LOCKE STUDIES

Vol. 17

<https://doi.org/10.5206/lis.2017.882> | ISSN: 2561-925X

Originally published: 2017

Published online: 01 JANUARY 2019

© Locke Studies, 2017

Locke and Hate Speech Law: A Critical Review

J.K. NUMAO (KEIO UNIVERSITY)

Recommended citation:

Numao, J.K. "Locke and Hate Speech: A Critical Review." *Locke Studies* 17 (2017): 173–96.

<https://doi.org/10.5206/lis.2017.882>.

For more information about this article:

[see this article's webpage.](#)

Locke Studies is published by [The John Locke Society](#).

This is an open access article published under the terms of the [Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International](#) license, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and shared under the original license.

LOCKE AND HATE SPEECH LAW: A CRITICAL REVIEW

J. K. NUMAO

§1. Introduction

Hate speech is a high profile issue in many liberal democracies today. While commentaries by constitutional experts and jurists abound in the press, and by legal and political philosophers in academia, it is remarkable that there is far less contribution from students of history of political thought and intellectual history, especially of the early modern era, considering how largely the theme of religious toleration and intolerance featured in this period.¹ Jeremy Waldron's *The Harm in Hate Speech*, and more specifically Chapter 8 entitled 'Toleration and Calumny', helps to break this silence, making a case for how Enlightenment toleration theories from Locke to Voltaire might connect and enrich our discussions about hate speech today.²

In this article, I try to further the discussion on Locke, exploring how his views about toleration and law might fit into modern debates over hate speech legislation. To my knowledge, there are no sustained or extensive discussions of Locke and hate speech law in the scholarship apart from Waldron's chapter and Teresa Bejan's response to it.³ While both works help to rectify

¹ For example, in Michael Herz and Peter Molnar (eds.), *The Content and Context of Hate Speech* (Cambridge, 2012), a recent collection of essays and interviews reflecting on the possibility of regulating hate speech, only a few contributions make use of early modern resources. See especially Alon Harel, 'Hate Speech and Comprehensive Forms of Life', and Stephen Holmes, 'Waldron, Machiavelli, and Hate Speech'.

² Jeremy Waldron, *The Harm in Hate Speech* (Cambridge, MA, 2012). Chapter 8 was originally delivered as an Oxford Amnesty Lecture in 2010. See Jeremy Waldron, 'Toleration and Calumny', in Kate E. Tunstall (ed.), *Self-Evident Truths?: Human Rights and the Enlightenment* (New York and London, 2012), 209–37. While the two texts are for the most part the same, there are some key changes in the book—notably, there is a stronger claim about Locke's call for a legal response to hate speech.

³ Teresa Bejan, 'Locke on Toleration, (In)civility, and the Quest for Concord',

the gap in the scholarship, I show that there is a different story to uncover and so also a different lesson to draw from Locke. In what follows, I first examine Waldron's chapter and his case for a Locke who tends towards the regulation of hate speech. I then consider Teresa Bejan's response to Waldron and her case for a Locke who is sceptical of regulation. By reassessing the evidence offered by Waldron and Bejan, and also by drawing attention to other resources from Locke's political and legal philosophy as well as his philosophy of language, I hope to offer a more nuanced interpretation of Locke, a Locke who is open to the idea of regulating certain forms of speech including what we might categorize as hate speech, but also one who is circumspect about introducing a law suppressing 'hate speech' in a catch-all sense. Finally, I reflect on the legislative implication of the revised Locke with reference to the recent hate speech debates in Japan.

§2. Waldron's Case Considered

In Chapter 8 of *The Harm in Hate Speech*, Waldron offers an interpretation of Locke as one who sees religious hate speech to be incongruent with the idea of a tolerant society and further believes such speech is the subject of legal regulation.⁴ In this section, I review Waldron's case critically, pointing out some

History of Political Thought 37 (2016): 556–87. There are works that make brief mentions of and connections between Locke and hate speech—see, for example, Greg Forster, *John Locke's Politics of Moral Consensus* (Cambridge, 2005), 20; Harel, 'Hate Speech', passim; Milton Heumann and Thomas W. Church, 'Introduction', in Milton Heumann and Thomas W. Church (eds.), *Hate Speech on Campus: Cases, Case Studies, and Commentary* (Boston, MA, 1997), 18; Steven J. Heyman, 'Hate Speech, Public Discourse, and the First Amendment', in Ivan Hare and James Weinstein (eds.), *Extreme Speech and Democracy* (Oxford, 2009), 167–68; Alex Tuckness, *Locke and the Legislative Point of View* (Princeton, NJ, 2002), pp. 51–56; idem, 'Rethinking the Intolerant Locke', *American Journal of Political Science* 46, (2002): 297. For a more substantial treatment of the topic, see Samuel V. LaSelva, 'Toleration Without Hate Speech: The *Keegstra* Decision, American Free Speech Exceptionalism and Locke's *Letter*', *Canadian Journal of Political Science* 48 (2015): 699–718. I regret that LaSelva's article came to my attention after I had submitted this article, but fortunately we make different points.

⁴ On what we would call 'religious hate speech' in Enlightenment theories, see Waldron, *The Harm*, 208–10.

crucial weaknesses in his arguments. To help put my discussion into context, I start by saying something briefly about Waldron's general aims in the book and Chapter 8.

In the book as a whole, Waldron tries to show that hate speech causes a serious harm that calls for regulation. Hate speech undermines a person's 'dignity', defined as a person's 'basic entitlement to be regarded as a member of society in good standing'. In undermining this basic entitlement, Waldron argues, hate speech destroys the public good of 'inclusiveness' to which we are committed—that we can go about our business with the assurance that we do not have to face discrimination or exclusion by others. The aim, therefore, of hate speech laws is to protect this public good.⁵ Taking a historical turn in Chapter 8 of the book, Waldron tries to bridge contemporary debates on hate speech and historical ones on religious toleration in the seventeenth and eighteenth century, which, he observes, are pursued in the legal and philosophical literature 'as though they had nothing to do with each other'. In doing so, he aims 'to add a dimension of historical richness' to what he sees as the 'often flat and colorless' constitutional debates about hate speech.⁶

The more specific questions Waldron asks in this chapter are threefold: first, whether the Enlightenment thinkers touch on the theme of hateful defamation in their theories of toleration; second, (if they do say something about the theme) whether they see hateful defamation as something to be avoided or regulated; and third, what a 'tolerant society' looks like in their minds.⁷ How does Locke turn out? Concerning the first question, Waldron observes that while Locke does not say much about 'how we should regard vituperation in the context of religious diversity', the theme is there in the *Letter Concerning Toleration* (1689) if

⁵ Ibid., 4–6, 105.

⁶ Ibid., 16–17, 204, 231. For a reflection on the significance of Waldron's recourse to history, see Liora Lazarus, 'Rights Persuasion: A Response to Jeremy Waldron', in Tunstall (ed.), *Self-Evident Truths*, 239ff.

⁷ Waldron, *The Harm*, 207.

we ‘dig a little’.⁸ Waldron’s excavation then yields three themes from Locke, which together constitute the responses to the second and third questions. The first theme is that ‘public expressions of hatred and vilification are typical of an intolerant rather than a tolerant society’ (a response to the third question). The second is that ‘there is a specific duty—perhaps even a legal duty—to refrain from rough usage of word, as well as rough usage of action, if that is calculated to have a detrimental impact on an individual’s person or honour or estate’ (a response to the second question). And the third that ‘the duty of toleration is bound up with a general duty of charity, civility, and good fellowship’, or in other words, a duty to maintain a peaceful relationship amid diversity (likewise, a response to the second question).⁹

Now the fact that Waldron had to ‘dig’ for evidence might suggest that Locke was not really concerned with religious hate speech or committed to suppressing it. Waldron anticipates such criticism and rejects it as being ‘premature’. For one, while these arguments may not be at the forefront, the fact remains that Locke did say something against expressions of religious hatred. There is also the ‘direction’ or ‘tendency’ of Locke’s overall arguments, which, Waldron argues, shows that he did not think that religious defamation merited the usual toleration.¹⁰ What, then, does Locke actually say about religious hate speech and where are his overall arguments heading on this issue? Does the evidence justify Waldron’s reading of Locke as one in favour of regulation?

As evidence of our duty to refrain from using abusive language to harm someone’s person or honour or estate, Waldron cites and notes Locke’s objection to the ‘rough usage of word or action’ by the church against excommunicates, and his doubts regarding the magistrate’s use of attacks on people’s honour; and as evidence of Locke’s approval of regulating hateful speech by law, Waldron adduces his likely involvement in the drafting of the *Fundamental Constitutions of Carolina* (1669), which includes an article

⁸ Ibid., 211, 221.

⁹ Ibid., 213–15.

¹⁰ Ibid., 221–22.

(Article 97) forbidding the use of ‘reproachful, reviling, or abusive language against the religion of any church or profession’.¹¹ While Waldron helpfully sheds light on these relevant but overlooked points, I think that there is a gap between what Waldron wants to show and what the evidence can justify, and thus a problem in the way he uses the passages to build his case.

Broadly speaking, there may be two general levels on which we could say Locke is opposing religious hate speech. One is on the state level (political and legal level), where we could say Locke is arguing for the state to intervene through law or other means to counter such speech. The other is the societal level (moral and social level), where we could say Locke is calling on the members of society not to engage in hateful speech. As I see it, Waldron wants to connect the two levels.

Both in the book as a whole and in Chapter 8, Waldron’s argument seems to run as follows. A tolerant society requires more than people being able to say or believe what they want; it requires an environment in which people can rest assured that they will be able to enjoy a certain degree of respect as members of that society. Where there is moral slack in society, there is a role for law to uphold such an environment. Waldron brings to light passages from Locke that purportedly touch on our moral, social, political, and legal obligations to counter religious hate speech. But while the evidence he provides to establish our societal duties of treating other fellow citizens with a certain degree of respect (e.g. our duty to maintain ‘love and charity in the diversity of contrary opinions’) is strong, the evidence to establish our political and legal duties to counter religious hate speech is much weaker. As a result, I believe that the connection Waldron wants to make cannot be sustained in the case of Locke.

¹¹ Ibid., 211–13. Article 97 in full reads: ‘No person shall use any reproachful, reviling, or abusive language against the religion of any church or profession, that being the certain way of disturbing the public peace, and of hindering the conversion of any to the truth, by engaging them in quarrels and animosities, to the hatred of the professors and that profession, which otherwise they might be brought to assent to’. John Locke, ‘The Fundamental Constitutions of Carolina’, in Mark Goldie (ed.), *Locke: Political Essays* (Cambridge, 1997), 176.

For the most part, Waldron does not say explicitly which level of duty he is trying to justify by citing a certain passage. I assume, however, that he cites the passage where Locke expresses doubts against attacks on people's 'honour' as well as their 'estate' and 'liberty' to justify Locke's opposition to hate speech on the state level qua threat to one's honour and dignity. The relevant passage is as follows: 'Loss of estate and dignities may make a proud man humble...But will you therefore infer, that the magistrate may take away a man's honour, or estate, or liberty, for the salvation of his soul'.¹² The implied answer is that the magistrate may not, and therefore that the state cannot rightfully use force to strip people of these things for their religious beliefs. Without analyzing this passage, Waldron then quickly moves on, and referring to Article 97 of the *Fundamental Constitutions*, talks about how 'this duty' (i.e. the duty to refrain from attacking someone's honour as well as person and property) might be a duty upheld by law on Locke's account.

As Waldron suggests, honour does seem to be something to be upheld by law. There is an interesting passage overlooked by Waldron, which helps to link Locke's desire to protect one's honour with law. In the *Letter*, Locke explicitly states that 'slanderers' along with the 'seditious, murderers, thieves, robbers, adulterers' are to be 'punished and suppressed'.¹³ This seems to suggest that Locke would be in favour of having defamation laws, which include laws against slander and libel. However, even if this is the case, it would be premature to conclude that Locke would support current hate speech laws. This is because while some democracies (e.g. Japan) have defamation laws, which are designed to protect the reputation of individuals and corporations, these countries (e.g. US and until recently Japan) do not have robust hate speech laws, which, by contrast, are concerned with groups. Of course, it is possible to see hate speech in terms of 'group libel' as Waldron argues, and indeed there are countries

¹² John Locke, 'A Second Letter Concerning Toleration', in *The Works of John Locke*, (10 vols., London, 1823), VI, 69.

¹³ John Locke, 'A Letter Concerning Toleration', in Mark Goldie (ed.), *A Letter Concerning Toleration and Other Writings* (Indianapolis, IN, 2010), 58.

that have group defamation laws for what we are calling hate speech laws. However, just from these fragmentary remarks, we cannot say whether Locke is envisaging an individual-based or group-based application when talking about countering slander.

The important question to ask, therefore, is what the content of Lockian honour is. This takes us to the fundamental problem with Waldron's move from Locke's call to protect honour to a Lockean justification of the use of law to this end. I suspect that in this section, what Waldron wants to do is to link Locke's use of the terms 'honour' and 'dignity' with his idea of dignity as a person's entitlement to be regarded as a member of society in good standing, and thence to argue that Locke calls to protect this by law. However, Waldron provides no analysis of what Locke means by honour in the cited passage or what he is trying to do in calling for the protection of one's honour. This is fatal because, contrary to Waldron's expectation, I do not think that Locke uses these terms in the sense of a person's *basic* standing in society. Rather Locke's use of 'honour' and 'dignity' seems to fit more neatly with his language of 'property'—those things that rightfully belong to us and that government exists to protect.¹⁴ Locke uses both terms in his other writings and in different ways.¹⁵ He does talk about the 'dignity of man', which consists of being able to exercise religious liberty, but for the most part dignity and honour are about 'rank, title, worth' and 'virtue, praise, glory, respect' respectively, which we can see as being *added* to the list of proper things belonging to people who are in principle moral equals.¹⁶ These things rightfully belong to

¹⁴ John Locke, *Two Treatises of Government*, ed. Peter Laslett (Cambridge, 1988), I.90, 208.

¹⁵ See especially Laurie M. Johnson, *Locke and Rousseau: Two Enlightenment Responses to Honor* (Lanham, MD, 2012), esp. ch. 2.

¹⁶ For 'dignity of man', see, for example, John Locke, 'An Essay Concerning Toleration', in J.R. Milton and Philip Milton (eds.), *An Essay Concerning Toleration and Other Writings on Law and Politics, 1667-1683* (Oxford, 2006), p. 294. For dignity as rank, title and worth, see, for example, John Locke, 'Atlantis', in Goldie (ed.), *Political Essays*, 254; Locke, *Letter*, 24; Locke, *Two Treatises*, I.44, 172. For some examples of honour in the sense of virtue and respect, see Locke, *Two Treatises*, II.68, 313, II.111, 342. More generally, see Johnson, *Locke and Rousseau*, ch. 2.

people but they are not basic entitlements. Locke's was a society of titles, Locke himself being a 'gentleman'.¹⁷ It was conceivable for people to be moral equals and at the same time for some to have titles. Waldron might say that it is possible to talk about how today we accord to everyone a high and equal rank, which is the dignity people rightly enjoy.¹⁸ But this further development would be Waldron's, not Locke's. Thus, it does not seem to me that the passage about protecting honour from attacks can do anything to further Waldron's case.

Wherever Locke's comments about the protection of honour may leave us, Article 97 of the *Constitutions* seems convincingly to establish a case for regulating religious hate speech by law. What might make it more relevant in light of the discussion about hate speech laws is that it prohibits abusive language against more general entities than individuals, namely 'religion of any church or profession'. However, the problem of relying on this piece of evidence too much is that while it can help to show the existence of *seventeenth-century* oppositions to religious calumny, it is a bit weak as evidence to establish *Locke* as an advocate of regulating religious hate speech, since we do not know the precise extent to which he was involved and contributed to this work.¹⁹

Now, while the evidence Waldron provides for what Locke actually says about religious hate speech seems to be inadequate to put Locke in the pro-regulation camp, there is still the general 'tendency' towards regulation to consider. What then is the 'tendency' Waldron identifies in Locke? First, there is the tendency that denies calumny as an 'offensive strategy'. Locke

¹⁷ Peter Laslett, 'Introduction' [to his edition of *Two Treatises*], 42.

¹⁸ Waldron, *The Harm*, 46–47. See also Jeremy Waldron, *Dignity, Rank, and Rights*, ed. Meir Dan-Cohen (Oxford, 2012), Lecture 1.

¹⁹ See, generally, David Armitage, *Foundations of Modern International Thought* (Cambridge, 2013), 90–113; J.R. Milton and Philip Milton, 'General Introduction', in Milton and Milton (eds.), *An Essay Concerning Toleration*, 118. Even if we suppose that Locke could have amended this particular article in 1682 and did not, while this might suffice to suggest that he supported the article in view, it would not suffice to show that he would have supported the idea of anti-hate speech laws in general. Armitage, *Foundations*, 105–7.

denies the use of force by the civil magistrate in religious matters because of the ineffectiveness of it to bring about genuine change of belief and also the lack of a Godly commission to use such means; likewise, Waldron maintains, the use of public calumny to prompt conversion would fall prey to these same arguments. While religious insults might force people to change their religious behaviour, it would most likely only get them to conceal their real beliefs, and so would not be effective as a means of achieving genuine conversion. Next, there is also the tendency that denies calumny as a ‘defensive strategy’, or as a tactic to warn those of orthodox faith against the danger of conversing with heretics. This too, Locke seems to intimate, is ineffective. The proper way for religious leaders to keep ahold of believers is by ‘exhortations, admonitions, and advices’, not by raillery and abuse. Finally, there is the tendency that denies calumny ‘not as a means to an end’ but as a form of ‘self-expression’. Waldron conjectures that if self-expression is just a matter of ‘letting off steam’, then ‘Locke’s arguments for social peace and civility [would] require people to find other outlets’. But what if what one feels or perceives as hate speech is for another a natural form of expressing one’s own views? While Locke holds that people are commissioned ‘to admonish, exhort, convince another of error, and, by reasoning, to draw him into truth’, he also holds that ‘Nothing is to be done imperiously’, or, as Waldron translates, ‘by way of sanction’. So while expressing disagreement is acceptable, reviling others in order to hurt or punish them is not. Forceful disagreement might be acceptable to the extent that we can still engage intellectually with the opponent.²⁰

Waldron shows quite persuasively that there is a tendency towards prohibiting religious hate speech in Locke’s thought.

²⁰ Waldron, *The Harm*, 222–28. Locke’s own exchange with John Edwards and Richard Willis in the two *Vindications of the Reasonableness of Christianity* might serve as examples of this, although the former was testing the limits of Locke’s patience. Victor Nuovo, ‘General Introduction’, in Victor Nuovo (ed.), *Vindications of the Reasonableness of Christianity* (Oxford, 2012). In any case, if the attacks against one’s works are less about criticism for the sake of truth than slandering, ‘disregarding’ might be an option. John Locke, *The Correspondence of John Locke*, ed. E. S. De Beer (8 vols. [to date], Oxford, 1976–), VI, 679.

However, the difficulty with his case is that he relies mostly on material from the *Letters on Toleration*. To talk about the tendency of the ‘overall argument’, we need to broaden the scope of investigation. This is what Bejan does. However, as we shall see, she shows that there is material in Locke’s other writings that suggests he was critical of a legislative response to religious defamation.

§3. Bejan’s Case Considered

In her article, ‘Locke on Toleration, (In)civility, and the Quest for Concord’, Bejan shows contra Waldron, and through a more comprehensive and chronological look at his writings, that Locke was sceptical towards regulating religious hate speech. The article itself attempts to refute those who see Locke as representing an ethically minimal, ‘thin’ account of toleration, arguing that he in fact had a more complex and richer idea of toleration, which developed over time. Bejan sees the development of Locke’s thought as a struggle over how best to ‘civilize’ the disagreements amongst various religious groups and the societal hatred acting as a catalyst of this, and traces out Locke’s changing views of civility, toleration, and the role of legislation from his early to his later works. Bejan argues that Waldron’s earlier works on Locke come under the ‘thin’ heading, while his more recent works, which include the book on hate speech, align more with the revisionist ‘charitable’ stance. However, even Waldron’s more recent work is partial and, as we shall see, only captures some of the characteristics of the middle Locke according to Bejan.²¹

Bejan maintains that the early Locke saw religious diversity as a threat to civil peace, believing that it acted as a catalyst intensifying the hatred within society. In his view, English society at the time was full of sectarian hatred, religious enthusiasm and over-sensitivity. While toleration was an attractive idea, it was too dangerous to experiment with on a practical level. Thus, uniformity was needed for civil peace, and this would be achieved through external conformity to standards imposed by the

²¹ Bejan, ‘Locke on Toleration’, 557–61.’

civil magistrate concerning the rules of civility and religiously indifferent things. Civility at this point only involved conformity to contingent cultural norms of politeness. The middle Locke makes a partial but important break from the early Locke. Locke no longer sees religious diversity *per se* as an intensifier of societal hatred. Rather, the blame was to be attributed to incivility—the uncivil, harsh language—which led to enmity between people. Through the suppression of uncivil language and a policy of toleration rather than uniformity, Locke believed that society could foster a sense of mutual trust. Thus the *Carolina Constitutions* forbade ‘the use of reproachful, reviling, or abusive language’ and the ‘Essay Concerning Toleration’ called for religious toleration to the extent that it was compatible with public order.²²

However, the story of Locke’s development does not stop here. According to Bejan, the later Locke continues to see incivility as the cause of religious hatred. There is, however, a shift in his idea of civility and thus a shift in his concept of toleration. From a virtue of external conformity to contingent norms, civility became an inward disposition connected with charity. Toleration too was re-imagined as a positive idea of concord amongst society from a negative idea of permission without respect. Concurrently, the role of the state receded. For Locke, the ‘best method of popular instruction was *not* legislative proscription’; he encouraged, rather, that gentlemanly people set examples of civility.²³

Bejan’s crucial piece of evidence here to establish Locke’s doubts over a legislative response to religious defamation is not from the *Letter* but the largely neglected notes on William Penn’s ‘Pennsylvania Laws’ (1686).²⁴ Penn’s proposal, which resembles the anti-defamation article in the *Carolina Constitutions*, reads: ‘And if any person shall abuse or deride any other for his different persuasion and practice in matters of religion, such shall be looked on as a disturber of the peace and be punished accordingly’. In response to this, Locke notes critically: ‘a matter

²² Ibid., 561–70.

²³ Ibid., 570–86.

²⁴ Ibid., 579–81.

of perpetual prosecution and animosity'.²⁵ Bejan first notes Locke's objection to anti-blasphemy laws in Penn's proposal and the 1698 Blasphemy Act and how he thought they engendered definitional imprecision, which in turn could lead to their being abused. She then argues that Locke likewise feared that Penn's anti-religious hate speech law could cultivate an antinomian culture by allowing people to become the judges of offence and derision. Over-sensitive people would be encouraged to prosecute each other rather than working out disagreements in a spirit of civility. Thus as Locke writes in a different manuscript, while there were many things 'of great use', they may not be 'well established by laws'.²⁶

Bejan offers compelling evidence against Waldron for a Locke advocating non-legislative measures to counter hateful defamation. However, by recasting the evidence in a different light, I believe that it is possible to draw a different conclusion from Bejan's. First, I agree that Locke was critical in his commentary on Penn's proposed frame of government. However, by looking at the general flow of the criticisms and Locke's thoughts expressed in his other writings about the subjects he comments on in the 'Pennsylvania Laws', it is possible to see Locke objecting to the wording of the articles but not necessarily rejecting the existence of such legislation altogether.

Take, for example, Locke's note on adultery. Commenting on the article 'Adultery to be punished with twelve months imprisonment in the house of correction, and longer if the chief magistrate think meet',²⁷ Locke writes, 'Arbitrary power'.²⁸ This comment, we might suppose, is a reference to the latter part of the

²⁵ John Locke, 'Pennsylvania Laws', in Mark Goldie (ed.), *A Letter Concerning Toleration and Other Writings* (Indianapolis, IN, 2010), 182.

²⁶ John Locke, 'Marriage', in Goldie (ed.), *Political Essays*, 273. Other material includes, for example, Locke's 'Liberty of the Press', where he notes the definitional difficulty and hence the possibility of abuse of terms such as 'heretical, seditious, schismatic, or offensive'. John Locke, 'Liberty of the Press', in *ibid.*, 330–31.

²⁷ Meet: suitable, fit, proper. See *OED* meet, *adj.*, 2.

²⁸ Locke, 'Pennsylvania Laws', 182.

sentence ‘longer if the chief magistrate think meet’. But we needn’t conclude from this objection that Locke is against or expressing doubts about punishing adultery, not least because Locke argues in the *Letter* that the ‘seditious, murderers, thieves, robbers, *adulterers*, slanderers’ are to be ‘punished and suppressed’.²⁹ Likewise, Locke responds to the article, ‘Incest: forfeiture of half the estate’ with the question ‘to whom?’.³⁰ Again, he is expressing doubts regarding Penn’s proposal. However, Locke’s remark in the *Third Letter for Toleration* (1692) about the magistrate’s power to punish those who marry ‘within degrees thought forbidden’ suggests that he thought incest was punishable by law.³¹ In any case, both adultery and incest were breaches of the law of nature, by which our duties to ourselves, our fellow human beings, and God, are explained and upon which the positive laws of commonwealths are founded.³² Given this, we might see Locke’s comment not as a complete rejection of a law against incest as such, but rather a query over the punishment specified in the draft.

The case for Locke as one concerned with the details of a piece of legislation can also be supported by his remarks in the *Second Letter Concerning Toleration* (1690). In this work, responding to the Anglican divine Jonas Proast, critic of his *Letters* and the advocate of the use of indirect and moderate punishments by the magistrate to encourage dissenters to reconsider their faith, Locke demanded that the clergyman clarify the following set of questions should he wish to justify his policy: ‘1. Who are to be punished. /2. For what. /3. With what punishments. /4. How long. /5. What advantage to true religion it would be, if magistrates

²⁹ My emphasis. See also John Marshall, *John Locke, Toleration and Early Enlightenment Culture* (Cambridge, 2006), 715–16.

³⁰ This part is not included in the Goldie collection. John Locke, ‘Pennsylvania Laws’ (1686). Bodl. MS. Locke f. 9, 33. See also Maurice Cranston, *John Locke: A Biography* (London, 1968), 261–62.

³¹ John Locke, *A Third Letter for Toleration*, in *Works*, VI, 241–42; John Witte Jr., *The Western Case for Monogamy over Polygamy* (Cambridge, 2015), 364–67.

³² Locke, *Two Treatises*, I.59, 183, II.12, 275.

everywhere did so punish. /6. And lastly, whence the magistrate had commission to do so.' Only when he satisfies these points, which, as we can see, include not only fundamental political questions about the magistrate's rights but also such detailed matters as the length of punishments, would his laws be considered 'consistent and practicable'.³³

Returning now to Locke's comment on the anti-religious defamation article of the 'Pennsylvania Laws', if I am right, it is possible to see the comment 'a matter of perpetual prosecution and animosity' not as an objection to anti-defamation legislation as such, but as an objection to the proposal as it stands. One possible problem with the proposal might be the idea of punishing 'derision', as Locke himself considers some beliefs 'false and absurd'.³⁴ If people can prosecute others for simply deriding them, and not for something more concrete such as injury to one's possessions, this could lead to endless cases being brought to court. It is noteworthy therefore that the idea of punishing people for deriding others concerning religious matters was not part of the anti-religious hate speech legislation of the *Carolina Constitutions*.

§4. The Case for an Open-Ended Locke

Let me summarise the discussion so far. Looking at both what Locke says and where his arguments seem to be heading, Waldron tries to show that Locke would be in favour of regulating religious hate speech by law. He argues that we can bridge Locke's discussion about religious hate speech and modern day discussions, and intimates that there is material in Locke making it a political and legal issue. However, I have tried to show that both from what Locke says about religious hate speech and what his arguments seem to imply, there is not enough evidence to conclude that he would positively opt for legislation. Note that I am not saying that Locke would be against it. In fact, I believe that there is space in Locke's thought for legislation against hate

³³ Locke, *Second Letter*, 111.

³⁴ Locke, *Letter*, 44.

speech but not for the reason Waldron suggests. Bejan, on the other hand, tries to show that (the later) Locke would not be in favour of regulating religious hate speech by law. She presents us with compelling evidence against Waldron that Locke was aware of the difficulties accompanying legislation. However, I have tried to show that the evidence can also point to a different direction. While Locke can express worries about the definitional imprecision of hate speech laws and the possibility of their being abused, he does not necessarily have to object to the idea of introducing hate speech laws should they be able to satisfy the wording problem. Again, I believe that there is space in Locke's thought for legislation against hate speech, but with more reservations than Waldron's Locke.

Developing my critical observations above, I want to make the following two points regarding the place of Locke in hate speech law debates. First, that Locke would be open to the idea of regulating hate speech if such speech causes harm to our property. Second, that while the elusiveness of the wording could be seen as an argument against legislating a hate speech law, for Locke, elusive needn't imply impossible. Let me take these in turn.

4.1 Property vs Dignity

As we saw above, when Waldron considers Locke's political and legal response to religious hate speech, while taking note of the property-based harms, he emphasizes the harms to one's honour, with the apparent implication that this is going to be of special relevance. I maintain that Locke's arguments on this issue are best understood from the perspective of preserving our property. As Waldron observed, Locke's main concern in the *Letter* is to clarify when it is appropriate for the state to intervene with the use of force. Locke's answer is straightforward here: the state exists to protect and preserve our '*Civil Interests*'—which includes 'Life, Liberty, Health, and Indolency of Body; and the Possession of outward things, such as Money, Lands, Houses, Furniture, and the like'. To this list we may indeed include 'honour' or 'dignities' (but not in Waldron's broad sense)—or, in the language of *Two Treatises of Government*, 'Property'—which

is a general name for ‘Life, Liberties, and Estates’.³⁵ A threat to our civil interests therefore becomes the proper subject of the state. So, if hate speech harms our civil interests, then it becomes an actionable threat. Locke was not a defender of absolute free speech; there were clear limits to what people could say. Well-known among these were misanthropic opinions, politically seditious doctrines based on arbitrary power, and the denial of a God.³⁶ This was necessary to preserve the fabric of civil society. Less known until Waldron’s efforts is that there were also limits to *how* people could say things. But the crucial point in the latter case is where the emphasis lies.

Consider again Locke’s objection to ‘rough usage’ of language. Locke argues that when a church excommunicates its members, it should carry ‘no rough usage of Word or Action, whereby the ejected Person may any wise be damnified in Body or Estate’.³⁷ When Waldron cites this passage, he emphasizes the former part about the use of rough language being a limit to what someone can say, but the key is not so much the roughness of the speech as what it *does*, as we can see from the dependent clause modifying ‘rough usage of word or action’—whereby a person might be ‘damnified in Body or Estate’. Speech that causes damage to a person’s bodily and material integrity, or in other words their civil interests or property, becomes the subject of regulation.

So the story is quite simple. For Locke, any action or speech that in content or manner causes damage to a person’s bodily and material integrity becomes the subject of regulation; if not, it does not. The emphasis here is on what causes harm to our civil interests; what that harm is called seems to be less of a concern. Given that there may be a range of hate speeches, not all of which may be harmful to our civil interests, then the question is not whether it belongs to the category ‘hate speech’, but whether it does or does not threaten our civil interests. It is possible to take a

³⁵ Locke, *Letter*, 12; Locke, *Two Treatises*, II.123, 350, I.90, 208.

³⁶ Locke, *Letter*, 49–53.

³⁷ *Ibid.*, 19.

blanket approach and regulate everything that falls into the category ‘hate speech’, but Locke does not seem to be fond of such an approach. For a perceived threat to be truly a threat depends on it being a threat to our civil interests. That threat being a threat, however, could also be a contingent fact. Thus in the *Letter*, Locke did not call for an anti-Catholic laws as such: without naming Catholics and so speaking more generally, he called for the suppression of a threat to the stability of the state, which captured contingent characteristics of a certain kind of Catholic faith (of course, contingent does not mean it is easy to dissociate oneself from those characteristics).³⁸ By analogy, while Locke may be in favour of regulating certain speeches, it seems likely he would not support a blanket regulation of hate speech; he would try instead to isolate the relevant speeches that cause harm.

4.2 Law and Locke’s Philosophy of Language

Let me turn to the second point. For Bejan’s Locke, legislating a hate speech law engenders definitional imprecision, and because of this imprecision, may open it up to abuse by individuals. The problem of uncivil language therefore is best dealt with on the moral and societal level rather than the political and legal. I have suggested that Locke may have simply been pointing out the inadequacy of the wording of Penn’s proposal and so was not necessarily objecting to the legislation of a religious hate speech law. I have also suggested above that Locke would not call for a blanket regulation of hate speech. While I agree with Bejan that Locke would be concerned about the definitional problem, this difficulty does not have to mean that Locke’s default response would be non-legislative. Of particular interest and relevance in this context is Locke’s philosophy of language.

Locke’s philosophy of language at first sight seems to lend support to Bejan’s case. In Book III of the *Essay Concerning Human Understanding* (1690), Locke talks about the imperfection in the relation between words and the signification they have for different people, and also about those who wilfully

³⁸ Ibid., 50–52. See also Locke, ‘An Essay’, 284–85.

abuse or are negligent of the senses of words.³⁹

For Locke, words stand for, and only stand for, ideas in the mind of speakers. Ideas are ‘simple’ or ‘complex’, the former serving as the building blocks of the latter. The names of complex ideas are of two types: ‘substances’ or (for simplicity’s sake) ‘things’, and ‘modes’ or (again for simplicity’s sake) ‘affections of substances’ (under which a third type, ‘relations’, also falls).⁴⁰ While there are real world reference points for substances (think of ‘gold’ for example), ‘mixed modes’ (which consist of ‘several Combinations of simple *Ideas* of different kinds’) for the most part do not have such reference points in nature.⁴¹ They are ‘made by a voluntary Collection of *Ideas* put together in the Mind, independent from any original Patterns in Nature’, and this is how lawmakers can legislate about actions which have never been committed.⁴² Mixed modes such as ‘*Honour, Faith, Grace, Religion, Church*’ give rise to controversies, and concerning the consequent implication for laws, Locke says, ‘in the interpretation of Laws, whether Divine, or Humane, there is no end; Comments beget Comments, and Explications make new matter for Explications: And of limiting, distinguishing, varying the signification of these moral Words, there is no end’.⁴³ On the philosophical level, substance as well as modal ideas are plagued by an unstable relationship between the words and their senses, the former due to lack of knowledge about real essences and the latter, as indicated, by the lack of real world standards.⁴⁴ Thus,

³⁹ John Locke, *An Essay Concerning Human Understanding*, ed. Peter H. Nidditch (Oxford, 1975 [repr. 1979]), III.ix.1, 475–76, III.x.1, 490.

⁴⁰ Locke, *An Essay*, III.ii.2, 405; II.xii.3–4, 164–65. See Hannah Dawson, *Locke, Language and Early Modern Philosophy* (Cambridge, 2007), 222. For a discussion about the complexity of substances and modes, Antonia LoLordo, *Locke’s Moral Man* (Oxford, 2012), 74–82.

⁴¹ For ‘mixed modes’, Locke, *An Essay*, II.xxii.1, 288. For ‘for the most part’, see III.ix.7, 479.

⁴² *Ibid.*, III.v.5, 430.

⁴³ *Ibid.*, III.ix.9, 480.

⁴⁴ *Ibid.*, III.ix.11, 481–82.

there is an inherent imperfection in language, and the words speakers use to express their ideas do not necessarily evoke the same ideas in their hearers.

This natural imperfection is not the only problem. In addition to the inherent instability, there are abusers of words, who are guilty of '*wilful Faults and Neglects*'. One such abuse is '*affected Obscurity*', by which the speaker uses existing terms in an unusual way and confounds the ordinary sense. Because of this 'Art of Disputing', words are 'much more obscure, uncertain, and undetermined in their Meaning, than they are in ordinary Conversation'.⁴⁵ Such abuses have consequences beyond academic discussions of logic; they also impact on society itself. Of its effects on law, Locke writes: 'How else comes it to pass, that Princes, speaking or writing to their Servants, in their ordinary Commands, are easily understood; speaking to their People, in their Laws, are not so?'.⁴⁶

While this initial sketch of Locke's philosophy of language seems to justify Bejan's pessimistic reading of Locke's stance towards the regulation of hate speech through law, there is a more optimistic aspect, which leaves space for contemplating hate speech laws. When communicating our thoughts to others by words, Locke believes that there are two contexts for their use: 'philosophical use' and 'civil use'. Philosophical use of words call for great precision as it aims to express 'certain and undoubted Truths', whereas civil use of words, which is concerned with 'upholding common Conversation and Commerce, about the ordinary Affairs and Conveniencies of civil Life', require far less exactness.⁴⁷ Thus while Locke seems to be quite pessimistic with regard to the prospects of philosophical discourses, in the context of civil use he thinks that 'common use' or 'Rule of Propriety' can help to regulate the meaning of mixed

⁴⁵ Ibid., III.x.6, 493-4.

⁴⁶ Ibid., III.x.12, 496.

⁴⁷ Ibid., III.ix.3, 476.

modes ‘pretty well for common conversation’ and likewise for substances, it does ‘*well enough*’.⁴⁸ At the same time, he deplores the fact that the rules of common use ‘being no where established’ leads to the endless begetting of rival interpretations of e.g. law, as was quoted above.⁴⁹ Interestingly, in Article 73 of the aforementioned *Carolina Constitution*, it is written:

Since multiplicity of comments, as well as of laws, have great inconveniences, and serve only to obscure and perplex, all manner of comments and expositions on any part of these fundamental constitutions, or on any part of the common or statute law of Carolina, are absolutely prohibited.⁵⁰

Leaving aside whether this was Locke’s favoured solution, there is still the problem of mixed modes lacking an objective standard. But apparently Locke seems to have enough confidence in the stability of common use. Adultery and incest are both examples of mixed modes.⁵¹ Yet he obviously thought they were fit subjects for legislation: in the comments to the ‘Pennsylvania Laws’, for example, he didn’t feel it necessary to ask ‘what is meant by adultery?’ One might respond by saying that for Locke adultery was a natural law precept, and so for that reason could be apprehended simply and plainly by all. However, while the term ‘hate speech’ may not have such an established standing in the vocabulary as adultery, there is no inherent reason why it cannot rise to a stable level for common use. Even if hate speech engenders definitional difficulties, some aspects of it might be explicated in more widespread common terms. In Locke’s thought, there is therefore space to talk about hate speech legislation despite the linguistic difficulties.⁵²

⁴⁸ Ibid., III.ix.8, 479, III.ix.14, 483–4, III.ix.15, 484, III.ix.20, 488, III.xi.10, 513–4.

⁴⁹ Ibid., III.ix.8, 479. Cf. III.ii.8, 408.

⁵⁰ Locke, ‘The Fundamental Constitutions’, 176; Thomas Gustafson, *Representative Words* (Cambridge, 1992), 66.

⁵¹ Locke, *An Essay*, III.v.3, 429, III.v.5, 430, III.vi.44–51, 466–71.

⁵² Cf. Karol Janicki, *Language and Conflict* (New York and London, 2015), 78.

§5. Legislative Implications

Where then does this more nuanced, open-ended Locke fit in discussions about hate speech legislation? While it may seem as though I have been only making rather subtle distinctions so far, when these are translated into public policy, they point in a rather different direction to both Waldron's and Bejan's Locke.

Waldron's Locke adopts what we might call a 'harm identifying approach' to hate speech. He tries to identify the harm in hate speech, and on account of this harm (e.g. civic dignity), calls for its regulation. By contrast, Bejan's Locke adopts, or rather fears, what we might call a 'hate speech defining approach' to hate speech. The definition of hate speech, when left to individuals, leads to subjectivity, and on this account, creates the difficulty of limiting the scope of hate speech to those aspects or elements of it that might call for regulation. In contrast to both these approaches, the open-ended Locke adopts what we might call a 'property harm identifying approach', not to hate speech but to speech *simpliciter*. This approach involves making clear distinctions between what is or is not an injury to our civil interests, and remains indifferent to whether this injury is, or is produced by, hate speech or not. I suggest that this third approach may help to address some problems with hate speech while navigating around some of the difficulties the other two approaches might encounter. The recent Japanese debate over the introduction of a hate speech law will be helpful in illustrating this point.⁵³

Before the Japanese government legislated its hate speech law (albeit with no penalties) in June 2016, the debate over the introduction of such a law had been stalled for some time. Lawmakers of all stripes were agreed that hate speech was both deplorable and unacceptable but were divided over whether introducing a law was the right move, considering the

⁵³ Much of the following account is available through online English language newspapers such as *The Japan Times* and *The Asahi Shimbun Digital*.

consequences such a law could have on the constitutionally enshrined right to free speech.⁵⁴ One central point of controversy was what exactly hate speech was and what could be counted in it. Hate speech in Japan referred mainly to the anti-Korean demonstrations orchestrated by the chauvinist group *Zaitokukai* in areas such as Shin-Okubo, Tokyo, and Tsuruhashi, Osaka. The typical chants and slogans ranged from calling Korean residents ‘cockroaches’ or ‘maggots’ to calling their children ‘children of spies’; others went further, calling for their death—‘Kill them’ or ‘Burn yourself to death’—and even as far as calling for their extermination—‘We will carry out a massacre here’. But the expression was also used quite loosely, and one newspaper chose this expression to describe anti-government slogans such as ‘We don’t need a PM that wants war’.⁵⁵ Thus, those critical of legislation argued that because the particular harm caused by an array of speeches that was grouped into hate speech could not be established, or because less extreme speeches that were grouped into hate speech could very well be censored, a law should not be introduced. Waldron’s Locke would be confronted by the problem of trying to identify the harm shared by a range of speeches that are dubbed hate speech. Bejan’s Locke would hesitate to act given that the subjectivity in the definition of hate speech may lead to less extreme speeches being censored. Yet to leave the extremer speeches unaddressed might not seem to be an ideal state of affairs.

In this context, Locke’s ‘property harm identifying approach’ seems to offer a way around the difficulty. Rather than trying to find an all-encompassing definition of hate speech or harm caused by an array of speeches, this approach would encourage

⁵⁴ Article 21 of the Japanese Constitution declares: ‘Freedom of assembly and association as well as speech, press and all other forms of expression are guaranteed. No censorship shall be maintained, nor shall the secrecy of any means of communication be violated’.

⁵⁵ The digital version of the conservative *Sankei Shimbun*, one of Japan’s well-read national newspapers, referred to the slogans used to protest against the government’s new security bill as ‘hate speech’. Chants problematized here include ‘We don’t need a PM that wants war’, and ‘Isn’t the LDP [i.e. the ruling party] a bit nasty?’ (See 29 July 2015. <http://www.sankei.com/premium/news/150729/prm1507290004-n1.html>).

lawmakers to focus on specific speeches and explore more tailored laws to them. For example, in the case of a call for death or massacre, it may encourage us to consider a law against death threats (and if such a law exists already, to enforce it) or against public speeches glorifying and condoning genocide or mass murder. Thus this approach would be sensitive to today's reality that there are various levels of hate speeches and also various types of hate speech laws.⁵⁶ It can also avoid stalling legislative efforts due to worries concerning the precise definition of hate speech. Instead, it might recommend using terms that are more established in the language and so less ambiguous. While being sceptical of the precision and consensus about the sense and reference of words, Locke did believe that some words had a more stable relationship with senses than others at the level of common use. Thus, we might say that the idea of a threat to one's life is more commonly understood than hate speech. This would enable lawmakers to address and respond to some aspects of hate speech without using the problematic expression itself. In the interpretation of rights and laws, Locke was aware that in practice things may not always be so clear-cut, and a total resolution would only be available at God's court; but he did think that communication and law were both possible and necessary here on earth.⁵⁷

§6. Conclusion

The aim of this article was mainly corrective, making an attempt to offer a more nuanced interpretation than Waldron and Bejan's accounts of what Locke says about hate speech and how he would fit in modern discussions of hate speech legislation. Locke does not defend regulation based on civic dignity, but neither does he adopt legislative scepticism. His is an account based on harm to property. While there may be nothing outstanding about this conclusion, it does not deny Locke's relevance to and for us. We

⁵⁶ See for example, Alexander Brown, *Hate Speech Law: A Philosophical Examination* (New York and London: Routledge, 2015), esp. ch. 2.

⁵⁷ Locke, *Two Treatises*, II.240–2, 426–8; Locke, *Letter*, 48–9.

can say that the revised account of Locke is in line with and helps us to be sensitive to recent scholarly efforts emphasizing the plurality of types of regulation concerning hate speech, and to appreciate the complexity of the issue in the face of an often polarized and simplified ‘for or against’ debate over the introduction of hate speech law. Furthermore, by warning us not to be too optimistic about the prospect of a catch-all understanding of hate speech, and similarly, a successful search for a common denominator of what we categorize under the heading of hate speech, it may offer an alternative approach to hate speech—to address it without becoming embroiled in the difficulties to which, in their different ways, Waldron and Bejan draw our attention.⁵⁸

Keio University

⁵⁸ I would like to thank Dr. Bejan for kindly giving me access to her article in advance of its publication. I would also like to thank Dr. Takuya Furuta for his comments on earlier drafts of this article.