by blanket protections for freedom of expression, evidence suggests the contrary. In the Keegstra ruling, the Canadian Supreme Court recognized that individuals who are the targets of pervasive hate speech tend to avoid participating in society. The hostile environment that is created does not facilitate the development of free thinkers. Rather, victims refrain from participating in society, thus failing to satisfy interest vii. Also, such agents have an interest in responding authentically. That is, developing their personalities, being moral agents, responding to reasons, in a manner that is independent of ‘distortive influences’ (interest vi). While Shiffrin does not develop what she means by a ‘distortive influence’, the kind of systematic stigmatization that Lawrence calls attention to seems a plausible candidate (1987). In the Keegstra case, the courts recognized such agents suffered loss of self-worth due to the influence of hate speech. Thus, hate speech undermines agents’ capacity to form authentic pictures of themselves.

Determining whether or not the thinker-based defense grounds reason to tolerate hate speech requires balancing the effects described above. If hate speech is tolerated, victims of hate speech risk forming distorted pictures of their own self and refraining from being with others. If hate speech is suppressed, those same victims might lack some information that is necessary for full moral agency – knowledge of the exact opinions of their interlocutors. I would, however, like to call into question the necessity of this knowledge for moral agency. Potential victims of
hate speech are not unaware of the existence of certain views about them – the claim that they lack moral agency seems to require further grounding. Indeed, I find it implausible that being constantly confronted by aggressive messages increases your knowledge of other people’s opinions substantially, let alone your agency. One might very well form a distorted picture of what the general opinion is, for example. It is not in the scope of this paper to fully balance the interests of different parties to hate speech. However, in light of the considerations above, I would like to suggest it is most likely that thinker-interests are best preserved by regulation.7

Conclusion

We have seen two facets of the personal autonomy and its role in legislation. On the one hand, personal autonomy is the capacity of agents to act on reasons and motives that are their own. Considerations for this kind of freedom permeate various commitments in liberal democracies – commitments to ban forced servitude, for example, or the writ of habeas corpus. Through each of those, the state recognizes that depriving an individual of her capacity to act on her own reasons amounts to doing her wrong, and that this should be avoided. Hate speech deprives agents of their capacity to act autonomously by violently modifying their self-perception. Thus, if a state is committed to individual liberty in this first sense, robust measures against this should be implemented.

7 This argument is close to the one David O. Brink develops in his ‘Millian Principles, Freedom of Expression, and Hate Speech’ (2001). There, Brink argues that what is at stake in freedom of expression for Mill is a set of higher capacities. Because those are undermined by hate speech, Millian principles do not necessarily ground a defense of freedom of expression.
Second, autonomy generates a set of interests in agents. Those include an interest in developing fully as human beings, among other things. Yet, here as well, we find that pervasive hate speech bars its victims from achieving this. Agents who are stigmatized by hate speech are barred from flourishing as human beings. Thus, whichever conception of autonomy one is preoccupied with, the concept supports suppressing hate speech and not tolerating it. A society where autonomy is protected is one where certain limits are implemented on what individuals can publish.

As Jeremy Waldron explains, certain countries that do restrict hate speech reflect the belief that regulation is something that might be “required in a human-rights context” (2012, 13). I suggest understanding hate speech as an assault on personal autonomy hints at a theoretical foundation for this claim. While it is outside the scope of this paper to trace the importance of those findings in the broader debate on freedom of expression, I would like to conclude that, if there is a case for protecting hate speech, which I believe is unlikely; it must be one despite considerations of autonomy. Thus, I suggest the notion of a ‘right to freedom of expression’ is made implausible in light of this argument.

Works Cited:


