Uppsala Universitet
Filosofiska Institutionen

A Right to Free Hatred?

Johan Apelgren
C-uppsats
HT-2019
Handledare: Per Algander

1. Introduction

If you live in a liberal democratic society, chances are good that you appreciate and support the right to free expression. Open and free discussion between citizens serves both as a democratic safeguard and as an important part of basic human autonomy. But while one may support freedom of expression in the general case, each society will have some topics which approach or cross the border of general moral acceptance. When it comes to these topics, the value of free speech is often put into question or outright denied, even by those who otherwise support it.

One such subject in contemporary western society is hate speech. That is, the propagation of inflammatory, accusatory, hateful or threatening speech directed not at a specific individual but at a social group (generally a minority). Many countries struggle with the complex issue of how to deal with those who want to publicly express support for neo-Nazism, fascism, racism or other ideologies that denigrate certain societal groups. While we want to protect the right for everyone to speak their mind, past experiences, not least those of World War 2, make it hard to accept such expressions as being simply another opinion in the marketplace of ideas.

So, what should we, as a society, do about it? Ought these expressions simply be banned; accusations of paternalism and suppression of opinion be damned? Or is there a significant upside, whether direct or indirect, to allowing these opinions in the public square? Ronald Dworkin argued that there was. He was a staunch supporter of freedom of speech throughout his life, on several occasions arguing specifically against imposing limits on hate speech. Such limits, in his view, undermines democratic norms.

This position was, of course, not without opposition. One of those Dworkin frequently debated with on this subject was Jeremy Waldron, a former student of his. In Waldron's 2012 book outlining his own theory on this subject, *The Harm in Hate Speech,*[1] Waldron also included a direct response to Dworkin's 'legitimacy

[1] Waldron, Jeremy, *The Harm in Hate Speech*. Harvard University Press, 2012.

argument'.[2] Unfortunately, since Dworkin passed away in early 2013, no comprehensive response to Waldron's criticism was ever presented.
In this essay I will examine Waldron's criticism of Dworkin's argument, as it is presented in chapter 7 of *The Harm in Hate Speech*. My goal is to evaluate whether the 'legitimacy argument' can withstand this criticism, and in turn whether it is reasonable to oppose (certain kinds of) hate speech legislation by referring to this argument.

I begin this work in section 2, where I first make an initial theoretical assumption, before I summarize both Waldron's 'dignity argument' and Dworkin's 'legitimacy argument' in order to provide the reader with the context necessary to fully appreciate Waldron's critique.

In section 3, I will begin the discussion by going through the critique step by step, evaluating the strength of each argument against Dworkin's theory. As a part of this process, I will consider possible replies Dworkin could have employed to defend his theory. These replies will, to the best of my ability, be in line with what I understand Dworkin's position to be. I will also consider possible counter-replies from Waldron, before deciding whether I think that the criticism ultimately succeeds. To conclude the discussion, I will explain why I believe that Waldron shows serious flaws with the 'legitimacy argument', and why I believe that it fails to provide adequate reason for opposing certain types of hate speech restrictions.

In section 4, I will summarize the discussion, my conclusions, and my views on how this relates to contemporary concerns about free speech.

[2] The argument is, to my knowledge, never given a name by Dworkin. I borrow this moniker from Waldron, who refers to it as such in his book.

2. Theoretical context

Before going into detail on the two theories, I want to establish a basic theoretical assumption. I will assume that no credible free speech-theory[3] can support an 'absolute', or 'unlimited' freedom of expression. By that I mean the right to express whatever you want, whenever you want to, however you want to, with no legal threat of interference from any outside actor under any circumstance.

This view seems to be ascribed to freedom of expression advocates quite regularly in the societal discourse, often by freedom of expression supporters who ascribe it to themselves. However, it seems to be a virtually indefensible position, and I have not been able to find any established philosopher who argues for it. As John Stuart Mill put it:

> "All that makes existence valuable to anyone depends on the enforcement of restraints upon the actions of other people. Some rules of conduct, therefore, must be imposed—by law in the first place, and by opinion on many things which are not fit subjects for the operation of law."[4]

There are many more things one could say to argue why such a theory is untenable (not least practical reasons – what if I use my right to free speech to deprive others of their right to free speech, say by shouting down everyone with opposing views?), but I think Mill explains it succinctly enough.

2.1 Waldron's 'dignity' theory

As Waldron's theory is not the focus of this essay, I will only provide the details needed to give some basic context to his critique of Dworkin. The key term in Waldron's theory is 'dignity', which he defines as the "public good of inclusiveness."[5]

---

[3] I use speech and expression interchangeably in this essay, and they should be understood as referring to the same thing. Note that this should be understood as referring not only to *spoken* speech, but also to written or nonverbal communication (e.g. the Nazi salute), as well as things like images, statues, etc. Basically, anything that can reasonably be understood as expressing a view, opinion, feeling, state of mind, etc.

[4] Mill, John Stuart, *On Liberty*. Hackett Publishing, 1978, page 5.

[5] Waldron, 4.

This public good is the result of the acknowledgement by each group in society that "society is not *just* for them; but it *is* for them too, along with all of the others."[6]

In other words, Waldron's 'dignity' has two aspects to it. The first aspect is accepting that your group needs to co-exist with other (often very different) groups in society. The second aspect is knowing that the rest of society accepts your group in the same way as you accept them, namely as members of society in good standing. Both of these aspects are important, assuming that we are committed to a multifaceted and well-functioning society. Waldron argues that we should not allow hateful rhetoric to sow distrust and drive wedges between different societal groups. He considers hate speech a "slow-acting poison, accumulating here and there, word by word, so that eventually it becomes harder and less natural for even the good-hearted members of the society to play their part in maintaining this public good."[7]

This is a rather pragmatic argument. It does not rest on any of the more idealistic assumption that are commonly put forward as reasons for banning hate speech (such as, e.g., that we should treat each other with a minimal level of respect). Instead it argues that disallowing it has instrumental benefits to all of us. Reasonably, we should all want our society to function as well as possible. As it happens, our society consists of many different societal groups, and it is not likely to become homogenous any time soon. Thus, we have a strong imperative to favor the societal conditions that are maximally conducive to good relations (friendliness, intermingling, trust, cooperation etc.) between these different groups. Since a society that allows hate speech is (almost undisputedly) less conducive to such relations than a society that doesn't, we should not allow it.

Of course, society has many other considerations besides good conditions for coexistence. These other considerations could potentially lead us to the conclusion that banning hate speech is, all things considered, a net negative for society. This possibility is what Dworkin's argument wants to highlight.

[6] Waldron, 4.
[7] Waldron, 4.

2.2 Dworkin's 'legitimacy' theory

The legitimacy argument was developed by Dworkin over many years. The first publication of this line of thought that I could find is in an article published in 1994.[8] It was then mentioned again by Dworkin in an article from 2006,[9] on the topic of the 2005 Jyllands-Posten Muhammad cartoons. Both of these sources have helped solidify my understanding of Dworkin's position, but the main source for this essay will be what appears to be the final, and by far most detailed, version of the argument. This version was presented by Dworkin in the foreword to the anthology *Extreme Speech and Democracy* from 2009.[10]

In this foreword, Dworkin makes clear his opinion that free speech should be viewed as a basic human right, and expresses a desire to find a theory that can justify the universal value of freedom of expression even to those who may not support the theoretical underpinnings this right commonly relies on (e.g. individual liberty, equal rights etc.).[11] He thinks that J.S. Mill's argument, by far the most well-known defense of free speech, has limited persuasiveness since it relies on the reader accepting a fallibilistic world-view.[12] In Dworkin's view, this makes the theory too narrow in scope since not everyone accepts such a world-view.

While we may think that fallibilism is generally prudent, the kind of broad freedom of expression that Dworkin advocates will necessarily have to deal with issues regarding religious convictions, moral convictions, or other such strongly held beliefs. When that happens, defending freedom of expression by appealing to the fact that those convictions might be mistaken will generally fall on deaf ears. Dworkin wants to be able to demonstrate the value of free speech even to those who are completely convinced that the speech they are trying to censor is wrong, immoral or dangerous. Free speech, Dworkin argues, should not be defended in terms of its instrumental value, but as a basic principle of life.

---

[8] Dworkin, Ronald, "A new map of censorship". *Index on Censorship,* 1994.
[9] Dworkin, Ronald, "The Right to Ridicule". *New York Review of Books*, 23/3/2006.
[10] Dworkin, Ronald, "Foreword". *Extreme Speech and Democracy,* 2009.
[11] Dworkin, Foreword, vii.
[12] Fallibilism being the epistemological stance which argues for some variation of the basic idea that one's beliefs should always be open to revision in the face of new evidence, no matter how infallible these beliefs may seem at the current moment. See e.g. https://www.iep.utm.edu/fallibil/.

It is in response to this issue that Dworkin proposes a principle that, in his view, illustrates the value of free expression regardless of one's personal attitude towards any specific subject that this right might touch upon:

> *The legitimacy principle:* "[I]t is illegitimate for governments to impose a collective or official decision on dissenting individuals, using the coercive powers of the state, unless that decision has been taken in a manner that respects each individual's status as a free and equal member of the community."[13]

The first thing we can note is that this principle almost certainly *won't* appeal to just anyone. It does not quite go so far as to presume a belief in democracy, but at the very least it presumes an acceptance of the idea that government should be restricted in how it may exercise its power over the populace.[14] Dworkin claims to find the principle a "condition of human dignity," which in itself is far from a universally accepted concept.[15]

However, while this principle may not be convincing to everyone, it will find plenty of support in the Western liberal-democratic societies, which is the context in which Waldron's criticism is levelled. Hence, we will consider the strength of the principle in this context only.

On its face, this principle does seem to have some intuitive strength to it. It certainly does seem in line with democratic values to suggest that every citizen of a democratic country should have a real opportunity to express their opinions on laws that are being considered in their society.

However, this is de facto not a common (or even uncommon) implementation of democracy. Most democratic societies consider the only required process of legitimization to be a simple majority vote, and only a representational vote at that.

---

[13] Dworkin, Foreword, vii.
[14] In fact, the principle as stated is similar to e.g. the "liberal principle of legitimacy" as discussed by Rawls among others, and should probably be seen as part of a wider liberal tradition rather than, as Dworkin argues, a principle unrestricted by ideology.
[15] Dworkin, Foreword, vii.

But Dworkin, while agreeing that a majority decision is a necessary condition for a fair democracy, argues that it is not a sufficient condition:

> "Fair democracy requires [among other things] that each citizen have not just a vote but a voice: a majority decision is not fair unless everyone has had a fair opportunity to express his or her attitudes or opinions or fears or tastes or presuppositions or prejudices or ideals, not just in the hope of influencing others (though that hope is crucially important), but also just to confirm his or her standing as a responsible agent in, rather than a passive victim of, collective action. The majority has no right to impose its will on someone who is forbidden to raise a voice in protest or argument or objection before the decision is taken."[16]

Here we have the core of Dworkin's argument. Passing a law through majority decision (or another appropriate form of governing) is not sufficient to establish legitimacy. A fair democracy requires that its citizens are allowed the means to oppose or support any majority decision considered for implementation. Not simply through voting, but also through opinionating in favor of their position, or simply by contributing to the moral climate of society. If someone has been deprived of their ability to voice their honestly held opinion on a proposed law, then that law, once implemented, would reduce the democratic legitimacy of the state.

In other words, you should be able to say anything you want as long as it relates to a proposal that the state is considering (or might consider in the future) adopting. To take an example, it means that if a government wants to pass a law that makes illegal the intentional discrimination of a minority group, then there should not simultaneously exist legislation that forbids citizens from expressing certain opinions towards that minority group. If such limits exist, then those who oppose the discrimination legislation have not been given the same standing as those who support it. When this happens, Dworkin claims that we "forfeit our moral title to

[16] Dworkin, Foreword, vii.

force such people to bow to the collective judgments that do make their way into the statute books."[17]

This is an issue because Dworkin sees it as appropriate, if not outright required, of a fair and inclusive democracy to adopt legislation that protects "women and homosexuals and members of minority groups from specific and damaging consequences of sexism, intolerance, and racism,"[18] and the loss in legitimacy will affect how credible the state is when trying to enforce such protections. Since legislation protecting against discrimination is of such importance, Dworkin argues that the loss of legitimacy cannot be justified.

Dworkin refers to the act of introducing laws against hate speech as 'intervening upstream'. Waldron adopts this nomenclature in his critique, calling laws restricting speech "upstream legislation" and laws restricting discrimination 'downstream legislation'. The analogy is straightforward: if you intervene upstream, your actions may have an effect downstream, but not vice versa. In this case, the upstream intervention is the introduction of hate speech restrictions, and the downstream consequence is the loss of legitimacy that those restrictions carry. It is the anti-discrimination law, i.e. the downstream law, that suffers this loss of legitimacy, while the upstream law remains unaffected.

To more easily discuss Waldron's criticism, I will be using this nomenclature as well. If we were to re-phrase Dworkin's argument in these terms, we would say that the existence of upstream legislation reduces the legitimacy of any corresponding downstream legislation, where 'corresponding' should be understood as "dealing with the same societal group."

To clarify what it means for two laws to correspond to each other, consider two examples: a downstream law that protects women against workplace discrimination would have reduced legitimacy if there simultaneously exists an upstream law that forbids the expression of misogyny. Similarly, a downstream law that establishes an affirmative action program for certain minority groups would be delegitimized by the

[17] Dworkin, Foreword, viii.
[18] Dworkin, Foreword, viii.

existence of upstream legislation that forbids hate speech towards those same minorities. But the upstream law forbidding misogyny would not delegitimize the affirmative action law, so long as women were not one of the groups benefiting from the affirmative action.

At this point, one could reasonably wonder if Dworkin's worry is not a moot point. Is it not entirely possible to have legislation that forbids the publication of hate speech, while simultaneously allowing people to oppose anti-discrimination legislation in other ways? Does one truly need to be able to vocalize the deepest depths of their hatred for an entire group of people, just to show their dissatisfaction with some piece of legislation?

Dworkin's answer would be that yes, they do need that ability. As I noted previously, his view of legitimacy requires not only the ability to vote or speak out against something, but the ability to contribute to the *general moral climate*. Since laws often (in the case of laws concerning discrimination and hate speech, practically always) have their origin in the morality and culture of the citizenry, not allowing certain opinions in the public square effectively excludes some people from the democratic process. Dworkin puts it this way:

> "It is as unfair to impose a collective decision on someone who has not been allowed to contribute to that moral environment, by expressing his political or social convictions or tastes or prejudices informally, as on someone whose pamphlets against the decision were destroyed by the police. This is true no matter how offensive the majority takes these convictions or tastes or prejudices to be, nor how reasonable its objection is."[19]

As we will see later on, I think this specific passage creates significant issues for Dworkin. But for now, we will move on to discussing Waldron's criticism of the legitimacy theory.

[19] Dworkin, Foreword, viii.

3. Waldron's critique

I will divide Waldron's criticism into two sections, one for each part of his criticism. The first part discusses whether the legitimacy argument is too broad, and the second part concerns whether there is a reasonable interpretation of the term 'legitimacy'.

3.1 The first argument

Waldron's first criticism centers on whether the legitimacy argument can be coherently defended when applied to other contexts where we *do* limit speech, and think that it is the right thing to do.

He mentions several possible ways that people could protest against downstream legislation. For example, they may "shout 'Fuck!' in public, or challenge the legislation's proponents to a fight, or urge mutiny by the armed forces, or display child pornography."[20] He notes that since the above ways of expressing opposition actually *are* prohibited by upstream legislation in most (if not all) nations, it seems that the legitimacy argument would have an issue with these actions as well. If Dworkin's argument is correct, then should these expressions also be legal?

This conclusion, if left unchallenged, would seriously damage the credibility of the legitimacy theory. It seems that either the theory is incorrect, or that we need to reconsider the legality of many actions which most people are convinced should remain illegal.

3.1.1 Dworkinian defense against the first argument

It is possible that Dworkin wouldn't mind (some of) these actions being legal, but I will defend Dworkin's position under the assumption that he would like to reject these conclusions. After presenting my defense of Dworkin, I will present and evaluate two possible replies to the defense, before giving my analysis on whether the defense succeeds.

[20] Waldron, 182.

I believe Dworkin's best recourse would be to attempt to differentiate the type of speech acts mentioned by Waldron from the type of speech acts that Dworkin is trying to defend (i.e. hate speech). Dworkin would, even without Waldron's challenge, likely need to do this at some point anyway, in order to justify banning some kinds of expression, which this essay has already assumed that any reasonable free speech theory must be able to do.

One way to draw this distinction could be to argue that Waldron's speech acts do not seem inherently linked to an argument for disapproval, while hate speech typically is. What I mean by this is that those who are inclined to use hate speech to oppose a law often use the hate speech *itself* as an argument for their opposition.

For example, someone might argue that we should not pass a law allowing same-sex marriage because, e.g. "I believe homosexual acts are an abomination in the eyes of God." In this example, the content of the hate speech *is* the argument. You cannot express the argument without expressing the belief underlying it. At most, you could perhaps water it down (e.g. "I believe God disapproves of homosexuals"), but this would cause the argument to lose some of the force it has for those who might be swayed by it, which would cause precisely the kind of unfairness that the legitimacy argument is concerned with.

With regard to Waldron's examples, I do not need the right to shout expletives in public in order to express my reasons for opposing a new tax bill. I might have a strong urge to do so, but it wouldn't add anything to my argument against passing the bill. Shouting "Fuck!" is not expressing an argument, it is at best expressing exasperation. Likewise, I do not need to urge mutiny by the armed forces in order to protest my government's involvement in an unjust war.[21] The speech act of urging mutiny is not an argument in favor of my position, it is a suggestion or plea for action – if anything, it is an attempt to bypass the democratic process altogether.

In fact, none of Waldron's examples seems like good examples of acts of protest. They could be considered acts of protests in a broad sense, simply because most people

[21] An example of such a protest might be the counter-culture movement protesting (among other things) against the Vietnam war during the 1960s and 70s.

disapprove of these actions, but they don't really *protest* anything specific. Put in terms of the legitimacy argument, there does not seem to be any downstream legislation that is obviously connected to these actions. They seem to be general forms of voicing discontent rather than attempts to convince others to take your position. Thus, banning Waldron's type of speech acts (shouting expletives/urging mutiny/etc.) does not seem to remove the possibility of expressing disapproval based on any specific view. It just seems to limit certain actions generally, which the legitimacy argument doesn't have an issue with.

On the other hand, banning the speech acts that Dworkin is defending (i.e., espousing racist or bigoted views) *does* seem to remove the possibility for expressing disapproval based on a specific view. Namely, a racist or bigoted one. Dworkin's position does not commit him to defending *any* expression just because that expression was used in connection with protesting a law. A threat of physical violence, for example, would not enjoy free speech protection under this interpretation of Dworkin's theory. The threat may be intended to shape public opinion somehow, but that alone is not a necessary condition for expressing any specific view (whether bigoted or otherwise). If anything, a threat of physical violence would be contrary to the idea of free speech, since such a threat could stifle opposing views (the proponents of which also enjoy the same rights to express themselves as the one making the threat does).

3.1.2 Possible replies

There are (at least) two possible objections to the defense I've put forward, where the second objection is a more specific version of the first one. The first objection is to appeal to the principle of free speech that the theory is supposed to protect. That is, to ask why it would matter to Dworkin's theory whether the expression is a part of an argument or not. Dworkin wanted to construct a theory that shows that free speech is a basic human right, independent of other considerations. Shouldn't it then be enough for something to be speech for it to be protected by free speech? Isn't making a distinction between hate speech and other unwanted speech just splitting hairs?

As I discussed at the start of section 2, there is in practice no such thing as an absolute freedom of expression. Thus, all reasonable theories of free expression must

make a distinction between different types of expressions. Dworkin is undoubtedly aware of this fact. If he did indeed intend his theory to treat all speech equally, then his theory is likely doomed from the outset. Both I and Waldron give him more credit than that, and I will assume (at least until proven otherwise) that he had a more nuanced idea in mind.

Therefore, the mere fact that something is an expression cannot be enough to protect it from censorship or rebuke. The circumstances under which an expression should not qualify for protection is not something Dworkin explicitly discusses, but an attempt to formulate a reasonable criterion might sound something like the following: an expression does not qualify for protection if the expression has a clear, obvious and significant downside, while at the same time also lacking any comparable redeeming features. The stereotypical example of such an expression is falsely screaming "Fire!" in a crowded theater. Such an expression serves no purpose except to create panic, distress, and damage, and can therefore be banned to everyone's benefit.

Allowing hate speech has a clear and significant downside. How significant is a topic for another debate, but Dworkin himself argues that the downsides are a cost worth paying.[22] In other words, for Dworkin these expressions *do* have a redeeming feature – specifically, they are justified by the fact that permitting them confers legitimacy onto other laws. Shouting expletives, issuing threats of physical violence, swindling money from unwitting seniors, displaying child pornography etc., are all expressions which have no such comparable upside (or if they do, it is at least not obvious to me).

Expressions, like all other types of actions, must be regulated to some degree. While our general commitment to the freedom of expression puts the bar for banning speech very high and makes clear that the burden of proof rests on those in favor of censorship, that is not the same thing as disallowing censorship altogether. On these grounds, I think it's reasonable to argue that Dworkin would not confer protection on just any expression.

However, the second reply puts the explanation I've just presented into question. I have defended Dworkin by appealing to the stated assumption I've made from the

[22] Dworkin, Foreword, vii.

outset, namely that no reasonable free-speech theory (which Waldron and many others judge the legitimacy theory to be) can support a blanket protection of all speech. Since I assume (at least until contrary evidence has been presented) that Dworkin's theory is reasonable, it follows that I also assume that his theory does not grant such a blanket protection, and that therefore interpretations that come to this conclusion must be mistaken. Dworkin's theory seemed defensible in this narrow context, when it related only to expressing views that oppose specific pieces of legislation.

But Dworkin is - by his own account - concerned not only with each citizen's ability to oppose specific legislation, but also with their ability to influence the *general moral climate* of the society. Thus, the second reply is this: While shouting expletives and urging mutiny might not be arguments against a specific law, aren't they nonetheless expressions that contribute to the general societal atmosphere and moral climate, and wouldn't we therefore incur a loss of legitimacy if they were to be prohibited?

I think this second reply is much more damaging than the first one. Precisely how damaging will depend on how well one manages to argue that these things contribute to the general moral climate. I will not spend much time arguing that here, but prima facie it appears likely that very nearly all expressions contribute to the general moral climate in some way. And even if some expressions could be shown not to contain moral content, it seems evident that the most controversial expressions (like the ones we might be inclined to ban, say child pornography or threats of violence) do. They are, after all, controversial precisely *because* they go against the norms and morality of our societies, and so by performing these expressions you are *de facto* making a moral statement.

Dworkin's express desire to protect any expression that affects the general moral climate, and the realization that this is likely to include virtually any expression imaginable, is in my view a strong reason to consider the legitimacy argument flawed. Assuming that it is indeed possible to make a strong case for why most expressions affect the general moral climate in some fashion, we would have to either drop the provision for protecting such expressions from the theory (which in the process

would greatly reduce the range of expressions it protects), or we would have to abandon the theory altogether.

But, since there is some potential wiggle room with regard to what types of expressions should count as contributing to the general moral climate of society, it is not necessarily a death blow to the argument. And, more importantly for the purposes of this essay, it is a line of thinking that Waldron does not explore.

He focuses his attention on the fact that hate speech is harmful, just like the acts he uses as examples are, and that we therefore cannot make a distinction between hate speech and the other illegal acts on that basis.[23] The issue of scope with regard to the 'general moral climate' clause is left untouched by Waldron. Therefore, I will leave the discussion here and move on to evaluate the rest of Waldron's critique.

3.2 Second argument

In the second part of his criticism, Waldron considers what exactly Dworkin could have meant by the term "legitimacy". Specifically, Waldron wants to figure out what it would mean for a law to *lose* legitimacy? Legitimacy, after all, is the founding block that we rest our democratic society on. It is the perception of legitimacy, of a government's power being willingly bestowed to it by the population, that makes citizens of democracies accept laws in a way that is not the case in less inclusive systems of government. It is one of the primary reasons for why we consider democracy to be the preferable system of government above all other options.

So, if a law was to lose that legitimacy, what would happen? Can that law no longer be legitimately enforced? Are people who would normally be bound to obey it no longer so? Waldron calls this the "binary interpretation" since it leaves only two possibilities open: a law either has full legitimacy, or it has no legitimacy at all.

[23] Waldron, 183.

3.2.1 Loss of legitimacy as a binary feature

If this is the interpretation of "losing legitimacy" that Dworkin intends, then it would lead to some disturbing consequences, because there currently *is* upstream legislation against hate speech in place in almost all liberal democracies.[24] If losing legitimacy means a complete loss of authority, it would mean that, according to the legitimacy theory, all corresponding downstream legislation can be ignored at will.

Hence, people, companies, government departments, universities, and other institutions would be able to discriminate against minorities free from any legal concerns. Granted, they would still have to accept the implementation of anti-hate speech laws, since it is the implementation of those laws that cause the loss of legitimacy to take place. However, it seems to me that even those that propagate hate speech would, if given the choice, much rather have the ability to discriminate against a group than the ability to "only" badmouth them in public.[25]

The silver lining would be that this loss of legitimacy would primarily (or perhaps only) affect anti-discrimination legislation. We could still legitimately prevent e.g. racial violence, since our society has an overarching commitment to protecting its citizens from violence which is not contingent on any racial factor.

But that is a small comfort given the circumstances, and it seems like this is a trade-off that the racists and hatemongers would welcome. Given these consequences, it seems doubtful that such an extreme interpretation is what Dworkin had in mind.

[24] The United States is the main exception to this due to how expansive the First Amendment clause of the US constitution is, explicitly forbidding *any* law restricting speech. However, there have been exceptions to this rule carved out over the centuries, giving further indication that no reasonable system can be completely devoid of restrictions.
[25] Waldron, 185.

3.2.2 Loss of legitimacy, but only for some laws

Having reached the conclusion that the first interpretation results in an unworkable theory, Waldron looks for other interpretations of what losing legitimacy could reasonably mean. The first interpretation that Waldron considers is that loss of legitimacy only affects some downstream laws, but not all. The theoretical basis for such a distinction would be something similar to the comment I made about racial violence in the previous section. Perhaps laws against racial violence could still be enforced even if laws against racial discrimination could not, since we want to prevent violence in general.

Waldron is not convinced by this possibility, and neither am I. First off, the source of the legitimacy (i.e. the possibility of having a free and open debate) is the same for both laws, and thus losing legitimacy for one should mean losing legitimacy for the other. Even if we have a commitment to prevent violence in general, it is not clear how we could retain the legitimacy of laws related to *racial* violence specifically, since the upstream law would affect all laws concerning the topic it restricts (i.e. race, in this example). And finally, even if we could figure out a satisfying solution to these issues, we would still lose all of our current anti-discrimination legislation. In light of that, this interpretation seems nearly as bad as the one we originally rejected.

3.2.3 Loss of legitimacy, but only for some people

Waldron's borrows his next interpretation from Robert Post: "If the state were to forbid the expression of a particular idea, the government would become, with respect to individuals holding that idea, heteronomous and nondemocratic."[26] This would mean that the law only loses legitimacy towards those who are inclined to express hate speech, and not towards all of society.

While this is an interesting take, it is problematic in two regards. Firstly, the people who are inclined to express hate speech are also likely to be the people we most want

[26] Waldron, 187.

anti-discrimination laws to apply to. Therefore, making the laws lose legitimacy only for these people does not really solve anything.

Secondly, it seems to me that the unfairness would occur regardless of what views you personally hold. A majority decision that has been reached without input from all willing parties is, according to the legitimacy argument, simply not a fair decision. A person's personal belief is not what makes it fair or unfair, but rather whether the decision came as the result of a fair process. Hence, this interpretation seems to be incompatible with the structure of the legitimacy argument.

3.2.4 Loss of legitimacy as a matter of degree

The final interpretation that Waldron considers is that legitimacy could be a matter of degrees.[27] Under this interpretation, the existence of upstream legislation would not immediately and completely nullify the legitimacy of a downstream law, as would be the case under the binary interpretation. Instead, it would merely reduce the legitimacy of the downstream law to a level below what it would have had, had the law passed after an unrestricted debate. The size of the reduction would depend on the size of the restrictions imposed by the upstream law.

Let's put aside the details of what this would mean for a moment, and consider if the interpretation is reasonable. My understanding of Dworkin's position is that he saw the issue of legitimacy as, ultimately, an issue of fairness. A law with reduced legitimacy is a law that is more unfair than it otherwise would be. I also do not believe Dworkin had intended for his theory to have any truly radical conclusions. In his writings, I have not found anything which I interpret as encouraging those affected by upstream legislation to break with society, stop obeying laws, or anything of the sort. He is simply arguing that we, as a society, are being unfair to a section of our citizens when we pass laws that restrict them from speaking their mind and attempting to convince others to agree with them.

If that is a correct understanding of Dworkin's position, it seems to me that Dworkin would never want to use unfairness to justify further unfairness. He would never

[27] Waldron, 188.

accept an interpretation of legitimacy that, like the binary interpretation, would allow widespread discrimination simply because the process of passing a law was somewhat unfair. If your ultimate concern is fairness, to do so would, to borrow a proverb, be like cutting off your nose to spite your face.

I also think Dworkin would readily agree that society is not immediately and fully undermined by the existence of some unfairness. Unfairness is, regrettably, an unavoidable condition of both daily human life and of society at large. We can (and arguably should) work to diminish unfairness whenever possible, but we should also keep in mind that fairness is not our only concern. Sometimes, accepting a little unfairness in order to achieve some other goal is a trade-off we have to accept. So, while we should take concerns or complaints about unfairness seriously, having unfairness as a consequence should not by itself disqualify an action from being considered and carried through.

But even if Dworkin would agree with the above, I think that he would nonetheless consider that kind of unfairness to be a significant blemish on the democratic process. After all, fairness is a foundational principle of modern-day democracies – next to self-determination, it is perhaps the principle associated most strongly with it. So, I imagine Dworkin would argue, while the presence of an upstream law does not suddenly give racists leave to discriminate and abuse at will, it *does* give them legitimate grounds to criticize the democratic process on. In his view, passing upstream legislation would be akin to willingly conceding the moral high ground, giving the bigots and racists a leg to stand on where they otherwise would not have had one. It allows them to point towards the upstream laws and say "Look! You preach inclusiveness, fairness, and equality for all, but our opinions are not allowed the light of day. This process is neither fair nor equal, it is simply another form of paternalism." And, most importantly, it would allow them to say it and be *correct*. This, I think, is what Dworkin means when he says that we lose some of the 'legitimacy' of our downstream laws.

Waldron provides further evidence in favor of this interpretation in the form of a quote by Dworkin from the book "Is Democracy Possible Here?", where Dworkin

describes legitimacy as "not an all-or-nothing matter."[28] Adding this quote to my reasoning above, I feel very certain in agreeing with Waldron that this is the most, and perhaps the only, reasonable interpretation of what legitimacy means in Dworkin's theory.

3.2.5 The consequences of legitimacy being a matter of degree

Let's now examine what effect this interpretation will have on the legitimacy argument. Assuming that losing legitimacy is a matter of degree, it follows that the amount lost can be of varying sizes. In the context of the legitimacy argument, the amount lost would be contingent on how great a restriction on speech the law in question imposes on the population. A great restriction would cause a significant reduction, but a small restriction would only cause a slight reduction. This could allow us to set thresholds, where a severe reduction of legitimacy is unacceptable (at some point perhaps completely nullifying its binding force on the population), while a minimal reduction of legitimacy is an acceptable (albeit regrettable) consequence of the practical difficulties of making and implementing democratic decisions. Thresholds are, however, hard to specify and therefore usually end up being arbitrary, which in turn makes them hard to justify.

But it is nonetheless true that a smaller loss of legitimacy would be easier to swallow than a large one. While a huge loss of legitimacy would be a compelling reason to hold up legislation, a small loss of legitimacy would be a much less compelling reason to do so. Even more so if the legislation in question has other desirable consequences. If the ratio of positive outcomes to lost legitimacy is sufficiently great, we would consider the trade-off to be satisfactory, even if we acknowledge that it does create some unfairness.

Accepting this interpretation would therefore introduce the possibility of restricting speech, even for someone who believes that the legitimacy argument has something important to say. It would put the legitimacy argument back in line with my original

[28] Waldron, 188.

assumption that all reasonable free speech theories allow restrictions, as well as with how we generally understand acceptable restrictions of speech in society.

After all, there are many cases, such as the example where someone falsely screams "Fire!" in a crowded theater, where the restriction imposed is so small, and the potential upside so great, that nobody seriously argues against the restriction being in place. To phrase that example in the terms of the legitimacy argument, we could say that having an upstream law which restricts people from falsely shouting "Fire!" in a crowded theater results in an insignificant reduction in legitimacy when compared to the benefits of having such a rule.

3.2.6 How large is the loss of legitimacy due to hate speech restrictions?

Let's return to the specific topic of hate speech, and to Waldron's case in favor of having restrictions on it. Waldron's goal from this point forward is to try to convince us that the legitimacy cost hate speech legislation incurs is, in most cases, very small. To the degree that he can successfully argue this, the force of the objections raised by the legitimacy argument will be correspondingly diminished.

With this possibility in mind, Waldron attempts to show us that hate speech laws, when carefully implemented, are minimally restrictive and thus incur a minimal loss of legitimacy. He gives us two sets of statements containing four statements each, which he uses to attempt to illustrate the scope of the hate speech laws that are currently in place in western-style democracies. I will quote all eight statements in full below. The first set of statements are four examples of possible reasons that some person, X, who is opposed to some anti-discrimination legislation, L, could have for opposing that legislation:[29]

(1) X opposes L because he thinks L will make him worse off.

(2) X opposes L because he thinks L will generate perverse economic incentives, undermining economic efficiency in society.

[29] Waldron, 189.

(3) X opposes L because he distrusts the bureaucracy necessary to administer L.

(4) X opposes L because he denies that the intended beneficiaries of L are worthy of the protection that it offers them.

The only option we are concerned with here is (4), since nothing else would be covered by hate speech legislation. Thus, Waldron divides (4) into four new statements, each giving an example of how X could express the disapproval outlined in (4):

(4a) X expresses his dissent from the broad abstract principle that governments must show equal concern and respect to all members of the community.

(4b) X expounds some racial theory which he thinks shows the inferiority, by certain measures, of certain lines of human descent.

(4c) X gives vent to the view that the citizens who are intended to be protected by the anti-discrimination law are no better than animals.

(4d) X prints in a leaflet or says on the radio that these citizens are no better than the sort of animals we would normally seek to exterminate (like rats or cockroaches).

Waldron claims that the kind of hate speech and group defamation laws that exist in most western democracies today are "almost certain to restrict the expression of (4d), quite likely to restrict (4c), and possibly likely to restrict some versions of (4b), depending on how virulently they are expressed."[30] He goes on to point out that many modern hate speech laws have strict definitions on what the law may or may not censor, often outlining so-called 'safe havens', i.e. ways in which opinions such as (4b) - (4d) can be expressed without resorting to threatening or virulent language, and often exempting expressions made in good faith for the purpose of public interest.[31]

Assuming that Waldron's analysis of the legality of these statements is accurate, I think it seems fair to say that these laws are quite lenient in what types of content they actually restrict.[32] While the most obscene utterances might face restrictions,

[30] Waldron, 190.
[31] Waldron, 190-191.
[32] It is, of course, possible that Waldron's analysis of (4a) - (4d) is incorrect, and that the laws in place in western democracies would in fact impose stricter restrictions. However, I do not myself have the expertise to make a judgment either way. I also do not have any reason to doubt Waldron's judgment,

these laws appear to grant enough leeway that you could express and defend virtually any ideology or position without violating the restrictions. Therefore, it seems difficult to argue that they represent a significant encroachment on the right to free speech, and consequently difficult to argue that they would result in a significant loss of legitimacy.

The fact that is it is possible to express your position without resorting to the most virulent wordings is perhaps why Dworkin insisted on extending free speech protection not just to the terms necessary to explain one's position, but also to any terms that might affect the general moral climate. However, as I have already discussed at length in section 3.2, arguing that we should extend protection to all statements that affect the general moral climate is likely an untenable position either way, and thus it presents little obstacle to whether or not we now judge it acceptable to restrict some such statements.

Satisfied that he has shown the restricted scope of these laws, Waldron attempts to further justify that restrictions are acceptable, or even desirable in many cases, and that we commonly accept comparable restrictions on other types of expressions. He argues that the instances when we most strongly oppose a restriction is when it seems arbitrary, or exceedingly limiting compared to the supposed good it is protecting. To illustrate this, he compares restrictions on expression that are content-based (e.g. those restricting hate speech, threats, or obscene materials) to restrictions on expression that do not primarily concern themselves with what the actual content of the expression being restricted is. In this case, laws restricting public protests.[33]

Most countries, whether democratic or not, have laws that restrict when, where, and under what circumstances you may or may not hold a public protest, commonly requiring a permit to be obtained for protests above a certain size. The government can generally decline to issue a permit for varying reasons and is thus the ultimate arbiter on whether a protest is allowed or not.

---

given his legal background. (Jeremy Waldron has a Bachelor of Law from the University of Otago, New Zealand, and is a professor at the NYU School of Law where he teaches, among other things, philosophy of law.) I will therefore continue under the assumption that his claims are accurate.

[33] Waldron, 191-192.

If these laws were to be used by the state to prohibit protest on an arbitrary basis, e.g. to silence voices of dissent against it, we would likely consider these laws both unfair and illegitimate.[34] Yet despite the possibility for abuse, we accept the existence and execution of these laws without serious complaint, because we recognize that there are legitimate reasons why the government might need to limit when and where you can hold a public protest (e.g. for reasons of public safety and order).

Waldron argues that restrictions on hate speech are no different. If the reasons for restricting speech were arbitrary (such as trying to prevent offense, or to further some aesthetic idea of how a society should portray itself), then hate speech laws would indeed be unfair, and would significantly de-legitimize downstream legislation. But Waldron points out that his basis for wishing to restrict hate speech is not arbitrary feelings of offense or aesthetical judgments, but to enable an integrated, well-functioning and secure society by way of upholding a minimum level of 'elementary dignity'.[35] There is no denying that such dignity has value. It is something that all people desire – both the bigots and the victims of bigotry. Furthermore, to the degree that this feeling of 'dignity' contributes to keeping our society orderly and safe, we can see a direct parallel to the concern for safety with regards to restricting protests.

3.2.7 Evaluating Waldron's closing argument and possible rebuttals

We have now reached the end of Waldron's critique. I must admit that, despite my previous attempts to defend Dworkin's theory, I find myself quite convinced by Waldron's reasoning in this section. I agree with his interpretation of what Dworkin means by legitimacy - while Waldron himself does not present many arguments for why it's the correct interpretation, it seems to me to be the only interpretation of the legitimacy theory that makes any sense. From that interpretation, Waldron's arguments fall into place. Anything that is measured on a continuous scale is necessarily subject to comparisons. Legitimacy is no exception to this rule. At a minimum, it enables us to compare one amount of legitimacy to another amount of legitimacy, in order to determine which is greater, more important, etc. Since a small loss of legitimacy is obviously not as bad as a large loss of legitimacy would be, the

[34] As we often do, e.g. when such laws are employed in an overbearing manner by an authoritarian state.
[35] Waldron, 192.

incentive to refrain from imposing some restriction is correspondingly diminished, provided that the restriction results in such an aforementioned minor loss of legitimacy. What then remains is convincing us that the legitimacy cost of hate speech legislation is, in fact, small. Which I believe Waldron manages to do.

Waldron's analysis of example expressions, and his comparison to laws regulating protests, will of course not satisfy everyone. Many may read his critique and conclude that the legitimacy lost is still too large a cost to bear. But I think most people, myself among them, will read it and agree that restrictions with such a narrow scope do not undermine the legitimacy of our society to such a degree that we should refrain from implementing such restrictions for that reason alone.

But if Dworkin were to mount a defense, what could he argue? One possible defense is to claim that legitimacy is infinitely more consequential than 'dignity'[36] is, in a fashion reminiscent of Mill's idea of higher/lower pleasures.[37] But given the many well-known issues with such an argument, I don't think it would be very successful. To give just a few possible rebuttals: Is legitimacy truly so much more important than 'dignity' that it is not even in the same ballpark? And if so, wouldn't that mean that we have to make many other changes in our society, where we accept diminished legitimacy in favor of e.g. convenience or efficiency?[38] It just doesn't seem credible to argue that this specific aspect is *so important* that it couldn't possibly be overruled by other concerns. And if it can be overruled by some concerns, why not by a concern for 'dignity'?

Another possible defense might be to deny that the legitimacy lost is really all that small. As I already mentioned, that's likely a convincing argument to many. But it is unlikely to convince most, especially those who aren't already ardent defenders of the

[36] Any upside that hate speech restrictions have can, of course, be measured not only in terms of Waldron's 'dignity' but in terms of whatever metric you prefer (e.g., utility). Dworkin can likewise claim that 'legitimacy' is vastly more important than all of them, although I think that argument quickly loses credibility as the list of things that legitimacy trumps grows.
[37] Mill's version of hedonistic utilitarianism divides pleasure into higher and lower such, where higher pleasure represents, e.g., the pleasure gained from reading a good book or discussing your favorite philosophical topic, while lower represents, e.g., the pleasure of sex or good food. Mill argues that no amount of lower pleasure can outweigh the value of higher pleasure. Essentially, quality over quantity.
[38] The entire system of representational democracy is one example of such a concept, where we trade some amount of fairness and legitimacy (albeit a different type of legitimacy than Dworkin's 'legitimacy') for the convenience of not having to vote on every single proposal.

most extensive free speech protections possible. After all, Waldron showed us just how narrow the scope of such laws tends to be, and how we already accept similarly narrow restrictions in other circumstances.

A third possible defense is the reverse of the second, i.e. to argue that while the loss of legitimacy might be minor, the gain from these hate laws is even smaller (or even nonexistent). This is perhaps the best avenue simply because it would, unsurprisingly, be extremely hard to measure any impact (both positive and negative) of such a law in real life. It is hard enough to measure concrete impacts of laws, let alone impacts in terms of vague concepts such as 'dignity' or 'well-being'. Even if such a poll was conducted and a statistically significant change was noted, what's to say that the change is a consequence of the hate speech law?

But while this defense might be hard to disprove, it is not clear to me that it is very convincing. There certainly would be *some* benefit to the people these restrictions would protect, because lessening someone's worry of getting a slur thrown in their face would likely make it easier for that person to get through the day. Dworkin himself refers to the "malign, chilling force of hate speech and hate gesture", calling it "particularly odious" in universities.[39] Surely then there would be some benefit gained from restricting it.

Additionally, there might be a positive response from the symbolic act of passing such a law, which by itself could make the affected minorities feel more accepted. It also seems likely that more people would be positively affected by such a law than would be negatively affected by it, by virtue of there likely being more people in minority groups than there are outspoken bigots/racists.

These considerations, together with the reasonable assumption that the loss of legitimacy really is as small as it seems in Waldron's examples, makes me convinced that the gain from such a restriction is greater than the cost to legitimacy would be.

[39] Dworkin, Foreword, vi.

4. Conclusion

The proper limits of the right to free expression has been a topic of wide disagreement for as long as the right has been conceptualized in the minds of people. This is not likely to change any time soon, or indeed ever. This essay set out to analyze a specific instance of this wider disagreement, an instance which I think has special importance in our society today. As we move towards a more global and more culturally diverse society, it is inevitable that more and more intergroup conflicts will emerge. These conflicts will take on many different forms, including armed conflict, but the vast majority of the conflicts will involve some form of speech. It is important to have discussions not only about which forms of expressions we as a society think are acceptable and which are not (as is done at legislative and societal levels), but also to have meta-discussions about which types of rules we should be allowed to implement.

In the case of this essay, I have come to the conclusion that Dworkin's legitimacy argument does not provide as compelling a reason to abstain from restrictions as Dworkin had hoped it would. Under the most sensible interpretation of Dworkin's argument, it boils down to one question: how big is the loss of legitimacy that would be incurred by a restriction? If the loss of legitimacy is great, the legitimacy argument provides a compelling reason to abstain from implementing the restriction. But if the loss of legitimacy is miniscule, the strength of the argument diminishes correspondingly.

As Waldron demonstrates, the restricted scope of hate speech laws seems to greatly limit the loss of legitimacy incurred. Such a miniscule loss of legitimacy does not seem a compelling enough downside to ignore the upsides such restrictions could produce for our society. This rationale rests on the assumption that hate speech legislation actually *does* produce an upside for our society. I do not think that is a terribly controversial statement, but it is nonetheless something that cannot be taken for granted. For Waldron's argument to work, the upside still has to outweigh the lost legitimacy.

Additionally, while this was not a part of Waldron's original criticism, I believe that Dworkin's insistence that his argument must apply to all expressions that influence the 'general moral climate' causes his argument to grant protection to a wide range of expressions that he probably did not intend to endorse. Since virtually any kind of expression could be interpreted as having an effect on the general moral climate, it is not clear if we could draw a line of acceptability at all. As I have argued, any reasonable theory of free expression must allow for regulation, as there are some forms of expression that we all greatly benefit from restricting. By that standard, Dworkin's theory is not reasonable, and should thus either be heavily reworked or abandoned.

However, while the theory as Dworkin has presented might be flawed, I believe that the concern that Dworkin identifies is well worth paying attention to. In fact, the definitions of 'dignity' and 'legitimacy' each provide insights that may prove useful when debating such restrictions in the future. After all, if we want to uphold the values of liberal democracy, then we should care about any unfairness that our laws impose on our population, regardless of how small the affected group of people may be. Unfairness is generally a bad thing, but it is especially so when the unfairness impacts the credibility of our laws and our societal framework.

On the other hand, we should also strive to make our population feel safe, included and welcomed in our society. Not because (or at least, not only because) it's the nice thing to do, or because of some moral commitment, but because it will benefit us all in the long run. All societies need buy-in not only from its majority population, but from its minorities and sub-groups as well. Without it, a society will become less efficient, less safe, more corrupt, more divided and, eventually, dysfunctional. Therefore, I believe that both 'legitimacy' and 'dignity' identify important considerations, some version of which any conscientious lawmaker ought to have in mind when deciding whether or not to restrict speech in the future.

5. Reference list

Dworkin, Ronald. “A New Map of Censorship.” *Index on Censorship* 23, no. 1–2 (May 1994): 9–15.

Dworkin, Ronald. “Foreword.” *Extreme Speech and Democracy.* Oxford, Oxford University Press, 2009.

Dworkin, Ronald. “The Right to Ridicule”. *The New York Review of Books,* 23/3/2006. Available at https://www.nybooks.com/articles/2006/03/23/the-right-to-ridicule/

Mill, John Stuart. *On Liberty*. Indianapolis, Indiana: Hackett Publishing, 1978.

Waldron, Jeremy. *The Harm in Hate Speech*. Cambridge, Massachusetts: Harvard University Press, 2012.