AVERTING YOUR EYES IN THE INFORMATION AGE: ONLINE HATE SPEECH AND THE CAPTIVE AUDIENCE DOCTRINE

Alexander Brown[*]

I. INTRODUCTION

There is growing interest in whether or not United States First Amendment doctrine can, and should, accommodate certain regulatory strategies for safeguarding the public from potentially harmful online hate speech, or cyberhate.[1] This article proposes significant reforms of American free speech doctrine in relation to cyberhate regulations by repurposing the captive audience doctrine. According to this doctrine, it may be permissible, even under the First Amendment, for governmental authorities to

[*]Reader in Political and Legal Theory, University of East Anglia (UEA), UK. This Article is based on the paper I presented at a workshop on global expressive rights organized by Susan Brison and Katharine Gelber at Dartmouth College in spring 2015. I am very grateful for comments and insights from all those who attended.

1. *See*, *e.g.*, DANIELLE KEATS CITRON, HATE CRIMES IN CYPBERSPACE (2014); Richard Delgado and Jean Stefancic, *Hate Speech in Cyberspace*, 49 WAKE FOREST L. REV. 319 (2014); Julian Baumrin, *Internet Hate Speech and the First Amendment, Revisited*, 37 RUTGERS COMPUTER & TECH. L.J. 223 (2011).

pass laws that abridge freedom of expression for the sake of protecting the interests of unwilling recipients of unwelcome speech. More specifically, this article examines the issue of whether or not the captive audience doctrine could be plausibly applied to circumstances in which persons are compelled by the facts of life in the Information Age to access messages and content through the Internet and the web and, subsequently, become unwilling recipients of unwelcome cyberhate. Although my arguments about online captive audiences may well have implications for other sorts of unwelcome speech, I do not intend to discuss those other types here. The primary goal is to show how the doctrine could be used to strengthen arguments for the constitutional permissibility of online hate speech regulations and to act as a bulwark against other aspects of free speech doctrine that discount such regulations, not least being the principles of content and viewpoint neutrality.

I take it as read that there are two necessary conditions for applying the captive audience doctrine. First, the speech in question is unwelcome; second, the audience is unwilling. As I shall show in Part II, hitherto, courts have tended to unpack the notion of unwanted speech in terms of interests in privacy, autonomy, tranquility, and so forth, and have given substance to the concept of unwilling audiences by appealing to the unreasonable burdens of avoiding the home, the workplace, public transit vehicles, and so forth. In both respects, courts have ably developed the doctrine without recourse to examples of online hate speech. The existing case law on the captive audience doctrine does not include any cyberhate cases. Furthermore, in no existing hate speech cases do courts consider the captive audience doctrine. In other words, as it stands, the captive audience doctrine is not dependent on cases of captive audiences to online hate speech. I do not believe that there is any need to create this dependency. However, I do wish to argue that cases of captive audiences to online hate speech should not be excluded from the possible scope of the captive audience doctrine and that there are at least some analogical reasons why they should be included. This, in turn, could lend weight to the justification for cyberhate regulations.

In particular, I shall argue that the notion of unwanted

speech is best understood as speech which harms some or other significant interest—an interest of sufficient importance as to be worthy of attention by courts. One key aim of my article is to try to reveal the nature of the significant interests involved in cases of hate speech on the Internet and the web. I shall do this by drawing an analogy with significant interests associated with the home. Moreover, I shall argue that the concept of an unwilling audience is best understood in terms of persons being unable to take practical steps to avoid the speech in question whilst, at the same time, not incurring harm to (yet further) significant interests. A second key aim of my article is to investigate the practicalities and burdens of avoiding hate speech on the Internet and the web. This includes not only a critical examination of various privacy controls, including filtering, blocking, unfollowing, and reviewing, but also an account of when it might be unfair to expect people to avoid online hate speech because, in the circumstances, doing so would involve unreasonable burdens. These burdens are to be measured, once again, in terms of harm to significant interests, only this time associated with abstinence from the Internet and the web.

The idea that the captive audience doctrine could be a powerful weapon in the arsenal of defenders of hate speech regulations is already implicit in the work of other writers on that topic.[2] Indeed, some have toyed with the idea of restricting what may even count as hate speech, so that it must be speech that is addressed to captive audiences as a necessary condition.[3] Much has been said about the appropriateness of extending the captive audience doctrine to hate speech when the latter amounts to discriminatory intimidation of people in their homes or on residential streets, for example. Thus, it has been suggested that cross burning can involve a violation of the privacy rights of captive audiences.[4] Likewise, it has been argued that Nazi

2. *See*, *e.g.*, Richard Delgado and Jean Stefancic, *Four Observations About Hate Speech*, 44 WAKE FOREST L. REV. 353, 362 (2009).

3. David Brink, *Millian Principles, Freedom of Expression, and Hate Speech*, 7 LEGAL THEORY 119, 135 (2001).

4. STEVEN HEYMAN, FREE SPEECH AND HUMAN DIGNITY, 165–66 (2008); Edward J. Eberle, *Hate Speech, Offensive Speech, and Public Discourse in America*, 29 WAKE FOREST L. REV. 1135, 1211–12 (1994).

marches may not be an attack *in* Jewish homes, but they are an attack *at* their homes and a violation of privacy rights.[5] In a similar vein, it has been claimed that to print cartoons depicting a handdrawn pig wearing a Muslim headdress with the name "Muhammad" sketched across its torso, and to post them on the external walls of the homes of Palestinians "living under the regime of belligerent occupation," is to harm a captive audience.[6] In relation to the workplace, it has been argued that employees who are subjected to gender-based and other forms of hate speech, which creates a hostile working environment, ought to be considered a captive audience.[7] Similar arguments have also been applied to students in classrooms on university campuses.[8] Others extend the doctrine to hate speech occurring in other parts of the university campus, most notably halls of residency or

5. *See*, *e.g.*, Rosalie Berger Levinson, *Targeted Hate Speech and the First Amendment: How the Supreme Court Should Have Decided Snyder*, 46 SUFFOLK U. L. REV. 45, 67 (2013).

6. Amnon Reichman, *The Passionate Expression of Hate: Constitutional Protections, Emotional Harm and Comparative Law*, 31 FORDHAM INT'L L.J. 76, 120 (2007).

7. *See*, *e.g.*, Marcy Strauss, *Sexist Speech in the Workplace*, 25 HARV. C.R.-C.L. LAW 1, 35–37, 45–46 (1990); Marcy Strauss, *Redefining the Captive Audience Doctrine*, 19 HASTINGS CONST. L.Q. 85, 89–103 (1991) [hereinafter Strauss, *Redefining the Captive Audience Doctrine*]; J. M. Balkin, *Some Realism About Pluralism: Legal Realist Approaches to the First Amendment*, 1990 DUKE L.J. 375, 423–4 (1990) [hereinafter Balkin, *Some Realism About Pluralism*]; J. M. Balkin, *Free Speech and Hostile Environments*, 99 COLUM. L. REV. 2295, 2312 (1999) [hereinafter Balkin, *Free Speech and Hostile Environments*]; Richard H. Fallon Jr., *Sexual Harassment, Content Neutrality, and the First Amendment Dog That Didn't Bark*, 1 SUP. CT. REV. 1, 43, 54 (1994); Delgado and Stefancic, *supra* note 2, at 362.

8. *See*, *e.g.*, Jack M. Battaglia, *Regulation of Hate Speech by Educational Institutions: A Proposed Policy*, 31 SANTA CLARA L. REV. 345, 376 (1991); Caroline Mala Corbin, *The First Amendment Right Against Compelled Listening*, 89 BOS. UNIV. L. REV. 939, 962–63 (2009); Melissa Weberman, *University Hate Speech Policies and the Captive Audience Doctrine*, 36 OHIO N. UNIV. L. REV. 553, 583–89 (2010). Note, however, that it has also been suggested that the captive audience doctrine only applies to classroom hate speech when uttered by a professor or teacher. *See* WOJCIECH SADURSKI, FREEDOM OF SPEECH AND ITS LIMITS 186 (1999). This more nuanced view could draw support from Martin v. Parrish, 805 F.2d 583 (5th Cir. 1986), and Bonnell v. Lorenzo, 241 F.3d 800 (6th Cir. 2001). In these cases, the courts held that the captive audience doctrine did apply to the abusive, in-class speech of university professors and teachers.

dormitories,[9] walkways, thoroughfares and corridors that students use to move between dormitories and classes,[10] and even university sporting arenas.[11] (Of course, there is no suggestion that every space on campus is apt to create captive audiences.[12]) By contrast, however, very little has been said about captive audiences for hate speech on the Internet and the web.

Online hate speech, or cyberhate, is extremely varied, of course, almost as varied as the spaces and networks which constitute the Internet and the web themselves. Not surprisingly, some writers on hate speech have already started to examine in detail the special nature of, and the particular regulatory dilemmas and challenges posed by, online as compared to offline hate speech.[13] But, once again, the specific issue of online captive

9. *See* Mari Matsuda, *Public Response to Racist Speech: Considering the Victim's Story*, 87 MICH. L. REV. 2320, 2372–73 (1989); Charles Lawrence III, *If He Hollers Let Him Go: Regulating Racist Speech on Campus*, 1990 DUKE L.J. 431, 456 (1990); Battaglia, *supra* note 8, at 376; Calvin R. Massey, *Hate Speech, Cultural Diversity, and the Foundational Paradigms of Free Expression*, 40 UCLA L. REV. 103, 177 (1992); TIMOTHY C. SHIELL, CAMPUS HATE SPEECH ON TRIAL, 110–11(2d ed., 2009); Corbin, *supra* note 8, at 963; Weberman, *supra* note 8, at 576–79.

10. *See* Lawrence, *supra* note 9, at 457; Weberman, *supra* note 8, at 579–83;SHIELL, *supra* note 9, at 110–11, 155.

11. *See* SHIELL, *supra* note 9, at 110–11. *Cf.* Gregory Matthews Jacobs, *Curbing Their Enthusiasm: A Proposal to Regulate Offensive Speech at Public University Basketball Games*, 55 CATH. UNIV. L. REV. 547, 551, 564–81 (2006).

12. Consider a case in which a student organization invites a university professor to present his controversial views on the innate differences between "the white" and "the black" races in the form of an extracurricular lecture, which students are free to attend or not attend and which takes place in a room on campus that the student organization has paid to hire out. It might be argued that the audience members are not captive 'because they can simply choose not to attend the lecture.' Charles H. Jones, *Regulating Campus Hate Speech: Is It Constitutional?*, 1992 NCCD FOCUS 1, 4 (1992).

13. *See*, *e.g.*, Alexander Tsesis, *Hate in Cyberspace: Regulating Hate Speech on the Internet*, 38 SAN DIEGO L. REV. 817 (2001); Barbara Perry and Patrik Olsson, *Cyberhate: The Globalization of Hate*, 18 INFO. & COMM. TECH. L. 185 (2009); Baumrin, *supra* note 1; CITRON, *supra* note 1; Delgado and Stefancic, *supra* note 1; RAPHAEL COHEN-ALMAGOR, CONFRONTING THE INTERNET'S DARK SIDE: MORAL AND SOCIAL RESPONSIBILITY ON THE FREE HIGHWAY (2015); Alexander Brown, *What is So Special About Online (as Compared to Offline) Hate Speech?*, ETHNICITIES (May 19, 2017), http://journals.sagepub.com/doi/abs/10.1177/1468796817709846.

audiences is noticeably absent from that body of literature.

For the purposes of framing my own discussion of online hate speech and captive audiences, I shall concentrate on six idealized examples. Although they are not intended to be exhaustive, they are meant to represent a cross-section of the sorts of styles and modes of hate speech that can be found on the Internet and the web. I believe they each involve the use of hate speech, either directly or indirectly, because they each exemplify one or more typical forms of hate speech (e.g., slur, derogation, negative stereotype, incitement to hatred, discriminatory harassment)[14] relating to members of groups or classes of persons identified by protected characteristics (e.g., race, ethnicity, nationality, citizenship, origin of birth, war record, religion, sexual orientation, gender or transgender identity, disability, age or physical appearance).[15] This is true even if some of the examples might *also* enact other types of speech acts, such as invasion of privacy, false light, libel, online stalking, and so forth.

It has also been suggested in the literature that people can constitute a captive audience only if speech is targeted *at* them, in the sense of being directly addressed to them specifically.[16] I shall not assume that position here, however. Some of my idealized examples involve targeted online hate speech, whereas some involve speech that sits in wait for any audience members who happen to find themselves in the wrong part of cyberspace at the wrong time.[17] I must also make it clear at this stage that

14. For an analysis of the concept *hate speech*, see Alexander Brown, *What is Hate Speech? Part 1: The Myth of Hate*, 36 L. & PHIL. 419 (2017); Alexander Brown, *What is Hate Speech? Part 2: Family Resemblances*, 36 L. & PHIL. 561 (2017).

15. For a discussion of the numerous characteristics that governments could potentially deem "protected" for the purposes of hate speech law, see Alexander Brown, *The "Who?" Question in the Hate Speech Debate: Part 1: Consistency, Practical, and Formal Approaches*, 29 CAN. J.L. & JURIS. 275 (2016); Alexander Brown, *The "Who?" Question in the Hate Speech Debate: Part 2: Functional and Democratic Approaches*, 30 CAN. J.L. & JURIS. 23 (2017).

16. Strauss, *Sexist Speech in the Work Place, supra* note 7, at 36–37, 43; Battaglia, *supra* note 8, 364.

17. To clarify, on my proposed reading of how online hate speech makes for captive audiences, the crucial factor is not whether the hate speaker targets the audience, but whether the hate speaker wrongfully intervenes in the audience's option-networks, meaning that the hate speaker closes down certain conjunctive

whilst I am interested in captive audiences for online hate speech, I am not solely interested in cases where the audience and the subjects of speech are one and the same. People can be captive audiences for online hate speech even if it is not about them. This may occur in cases of incitement to hatred, for example. Nevertheless, all of the examples are designed to test my hypothesis that *at least some* online hate speech can either create or exploit captive audiences.

Case 1: Tagging. A student posts on his Facebook personal page a photograph that he has taken of a group of Muslim students entering the dedicated prayer space on campus. Accompanying the picture is this comment: "You can't trust Muslims to live as peaceful members of any civilized community. You think they're meeting to pray? Think again. They're plotting attacks against us." A Facebook friend of both the student who has posted the picture and one of the Muslim students captured in the picture, tags the picture using the name of the Muslim student. Because of the tag, the picture and comment appear on the timeline of the Muslim student's Facebook profile pages. All of his Facebook friends can see the picture and the comment until he notices them and removes the tag.

Case 2: Hate Sites. Mr. Kawolski, who also happens to have arthrogryposis (a congenital joint condition that causes curving of his arm joints), has a disagreement with his neighbors over their desire to build a wall along a shared property boundary in order to keep out bears who often stray into gardens. Over time a website that was created by the local residents to share ideas and plans relating to the wall starts to also include hate speech against Mr. Kawolski, including substitutions of the name "Mr. Kawolski" with "the cripple" and "the freak," videos in which Mr. Kawolski's disability is mimicked, defamatory allegations that Mr. Kawolski is only pretending to be disabled in order to claim disability benefits from the U.S. Social Security Administration, and assertions that Mr. Kawolski's curved arm is punishment from God for his being a rapist in a previous life.

options consisting of not receiving the online hate speech and not incurring unreasonable burdens by taking practical steps to avoid receiving the online hate speech, irrespective of whether or not the speech was directly addressed to, or directed at, him or her in particular.

Mr. Kawolski needs to stay abreast of what the neighbors are planning and discussing *vis-à-vis* the building of the wall, but accessing the website now means exposure to the disablist hate speech. Similarly, local officials and other interested third parties who also need to keep on top of the local residents' plans, including individuals who might also be disabled, must run the gauntlet of disablist hate speech on the website in order to access the information they need.

Case 3: Trolling. The friends of a deceased interracial couple set up a website dedicated to remembrance and memorialization. A self-confessed black nationalist gains access to the tribute site and defaces it with slogans denouncing interracial marriage and glorifying the death of the African-American man in question. The black nationalist proclaims that the man deserved to die for "selling out" his black brothers and sisters. Friends and family wishing to participate in the online memorializing, including individuals who might also be in interracial relationships, will be exposed to the trolling when they access the website.

Case 4: Online Vandalism. Someone edits the Wikipedia page entry for Eureka, California, adding the following text to an existing section. "Child abuse scandal: In recent years two Catholic priests in our city have been exposed as child molesters. The Catholic Church did nothing to protect the children involved and did its best to cover up the scandal. The Catholic community in this city knew what was going on and kept quiet. So good people of Eureka, please, I'm begging you, don't be scared to trust your feelings about just how much you can't stand the Catholics in this city. Speak it out at the mall and at the ball game, tell them what you think of their hypocrisy, tell them that you won't stand by any more and let them rape our youngsters. Stop hiding what you feel. These Catholics deserve only our hatred. Boycott St. Bernard's High School if you have to." It is two days before another user reverses the edits.

Case 5: Copycat Parodies. A fan of a transgender model posts on YouTube a video of the model performing on a catwalk in New York, and adds a positive comment about how good she looks. Soon after, another video appears on YouTube with an almost identical title showing someone imitating the transgender

model only with exaggerated male features including a deep voice, a beard, and barely disguised male genitalia. A link to the parody video is posted in the comments section under the original video. The transgender model comes across both the original video and the parody video in the course of researching her public profile and with a view to learning what ordinary people think about transgenderism.

Case 6: Cyberharassment. Court authorities have posted online the judgments from a series of rape cases in an effort to better educate members of the public about the nature of consent. The female judge who presided over the cases received special training to do so and is well regarded by her colleagues. However, not long after the judgments are posted online, the judge becomes the subject of a high volume of misogynist and threatening comments and messages. These are posted onto the comments sections linked to searchable databases of legal judgments, web forums, online newspaper articles, and blogs, and are posted on Twitter. These are online resources and Internet messaging services that judges typically use in the course of doing independent research, keeping up-to-date with public attitudes and mores, and communicating with colleagues. A typical comment runs as follows. "You stupid bitch. You cunt whore. I know where you live."

Do any of these examples involve captive audiences? Some people would be disposed to answer in the negative because of aspects of received wisdom about the nature of the Internet and the web, and of the speech made or received therein. These aspects of received wisdom include (1) that "the Internet is not as intrusive as television or radio"[18] partly because "unlike the television, radio, or telephone message service, the Internet is not an uninvited guest,"[19] which is to say that "[c]ommunications over the Internet do not appear on computer screens without the

18. Rachel Weintraub-Reiter, *Hate Speech over the Internet: A Traditional Constitutional Analysis or a New Cyber Constitution?*, 8 BOS. PUB. INT'L L.J. 145, 165 (1998).

19. Dawn L. Johnson, *It's 1996: Do You Know Where Your Cyberkids Are? Captive Audiences and Content Regulations on the Internet*, 15 J. MARSHALL J. COMPUTER & INFO. L. 51, 94 (1996).

user taking a series of affirmative steps"[20] and that "a user seeking information must affirmatively seek out harmful content by accessing an on-line chat room discussion or bulletin board service"[21]; (2) that "in most cases, it is possible to avoid undesirable messages [sent via the Internet]"[22] in virtue of the fact that "computer-based communication services provide the user with alternative methods with which to evade offensive content"[23]; and (3) that an online audience can never truly be held captive because "the audience can always terminate the speech encounter with a simple keystroke, click of the mouse, or voice command, just as she can refuse an unwanted pamphlet or avert her eyes from offensive speech."[24] I intend to challenge these aspects of received wisdom more fully in Part IV. But for now let me simply cast doubt on the notion that the Internet and the web are not as intrusive or invasive as television or radio.

I believe this can be done on several grounds.[25] First, for companies and organizations wanting to change people's minds or influence their consumption habits, the Internet and the web are now widely considered as being no less powerful than television or radio.[26] Second, the massive volume of messages and content sent through or available on the Internet and the web, the vast number of interconnected users, and the bewildering array of methods of access, means that the Internet and the web have become ubiquitous.[27] (In a normal working day, employees

20. Weintraub-Reiter, *supra* note 18, at 165.

21. Johnson, *supra* note 19, at 94.

22. Yulia A. Timofeeva, *Hate Speech Online: Restricted or Protected? Comparison of Regulations in the United States and Germany*, 12 J. TRANSNAT'L L. & POL'Y 253, 258 (2003).

23. Johnson, *supra* note 19, at 94.

24. Noah D. Zatz, *Sidewalks in Cyberspace: Making Space for Public Forums in the Electronic Environment*, 12 HARV. J.L. & TECH. 149, 233–34 (1998).

25. *Cf.* J. M. Balkin, *Media Filters, The V-Chip, and the Foundations of Broadcast Regulation*, 45 DUKE L.J. 1131, 1136–38 (1996).

26. One indication is increasing spending on digital advertising. For example, industry researchers are predicting that United States advertisers' spending on digital advertising will outstrip spending on television advertising in 2016. *See* Tim Peterson, *Digital to Overtake TV Ad Spending in Two Years, Says Forrester*, ADVERTISING AGE (Nov. 4, 2014), https://perma.cc/JL4F-VMCE.

27. Timothy Zick, *Clouds, Cameras, and Computers: The First Amendment*

are deluged with emails, messages, and notifications through the Internet and the web.[28] Most homes contain multiple technologies that enable access to the Internet and the web. And today most people can be perpetually online, even when they leave the home, due to the development of affordable mobile devices and the widespread availability of mobile phone networks and Wi-Fi hotspots in both public and private settings.) Third, the fact that the Internet and the web have so many users means that widely viewed videos (viral videos), for example, can have a substantial impact on popular culture.[29] Fourth, the colossal number of users also creates peer pressure to use the Internet and the web to avoid being the odd one out. Fifth, the Internet and the web have a similar magnetic attraction to television or radio,[30] and many people find it difficult to keep their usage within healthy limits.[31] Indeed, cases involving Internet addiction claims are now starting to come before courts in the United States.[32] Finally, there is significant potential for users (including but not limited to children) to view or receive unwanted content, either by accident or unavoidably, in the course of undertaking routine and normal activities online.[33] (I shall say much more about this phenomenon and its different varieties below.) All of this would suggest that the Internet and the web are at least as intrusive and invasive as television or radio. Of course, some people will argue that the key respect in

and Networked Public Places, 59 FLA. L. REV. 1, 26 (2007).

28. Indeed, this traffic has been shown to impair effective thinking. *See* Julie Rennecker and Daantje Derks, *Email Overload: Fine Tuning the Research Lens*, *in* THE PSYCHOLOGY OF DIGITAL MEDIA AT WORK 14, 21 (Daantje Derks & Arnold B. Bakker eds., 2013).

29. *See*, *e.g.*, MICHAEL STRANGELOVE, WATCHING YOUTUBE: EXTRAORDINARY VIDEOS BY ORDINARY PEOPLE (2010).

30. One indication is usage of digital media. Industry researchers found that in 2013 time spent by adults in the United States with digital media surpassed time spent with television for the first time. *Mobile Continues to Steal Share of US Adults' Daily Time Spent with Media*, EMARKETER (Apr. 22, 2014), www.emarketer.com/article.aspx?R=1010782&RewroteTitle=1&nlid=8.

31. *See* Cheng Cecilia & Li Angel Yee-lam, *Internet Addiction Prevalence and Quality of (Real) Life: A Meta-Analysis of 31 Nations Across Seven World Regions*, 17 CYBERPSYCHOLOGY, BEHAV., & SOC. NETWORKING 755 (2014).

32. *See* Pacenza v. IBM Corp., 363 F. App'x 128 (2d Cir. 2010).

33. *See* Baumrin, *supra* note 1, at 258.

which the Internet and the web are *not* intrusive or invasive is that they can be easily avoided, but I also intend to directly challenge this argument in Part IV.

In what follows I will argue that *at least some* of the aforementioned idealized examples *do* involve captive audiences. As a preliminary step, I want to briefly reflect on a recent article by John B. Major in which he addresses the related issue of whether or not cyberstalkers can create captive audiences.[34] Although this issue is both broader and narrower than the issue of captive audiences for hate speech on the Internet and the web, it nevertheless points in the direction of some potentially important questions for my own investigation. To be held captive is to be held somewhere, even if the sense of captivity being invoked is more metaphorical or figurative than literal.[35] In other words, the captive audience doctrine makes little sense unless a location can be specified. At the very least, there is a difference between saying that social networking services, like Twitter, can be used as instruments for sending messages directly into people's homes wherein they may be captive audiences, and saying that these services can create virtual locations or areas of cyberspace wherein people may be captive audiences. At one stage, Major claims that "when an individual is made *captive to speech on the Internet*, the captive audience doctrine can, and should, apply."[36] One way of reading this claim is to say that the location of captivity is the Internet itself. But this raises some potentially thorny issues that are not fully explored by Major. First, *who* exactly can be held captive on the Internet? Is it persons themselves, their online personae, or just their avatars? Second, in *what* (figurative) sense of captivity can it be meaningfully said that a person is being held captive on the Internet? Is the sense of captivity in which it can be said that someone is held captive *on* the Internet the same or different from the sense of captivity relevant to being held captive *in* the

34. John B. Major, *Cyberstalking, Twitter, and the Captive Audience: A First Amendment Analysis of 18 U.S.C. § 2261a(2)*, 86 S. CAL. L. REV. 117 (2012).

35. For more on the metaphorical aspects of the captive audience doctrine, see A. BOSMAJIAN HAIG, METAPHOR AND REASON IN JUDICIAL OPINIONS (1992).

36. Major, *supra* note 34, at 149 n.231 (emphasis added).

home? Third, *where* are people held captive? Is it on the Internet or can people also be held captive in places or spaces that exist on the Internet such as places and spaces on the web? These sorts of questions simply cannot be avoided when one is seeking, as I am here, to apply the captive audience doctrine to the Internet and the web.

Finally, before I begin, I want to make it clear at the outset that I am proposing a substantial change to American free speech doctrine. Presently, in the United States, there is no constitutionally proscribable category of speech called "hate speech"; indeed, some of the existing case law suggests that hate speech is a protected category. At the same time, the case law narrowly defines what counts as a "captive audience" that would justify regulation of otherwise nonregulatable speech. Specifically, the captive audience doctrine does not, as it stands, cover audiences who are, in a sense, held captive online. What is more, it is fair to say that past cases involving captive audiences have tended to involve content and viewpoint neutral restrictions on speech, whereas hate speech restrictions are typically content-based and are sometimes even viewpoint-based. Yet, under American free speech doctrine, content- and viewpoint-based restrictions on speech are more often than not constitutionally impermissible, albeit there are notable exceptions.[37] Indeed, arguably, hatemongers quickly took up the Internet and the web as means of disseminating their invective, not merely because of the ease of use and inexpensiveness of online communication, but also because they gambled that the expression of online hate speech would be as protected as offline hate speech under recent interpretations of the First Amendment.[38] However, I want to emphasize right at the start that I am not seeking here to address the question: What *is* the captive audience doctrine in the United States? But instead: What *should* the captive

37. *See*, *e.g.*, Steven H. Shiffrin, *Racist Speech, Outsider Jurisprudence, and the Meaning of America*, 80 CORNELL L. REV. 43, 50–64 (1994); Steven Heyman, *Spheres of Autonomy: Reforming the Content Neutrality Doctrine in First Amendment Jurisprudence*, 10 WM. & MARY BILL RTS. J. 647, 651–52 (2002).

38. *See* Brian Levin, *Cyberhate: A Legal and Historical Analysis of Extremists' Use of Computer Networks in America*, 45 AM. BEH. SCI. 958 (2002).

audience doctrine be? And not: What is the current doctrinal position on the constitutionality of content-based hate speech regulations in the United States? But rather: What ought to be the doctrinal position on such regulations?

The remainder of the article is structured as follows. Part II sets out the captive audience doctrine as it has been developed by courts in the United States, including its hitherto rare application to cases involving hate speech. Part III attempts to uncover some of the significant interests that might be harmed by online hate speech, based on analogies with the home. Here, I highlight three types of interests: privacy, autonomy, and tranquility. Part IV examines what it means to "avert your eyes," that is, to avoid unwelcome speech, in the Information Age. It has been assumed in some quarters that people's best defense against the threat of captivity posed by online technologies rests in those technologies themselves—not least in filtering, blocking, unfollowing, and reviewing controls. I shall argue that we have little reason to be sanguine about the ability of software to filter out hate speech without either overfiltering or underfiltering content, and that unfollowing and reviewing are only meaningful ways of avoiding unwelcome speech if Internet companies change the default settings so that these advanced privacy controls are automatically enabled. In addition, I explore the burdens associated with abstaining from the Internet and the web, either partly or entirely. I shall argue that given how important the Internet and the web have become to normal human functioning—as private persons, workers, consumers, citizens, and so on—it is now unreasonable to expect people to live their lives offline. It is the fact that receiving online hate speech may damage one or more significant interests and cannot be avoided without sacrificing (yet further) significant interests that makes talk of captive audiences *prima facie* plausible. Finally, Part V explores the implications of these arguments (that online hate speech can create or exploit captive audiences), including implications for whether or not courts and/or regulators should apply the First Amendment doctrine of content and viewpoint neutrality and/or the regulatory principle of net neutrality to laws/codes/regulations that restrict the use of online hate speech.

II. BACKGROUND CASE LAW

For more than half a century the U.S. Supreme Court has recognized as germane to First Amendment cases the issue of whether or not an audience is being held captive to speech, in the sense of being "practically helpless" to avoid or escape it.[39] The Court has also recognized that persons may be captive audiences even if the speaker is located in a traditional public forum, such as a public street or sidewalk. In other words, in order to reach a finding that persons are a captive audience it is unnecessary, doctrinally much less extradoctrinally, to redesignate traditional public forums as nonpublic forums.[40]

At the same time, however, courts in the United States have, in effect, narrowly circumscribed the application of the captive audience doctrine. In relation to the Internet, the following cases stand out. In *ACLU v. Reno*[41] and then later in *Reno v. ACLU*,[42] courts considered whether or not the application of the Communications Decency Act (CDA) to indecent materials received via the Internet was constitutional *vis-à-vis* the First Amendment. The courts eschewed the suggestion that people (children) could be held as captive audiences on the Internet on the grounds that "[u]sers seldom encounter content 'by accident'"[43] and that "the Internet is not as 'invasive' as radio or

39. Kovacs v. Cooper, 336 U.S. 77, 86–87 (1949).

40. I have in mind the majority decision in Frisby v. Schultz, 487 U.S. 474 (1988), where the majority held that "Our prior holdings make clear that a public street does not lose its status as a traditional public forum simply because it runs through a residential neighbourhood." *Id.* at 480. And that "No particularized inquiry into the precise nature of a specific street is necessary; all public streets are held in the public trust and are properly considered traditional public fora." *Id.* at 481. It went on to argue that the focused picketing of a house in a residential street, picketing located on the street but the noise of which carrying into the house, creates a captive audience. *Id.* at 487–48. And that the ban on residential picketing at issue this case served the important interest of protecting residential privacy. *Id.* at 479–80. This case stands in contrast to Lehman v. City of Shaker Heights, 418 U.S. 298 (1974), in which the majority did rely on its view that a city transit system is not a First Amendment forum in order to support its position that passengers are a captive audience. I thank Sonu Bedi for pressing me on these issues.

41. 929 F. Supp. 824 (E.D. Pa. 1996).

42. 521 U.S. 844 (1997).

43. 929 F. Supp. at 844.

television."[44] Similarly, in *United States v. Cassidy*[45] a U.S. District Court (D. Md.) made no mention whatsoever of privacy interests or of the captive audience doctrine, only of the emotional distress caused by the actions of a cyberstalker.

Likewise, in cases involving hate speech, more often than not, courts have explicitly rejected, declined to fully consider, and in some cases even mention, the potential application of the captive audience doctrine. In *Collin v. Smith*,[46] for example, a U.S. Court of Appeals (7th Cir.) upheld a lower court's ruling that a set of municipal ordinances designed to curb the intended marches of Frank Collin and other members of the National Socialist Party of America (NSPA) in the predominantly Jewish village of Skokie, Illinois were unconstitutional. According to the Court, "[t]here need be no captive audience, as Village residents may, if they wish, simply avoid the Village Hall for thirty minutes on a Sunday afternoon."[47] To take another example, the argument that student-on-student hateful abuse is unprotected speech under the First Amendment because students are a captive audience was put forward by the University of Michigan's defense team in *Doe v. University of Michigan*[48] (This argument appeared in the defense brief against the plaintiff's motion for preliminary injunctive relief against the University of Michigan's Policy on Discrimination and Discriminatory Harassment of Students in the University Environment.[49]) Judge Cohn consolidated the hearing on the motion with the trial and ultimately found no reason even to mention the captive audience doctrine in his opinion. In *R.A.V. v. City of St. Paul*,[50] a U.S. Supreme Court case concerning the burning of a cross in front of the home of an African American family, a unanimous Court struck down the City of St. Paul's Bias-Motivated Crime Ordinance. It held that since the Minnesota Supreme Court

44. 521 U.S. at 869.

45. 814 F. Supp. 2d 574 (D. Md. 2011).

46. 578 F.2d 1197 (7th Cir. 1978).

47. *Id*. at 1207.

48. 721 F. Supp. 852 (E.D. Mich. 1989).

49. *Id*. Defendant's Brief in Opposition to Plaintiff's Motion for Preliminary Injunctive Relief, at 13.

50. 505 U.S. 377 (1992).

construed the St. Paul ordinance as being limited to conduct which constituted "fighting words", the issue of the captive audience is "not before us in this case".[51] In *Dambrot v. Central Michigan University*,[52] a case involving a college basketball team coach's repeated use of the word "nigger" in team talks used to gee up the players, a U.S. Court of Appeals (6th Cir.) made no explicit mention of the captive audience doctrine. Nor did the U.S. Supreme Court make any direct reference to the doctrine in *Virginia v. Black*,[53] a case involving the burning of a cross in an open field in plain sight of a state highway and several residential homes.

I believe that courts can, and should, be more willing to apply the captive audience doctrine to cases of online hate speech or cyberhate. But just to be clear, I do *not* intend to make the unnecessary and implausible argument that whether government authorities may regulate otherwise unregulatable speech on account of the presence of captive audiences should also depend on the presence of online hate speech. That is, I shall not argue that the captive audience doctrine should be restricted to instances of cyberhate going forward. Clearly there have been in the past, and will continue to be in the future, many useful applications of the doctrine to circumstances that do not involve cyberhate. Rather, as I have already stated, my purpose is to argue simply that cases involving cyberhate should not be excluded from the possible scope of the captive audience doctrine and that there are at least some analogical reasons why they should be included.

III. IDENTIFYING INTERESTS

In order to be classed as a captive audience it is necessary that the relevant speech is unwelcome, meaning that the speech harms some or other significant interest. But how should one go about identifying the interests being harmed in potential captive audience cases? One method is to reflect on the very concept of a

51. *Id*. at 414.
52. 55 F.3d 1177 (6th Cir. 1995).
53. 538 U.S. 343 (2003).

captive audience and to identify a single core interest which is harmed in all meaningful instances of that concept. In 1953, for example, Charles L. Black Jr. suggested that the relevant interest is "a very old freedom, a freedom to which, in some sense, all the others are dedicated handmaidens—the freedom of the mind."[54] Since then, however, the captive audience doctrine has been applied to various different contexts and situations,[55] some of which may not have anything essentially to do with freedom of the mind, if that means being left in peace to contemplate the objects of some or other belief or disbelief. Sometimes the interest harmed by speech is of the non-cognitive or emotional variety; in other contexts or situations it is an interest in exercising a capacity for choice that is nonidentical with contemplating and feeling. So it may be extremely difficult to identify a single core interest that is implicated in all cases to which courts have, or might in the future, apply the captive audience doctrine.

A second method begins with a set of exemplar cases involving putative captive audiences and builds up from these exemplars a working list of significant interests. This seems to have been the method employed by the U.S. Supreme Court in *Cohen v. California*.[56] Highlighting *Rowan v. United States Post Office Department*[57] as an exemplar captive audience case, the Court reaffirmed its position that captive audiences in the home suffer an invasion of privacy.[58] But the Court also made it clear that people can be captive audiences even if they are outside the sanctuary of the home and, what is more, set down a test for determining when the doctrine can be applied outside of the home: namely, when "substantial privacy interests are being invaded in an essentially intolerable manner."[59] In short, the Court identified the interest at stake in the exemplar case and

54. Charles L. Black Jr., *He Cannot Choose but Hear: The Plight of the Captive Auditor*, 53 COLUM. L. REV. 960, 966 (1953).

55. *Cf.* Balkin, *Free Speech and Hostile Environments*, *supra* note 7, at 2312.

56. 403 U.S. 15 (1971).

57. 397 U.S. 728 (1970).

58. 403 U.S. at 21.

59. *Id.*

then utilized that interest as a basis for a more general test.[60] Employing a similar method, but starting with a larger set of exemplars, Marcy Strauss proffers the following list of rights which she believes are implicated in captive audience cases: the right to choose what information one receives, the right to repose, and the right to be free from offense.[61]

A third method of identifying significant interests at stake in potential captive audience cases is contextualism. This method involves reflecting closely on the nature of the contexts or situations in which persons might be said to be captive audiences and then trying to determine which interests are most naturally associated with or germane to those contexts. According to Richard H. Fallon Jr., for example, "[w]hen the so-called captive audience cases are read in conjunction, the character of the place seems more important than the degree of audience 'captivity' in explaining the applications of captive audience doctrine."[62] Although she is not explicit, Caroline Mala Corbin also seems to employ this method. She takes different contexts in turn—the home, the workplace, the polling station—and in each context sets forth the most germane interest which is being harmed—the right to privacy, the right to equal protection, the right to vote.[63]

My own hunch is that the best approach combines the second and third methods: that each would be incomplete without the other. On the one hand, in order to understand, using the second method, whether or not analogies can be drawn between exemplar cases of captive audiences in offline and online contexts, it will be necessary to engage in contextual analyses, using the third method, of the nature and function of different parts of the Internet and web. On the other hand, the third method is likely to produce plausible results only when combined with the second. For one thing, I think it would be mistaken to

60. This is a highly abstract test, of course, almost as abstract as the captive audience doctrine itself. For an attempt to inject greater precision into this test, however, see Geoffrey R. Stone, *Fora Americana: Speech in Public Places*, 1974 SUP. CT. REV. 233, 263–64 (1974).

61. Strauss, *Redefining the Captive Audience Doctrine*, *supra* note 7, at 106–16.

62. Fallon, *supra* note 7, at 18.

63. Corbin, *supra* note 8, at 951–65.

assume that the home, the workplace, the university, the cemetery chapel, crematorium, graveside, or anywhere that funeral services and death rituals are performed, for instance, must necessarily implicate very different interests. Not all relevant differences between contexts or situations equate to differences in the genera of interests at stake. Furthermore, it seems to me that courts in the United States have been willing to extend the captive audience doctrine to contexts other than the home only to the extent that analogies can be drawn between those other contexts and that exemplar. And part of what vindicates these analogies is the similarity of the *interests* at stake, and not merely the context.

For reasons of space, in what follows I shall investigate only two types of contexts. The first is the home, a context to which traditionally the captive audience doctrine has been applied; and the second is the Internet and the web, a family of contexts to which the application of the captive audience doctrine remains highly controversial. I shall argue that different species of the same genera of significant interests can be found in both types of contexts.

A. The Home

As touched upon already, one species of interest implicated in potential captive audience cases involving the home is *privacy in the home*. This interest reflects the basic idea that the home (and perhaps, also, residential streets in which homes are located) is a place where the right to avoid intrusion has particular force.[64] This interest may even rest on a conceptualization of the home as, by definition, the sort of place in which people have a right to be protected from unwanted speech, unlike *some* public spaces.[65]

A second significant interest is *autonomy in the home*. This interest can be divided into two types: *substantive* and *formal*. A substantive interest in autonomy in the home means, to borrow the words of Zechariah Chafee, that "home is one place where a

64. *See* 397 U.S. at 738; Frisby v. Schultz, 487 U.S. 474, 484–85 (1988).

65. *See, e.g.*, LORNA FOX O'MAHONY, CONCEPTUALISING HOME: THEORIES, LAWS AND POLICIES (2007).

man ought to be able to shut himself up in his own ideas if he desires."[66] Protecting a substantive interest in autonomy in the home means protecting the home as a place for personal reflection or quiet contemplation about what to believe and even how to behave.[67] Respecting a formal interest in autonomy in the home, by contrast, has to do with respecting a person's right to choose for him or herself what information or messages he or she will receive within the home.[68] In other words, what matters is the role played by the agent in exercising his or her right to protection from unwanted speech. Here, the thought is that it is not ideal for governmental institutions to instigate the blocking of unwanted mail, even if in the name of protecting recipients' substantive autonomy; it is better if the recipient exercises his or her own autonomy in taking steps to block mail using governmental institutions merely as instruments. Indeed, the U.S. Supreme Court has been more willing to uphold statutes or ordinances that place the onus on the recipient of unwanted mail to exercise his or her autonomy in taking action to stop the mail by informing the postal service that the mail is unwanted, rather than those that give the postal service the power to make a rebuttable presumption that certain mail is unwanted (on behalf of citizens).[69]

A third significant interest is *tranquility in the home*. This interest relies on an understanding of the home as a sanctuary; a refuge from the stresses and strains of working life or from the cacophony of the public square; a place of serenity and repose, rest and recuperation, repair and replenishment.[70] If the home is

66. ZECHARIAH CHAFEE, FREE SPEECH IN THE UNITED STATES 406 (1954).

67. *See* Kovacs v. Cooper, 336 U.S. 77, 97 (1949).

68. *See* Strauss, *Redefining the Captive Audience Doctrine*, *supra* note 7, at 109.

69. *Compare* Rowan, 397 U.S. 728 *with* Bolger v. Youngs Drug Prods., 463 U.S. 72 (1983). *See also* Troyer v. Town of Babylon, 483 F. Supp. 1135, 1138 (E.D.N.Y. 1980). The same point is made in Strauss, *Redefining the Captive Audience Doctrine*, *supra* note 7, at 109.

70. *See, e.g.*, Martin v. City of Struthers, 319 U.S. 141, 144, 152–53 (1943); City of Wauwatosa v. King, 49 Wis. 2d 398, 404–07 (1971); Frisby v. Schultz, 487 U.S. 474, 484 (1988); Olmer v. City of Lincoln, 192 F.3d 1176, 1185 (8th Cir. 1999).

"the last citadel of the tired, the weary, and the sick,"[71] then (it has been argued) the burning of a cross in front of the home of an African American family, for example, must surely constitute a most serious breach of the citadel walls.[72] How could a family be expected to enjoy relaxation and restful sleep amid the sights, sounds, and smells of a fiery cross?

Let us suppose for the sake of argument that the aforementioned interests are at stake in most cases of captive audiences in the home. Could not exactly the same interests be involved in each of the cases listed in Part I involving the Internet and the web? After all, "[t]he Internet, like broadcasting, can also be characterized as a medium that intrudes into the home."[73] So, for example, "Tweets do in a sense 'enter the home' as they come up on the computer screen of the victim."[74] Indeed, the fact that someone can be physically located within their own home but at the same time send and receive content which is virtually located within online public spaces may challenge traditional distinctions between public and private spaces.[75]

However, we must also be sensitive to the fact that with the increasing sophistication and prevalence of mobile devices (e.g., smartphones, tablets, laptops, smartwatches) people are increasingly gaining access to the Internet and the web outside of their homes. And this is also likely to change how we think about the distinction between public and private spaces,[76] as well as how we understand the scope of the captive audience doctrine and how we define the interests at stake. At the very least, it raises the question of whether or not it can be meaningful to say, even if figuratively, that an individual is made captive to speech *on* the Internet or *in* places and spaces found on the web. Putting this another way, if online hate speech is capable of reaching people not merely at home but also at work, in cars, on public

71. Gregory v. City of Chicago, 394 U.S. 111, 125 (1969).

72. *See* Eberle, *supra* note 4, at 1191–92.

73. William D. Araiza, *Captive Audiences, Children and the Internet*, 41 BRANDEIS L.J. 397, 398 (2003).

74. Major, *supra* note 34, at 153.

75. *See* Balkin, *Free Speech and Hostile Environments*, *supra* note 7, at 2311–12.

76. *Cf.* Zick, *supra* note 27, at 26–30.

transit vehicles, in public spaces, in restaurants, or quite simply wherever people take their mobile devices, then we at least ought to investigate if there is a plausible sense in which people can be captive audiences on the Internet or in places and spaces found on the web over and above the sense in which they may be held captive audiences in each of the aforementioned physical spaces. This further investigation, I believe, must involve a proper account of the significant interests that may be harmed by online hate speech—interests of a sort that can be relevant to the issue of whether or not the captive audience doctrine is applicable to the Internet and the web.

B. The Internet and the Web

In this subpart I shall argue that online hate speech can harm three significant interests which although are different species to the three interests outlined in the previous subpart, are nonetheless species belonging to the same genera.

The first is *privacy on the Internet and the web*. It is tempting to think that when accessing the Internet or the web one can no longer claim a right to privacy for the simple reason that one is choosing to let the outside world in. But I believe that this aspect of received wisdom about the Internet and the web is problematic. For one thing, users of the Internet and the web are able to create their own *virtual homes* within multi-user domains (MUDS), massively multiplayer online worlds (MMOWs), massively multiplayer online role-playing games (MMORPGs), and multi-user virtual environments (MUVEs). And the existence of these virtual homes surely opens up the possibility of violations of the privacy rights of avatars residing inside virtual homes that are located within larger virtual worlds.[77]

In addition to this, it seems to me that social networking profile pages, including timelines, could be viewed as close cousins of virtual homes, wherein the privacy rights of "real" people may also be threatened.[78] Now it is certainly true that the

77. *See, e.g.*, GREG LASTOWKA, VIRTUAL JUSTICE: THE NEW LAWS OF ONLINE WORLDS (2010). *Cf.* Darren MacLennan & Jason Sartin, *Review of F.A.T.A.L.*, RPG (Oct. 30, 2009), www.rpg.net/reviews/archive/14/14567.phtml.

78. *Cf.* Susanna Paasonen, *Immaterial Homes, Personal Spaces and the*

profile pages which users have the opportunity to create on mainstream social networking websites, like Facebook, Myspace, and Qzone, are quite different from virtual homes in that they do not contain virtual walls, ceilings, doors, furniture, and so on.[79] Nevertheless, these websites offer users a chance to create what might be called "online homes." These online homes (profile pages) share similar features with real homes. They are spaces that users can build themselves, the structure and form of which they have some scope to control and customize to look a certain way. They are spaces that users can fill with personalized meaningful objects, such as pictures, videos, and texts, so as to give the feeling of homeliness. They are spaces into which users can invite their friends and family, and in which a great deal of personal time can be spent relaxing. So there is a sense in which when people are spending time *in* their profile pages they are spending time "at home." Indeed, part of the lure of social networking websites for students studying overseas, for example, is that they can create a home away from home that exists online.[80] Moreover, because mobile devices enable people to access the Internet and the web wherever they go, leaving one's real or offline home is no longer a barrier to spending time in one's online home. No doubt when people carry their mobile devices with them this can make wherever they go feel more homely because of the devices *themselves*, akin to a child carrying around a much-loved teddy bear. But there is also a sense, I think, in which mobile devices enable people to take their online homes *with them*.[81] The important point here is that the existence of online homes might actually increase rather than decrease the extent to which people have privacy interests when

Internet as a Rhetorical Terrain, *in* HOMES IN TRANSFORMATION: DWELLING, MOVING, BELONGING (Hanna Johansson & Kirsi Saarikangas eds., 2009).

79. Interestingly, users are able to create virtual homes on more specialist social networking websites. Consider the Guest Rooms function on Habbo.

80. Of course, profile pages, like other spaces on the Web, are materially embodied by servers, cables, satellite transmitters, electromagnetic radio waves, receivers, computer screens, mobile device displays, pixels, and so on. But it is not these that I call home. The online home (profile page) merely supervenes upon these physical objects.

81. *See also* DAVID MORLEY, MEDIA, MODERNITY, AND TECHNOLOGY: THE GEOGRAPHY OF THE NEW 205 (2007).

they access the Internet or the web.

Taking the idea of online homes to its logical conclusion, I would argue that the act of Facebook tagging described in Case 1: Tagging, constitutes an intrusion into a place of privacy for the Muslim student. When hate speech finds its way onto people's social networking profile pages, and therefore into their online homes as I am calling them, they may experience a sense of invasion of privacy which is not dissimilar to the feeling they might experience if someone broke into their real home and painted hate messages on their walls and mirrors. It may be the case that the student can choose to remove the tag once he notices it, and thereby put a stop to the intrusion. But this does not prevent the intrusion itself. Unsurprisingly, then, in response to user requests for greater privacy in respect of tagging, in 2011 Facebook introduced a new Advanced Privacy Control which allows users to review tags that Friends add to their profile page timelines before they appear. Once this control is enabled, users can approve or reject any photo or post in which they are tagged before it becomes visible to anyone else on their profile page timelines.[82]

The interest in privacy on the Internet and the web need not be exclusively analogized to privacy in the home. Another analogy worth pursuing is with privacy in the context or situation of a funeral or death ritual. It is widely supposed that funeral mourners ought to be left alone to mourn in peace, without being distracted by unwanted speech.[83] Thus, courts in the United States have ruled in several cases that funeral picketing by the Westboro Baptist Church has either created or exploited captive audiences.[84] Arguably, a similar privacy

82. Although the reviewing control allows someone to prevent a tag from appearing on her own profile page Timeline, the tag will still remain on the tagger's Timeline and can be viewed by anyone who is friends with the tagger, including mutual friends of the taggee. So there is also potentially a false light privacy interest in not allowing other people to tag pictures that place oneself in a false, misleading, or distorted light before ones Facebook friends.

83. Levinson, *supra* note 5, at 67–69.

84. *See, e.g.,* McQueary v. Stumbo, 453 F. Supp. 2d 975, 992 (E.D. Ky. 2006); Phelps-Roper v. Nixon, 504 F. Supp. 2d 691, 696 (W.D. Mo. 2007); Phelps-Roper v. Taft, 523 F. Supp. 2d 612, 618–19 (N.D. Ohio 2007); Phelps-Roper v. Strickland, 539 F.3d 356, 362–72 (6th Cir. 2008).

interest is at stake in the set of circumstances found in Case 3: Trolling. In this case, the friends of a deceased couple are not left alone to partake in processes of online grieving, remembrance, and memorialization. Instead, their privacy is invaded by people who wish to hijack the virtual space to espouse and promote their ideology of racial difference, separation, and segregation. The mere fact that the processes of grieving, remembrance, and memorialization that are being intruded upon take place over the Internet and on the web as opposed to at a cemetery chapel is an immaterial difference. If people want to take part in funerals, they must go to the places designated for that event by common agreement or convention. If people want to take part in online memorials, they too must visit the websites set up for such activities by conventional practice.

A second significant interest is *autonomy in using the Internet and the web*. Of course, there are different senses in which cyberhate can affront the autonomy of others. A good deal of online hate speech takes the form of *denying* that a certain group of people are even capable of autonomously choosing for themselves how to think and how to behave.[85] However, I am more interested in the ways in which some online hate speech can harm substantial interests in autonomy of audiences, by undermining, subverting, or circumventing their normal processes of rational reflection.[86] Once again, this argument can be applied to different types of audiences of online hate speech, including not only people who are the subject matter of cyberhate speech but also people who receive cyberhate but who are not its subject matter. Consider Case 4: Online Vandalism. Here the hate speaker is exploiting an online encyclopedia, Wikipedia, in order to influence and persuade the audience to react to Catholics as he would like them to. He is attempting to change attitudes, to stir up hatred, and even to incite acts of discrimination through both the power of his rhetoric and the power of the brand,

85. *See* Heyman, *supra* note 4, at 175.

86. *See, e.g.*, Susan J. Brison, *The Autonomy Defense of Free Speech*, 108 ETHICS 312, 328 (1998); Brink, *supra* note 3, at 138–40; Seana Shiffrin, *Reply to Critics*, 27 CONST. COMMENT. 417, 437 (2011). For an application of this idea to online hate speech, see Tsesis, *supra* note 13, at 870; ALEXANDER BROWN, HATE SPEECH LAW: A PHILOSOPHICAL EXAMINATION 63 (2015).

Wikipedia. The strategy is partly to reach a large audience but also partly to make that audience believe that the ideas have some credibility or standing; the fact that they are posted on a well-known, and to some extent trusted, website could make readers give the ideas more credence than if they had stumbled across them somewhere else on the web. The question is: should people be free to access online encyclopedias without being confronted with such forms of unwanted influence and persuasion?

No doubt some people will argue that the moment individuals go online they can no longer reasonably expect to shut themselves up in their own ideas. The web is a public space, or is partly composed of public spaces, after all. What is more, these are spaces to some extent defined by attempts at influence and persuasion. To expect the web to provide opportunities for pristine reflection—akin to somebody's study or drawing room—is to fundamentally misconceive what the web is or should be (so the objection runs). But whilst it is certainly true that parts of the web do serve the purposes that the hate speaker seeks, it would be overly simplistic to assume that the same holds for *all* parts of the web. Surely the primary function of online encyclopedias is to provide user-generated sources of reliable information: to create, organize, and provide quick access to vast numbers of entries which are capable of acting as reference works for people wishing to discover or confirm facts or as concise summaries of bodies of knowledge which people can easily digest and utilize as starting points for further research. Users of online encyclopedias will often treat the information they find as source material for their reflections about what to believe. I would suggest that such information-gathering sessions might be among the situations when we have a duty to let other people alone to reflect about what to believe. Rules against online vandalism may be on a par with library rules requiring patrons not to write in or add handwritten annotations to library books (so-called library-book vandalism) so as not to distract or disturb the information-gathering of other patrons.

Third, I believe that some online hate speech can harm a significant interest in *emotional tranquility on the Internet and the web*. The general idea that hate speech can harm emotional

tranquility is not new of course. In the 1980s, Richard Delgado and Mari Matsuda pointed to the immediate or short-term severe emotional distress that may be suffered by individuals as a result of experiencing interpersonal racist abuse.[87] More recently, Eric Barendt has suggested that "[t]he best argument for restricting racist hate speech is undoubtedly that a state has a compelling interest to protect members of target groups against the psychological injuries inflicted by the most pernicious forms of extremist hate speech."[88] Steven Heyman highlights the harm that may be done to the victim's emotional tranquility by cross burning and Nazi marches, which he calls a violation of the right to personality.[89] Elsewhere I have argued for a retheorization of certain torts and delicts in cases of targeted hate speech as involving violations of dignity through degradation and humiliation, which include subjective as well as objective dimensions.[90] I wish to extend these lines of reasoning to hate speech on the Internet and the web. I want to focus on types of online hate speech that are capable of causing intense feelings of fear, anger, resentment, shame, embarrassment, humiliation, self-loathing, distress, loneliness, and exclusion.

It does not take much of a leap of imagination to see how the people on the receiving end of the online hate speech in Case 2: Hate Sites, Case 5: Copycat Parodies, and Case 6: Cyberharassment, could suffer damage to their emotional tranquility as a consequence. In Case 2: Hate Sites, for example, the father who comes across a website dedicated to tarnishing the memory of his deceased son and reviling him personally cannot remember his son or think about himself as a father without a significant degree of mental anguish—because he cannot separate his memories and thoughts about his son from

87. *See* Richard Delgado, *Words That Wound: A Tort Action for Racial Insults, Epithets, and Name-Calling*, 17 HARV. C.R.-C.L. L. REV. 133, 137 (1982); Matsuda, *supra* note 9, at 2336.

88. ERIC BARENDT, FREEDOM OF SPEECH 174–175 (2d ed., 2005).

89. Heyman, *supra* note 4, at 165–70.

90. *See* Alexander Brown, *Retheorizing Actionable Injuries in Civil Lawsuits Involving Targeted Hate Speech: Hate Speech as Degradation and Humiliation*, 9 ALA. C.R. & C.L. L. REV. 1.

his recollection of the words on the hate site.[91] Does it matter that his exposure to the hate speech was short-lived? Could there be a relevant difference between semipermanent hate speech that is painted onto neighborhood walls and which someone is forced to walk past repeatedly over a period of weeks or months, and hate speech that temporarily appears in an email or on a website and which someone sees only for a fleeting moment and then never again?[92] Even if some people would suffer severe emotional distress only as a result of accumulated exposure to hate speech over a prolonged period, it is still possible that other people could suffer lasting damage to their emotional tranquility even after a limited exposure to cyberhate. If so, then this may undermine an argument that says recipients of online hate speech are not captive audiences simply because they can delete emails and refrain from revisiting hate sites. I shall come back to this issue in Part IV.

In order to fully understand the special ways in which online hate speech may pose a threat to emotional tranquility, it is necessary to reflect further on the nature of the Internet and the web. It has been argued by critical race theorists in the past that racist hate speech "can cause mental, emotional, or even physical harm to their target, especially if delivered in front of others."[93] But what if the number of others is potentially vast? Social networking websites (e.g., Facebook, Myspace, Qzone), blogging platforms and instant messaging services (e.g., Twitter, WordPress, Tumblr), photo and video sharing websites (e.g., YouTube, Vimeo, Flickr, and Instagram), and Internet forums and message boards (e.g., 4chan) enable people to cultivate and enjoy *supersociability*. Supersociability has partly to do with the fact that ordinary people can post opinions and information, exchange messages, be friends, and share pictures and videos with more people than at any time in human history. But this

91. The hate speech qualifies as such not merely by its tone and content but also by virtue of picking out a protected group, namely, parents of United States military personnel.

92. *Cf.* Jeremy Waldron, *2009 Oliver Wendell Holmes Lectures, Dignity and Defamation: The Visibility of Hate*, 123 HARV. L. REV. 1596, 1600–01, 1624 (2010).

93. *Cf.* Delgado, *supra* note 87, at 143.

supersociability can be a double-edged sword. An awareness that comments, messages, images, and videos are capable of being seen by a large number of strangers on the Internet and the web, as well as by countless friends and family members, could potentially increase the inherent risk, raise the intensity, or even transform the quality of feelings of fear, anger, resentment, degredation, shame, embarrassment, humiliation, self-loathing, or loneliness which may come about as a consequence of exposure to hate speech.[94] This may be an important dimension of Case 1: Tagging, where the tagged picture and comment is visible to all of the student's Facebook friends.

The supersociability made possible by the Internet and the web might also be responsible for amplifying a deep-seated human need to be esteemed and for exacerbating the emotional toll of not being so. This can be seen most vividly in Case 5: Copycat Parodies. *Part* of what is at stake in this case is something that explains why many people are attracted to the Internet and the web in the first place: the impulse to cultivate and enjoy *likeability*. It matters to people whether or not they are popular among online communities—and it matters socially as well as professionally. Social networking websites and messaging services, blogging platforms, video sharing websites, and Internet forums and message boards enable individuals to become the sorts of people whom it is possible for very large numbers of other people to "like" because, for example, they are funny, cool, nice to look at, well-informed, chatty, savvy, and so forth. Indeed, there is an increasing tendency for online technologies to facilitate and foreground quantitative tools for measuring online likeability. Consider the number of Twitter followers, retweets, mentions, and heart clicks a user can amass; the number of Facebook friends and post likes; the number of Internet forum and message board comments and webpage visits; the number of Instagram likes; and so on. Online hate speech poses a threat to likeability. Realizing or fearing that one is not or will not be liked by other members of the online communities to which one belongs because one has been the target or subject matter of online hate speech may harm the emotional tranquility of people who are

94. *See* Brown, *supra* note 13; Brown, *supra* note 90.

emotionally invested in being liked online. This in turn may be related to very basic, hardwired human fears of social ostracization and isolation. (Of course, the fear might prove to be unfounded if the parody actually increases the likeability of the victim in the eyes of others. But even this eventual good outcome might not transform the parody from unwelcome to welcome speech if the unwelcomeness is defined in terms of the risk of a bad outcome.)

I wish to make two things clear at this juncture. First, what I have been presenting in the last two paragraphs are hypotheses for consideration as opposed to proven facts. (I have written elsewhere of the importance of being upfront in debates on hate speech law about what evidence does or does not exist.[95]) Nevertheless, if these hypotheses are accurate, then, I believe, they would add considerable force to the proposition that some online hate speech can be unwelcome because it harms significant interests.

Second, I do not take myself to have exhausted the list of significant interests that might be harmed by exposure to online hate speech. I offer interests in privacy, autonomy, and emotional tranquility on the Internet and the web merely with a view to repurposing the captive audience doctrine. If the aim were to identify interests harmed by online hate speech *simpliciter*, no doubt one could point to various other interests, such as interests in nonsubordination; interests in freedom from oppression; interests in the protection of human dignity; and interests in the public assurance of civic dignity.[96] For instance, it might be argued that part of the harm of hate speech in public forums on the Internet and the web is that such speech can function like "keep out" signs. Online hate speech either directly or indirectly tells certain groups of people that they are not welcome in cyberspace, as well as not being welcome in society as whole.

95. *See* Brown, *supra* note 86, at chs. 3 and 9; Brown, *supra* note 13.

96. *See* Brown, *supra* note 86, at chs. 3 and 5; Alexander Brown, *The Meaning of Silence in Cyberspace: The Authority Problem and Online Hate Speech*, in FREE SPEECH IN THE DIGITAL AGE (Katharine Gelber & Susan Brison eds., forthcoming).

IV. AVERTING YOUR EYES IN THE INFORMATION AGE

In order to show that online hate speech can create or exploit captive audiences, it is not enough merely to show that the hate speech received or accessed through the Internet or the web is unwelcome (i.e., damages a significant interest). It must also be demonstrated that the recipients are unwilling. Part IV aims at making that case.

A. Practically Helpless to Avoid Speech

In this subpart, I want to challenge the assumption that in most cases it is possible to avoid or escape unwelcome messages and content received or accessed through the Internet or on the web. If left unchallenged this assumption would obstruct the proposed extension of the captive audience doctrine to online hate speech. After all, the U.S. Supreme Court has been consistent in affirming that people are a captive audience to speech only if they are practically helpless to avoid or escape it. In *Cohen*, for example, the Court ruled that when Cohen wore a jacket that displayed the phrase "Fuck the Draft" in the public corridors of a courthouse this did not create a captive audience because others "could effectively avoid further bombardment of their sensibilities simply by averting their eyes."[97] Similarly, in *Spence v. Washington*[98] the Court held that passersby who might have been offended by a flag hung from a second-floor apartment window were not a captive audience because they "could easily have avoided the display."[99] And, in *Heffron v. International Society for Krishna Consciousness, Inc.*,[100] Justice Brennan opined that "[b]ecause fairgoers are fully capable of saying 'no' to persons seeking their attention and then walking away, they are not members of a captive audience."[101] I want to explode some of these ways of thinking about captive audiences or, more precisely, some analogous ways of thinking about the

97. Cohen v. California, 403 U.S. 15, 21 (1971).
98. 418 U.S. 405 (1974).
99. *Id*. at 412.
100. 452 U.S. 640 (1981).
101. *Id*. at 657–58 n.1.

practicalities of avoiding or escaping online hate speech.

Let me begin with this question: Is it fair to say that someone who receives an email or Tweet containing hate speech or who comes across a hate site is not a captive audience simply by virtue of the fact that he or she can delete the email or Tweet upon first reading and can elect not to look again at the hate site? I believe not. The point of averting one's eyes is not merely to avoid the speech but also to do so whilst at the same time avoiding harm to significant interests. If the harm occurs from the first moment that the hate speech is received and can persist even after one has averted one's eyes, then averting one's eyes is not a practical means of avoiding the harm. Indeed, in *FCC v. Pacifica Foundation*[102] the U.S. Supreme Court considered and rejected the idea that listeners could be reasonably expected to turn off the radio upon hearing offensive material.

To say that one may avoid further offense by turning off the radio when he hears indecent language is like saying that the remedy for an assault is to run away after the first blow. One may hang up on an indecent phone call, but that option does not give the caller a constitutional immunity or avoid a harm that has already taken place.[103]

The fact that it may not always be possible to psychologically *un*read online hate speech once it has been read is especially germane to the interest in emotional tranquility discussed in Part III.B. In Case 2: Hate Sites, for example, the father's emotional distress occurred as soon as he read the hate site, and so he could not have avoided the harm simply by closing down the page and never looking at it again. Whilst the father can avoid the website from that point onwards, the emotional damage cannot be so easily erased or undone. These reflections point to a relevant distinction between two kinds of averting one's eyes: *ex ante*, before one has even received or read unwelcome speech; and *ex post*, after one has already received or read the relevant speech. Averting one's eyes *ex ante* may well be the more important kind when it comes to applying the captive audience doctrine to online hate speech.

102. 438 U.S. 726 (1978).
103. *Id*. at 748–49.

Following on from this distinction, I now want to explore another aspect of received wisdom about the Internet and the web mentioned in Part I: that computer-based communication services provide the user with a range of effective methods with which to avoid receiving unwelcome content.[104] This is averting one's eyes *ex ante*. In practical terms it could mean seeking out only those mobile network providers and Internet service providers who offer content-limited access to the Internet and the web through network-based Internet-security filters which block access to hate sites (amongst other unwelcome websites). Or it could mean users availing themselves of the browser-based Internet-security controls offered by the major Internet browsers and web search engines (e.g., AltaVista, Bing, Internet Explorer, Mozilla Firefox, Google, Safari, Opera, Yahoo), which block websites based on content. Or it might involve the use of email filtering functions provided by email software packages (e.g., Outlook, Hotmail, Gmail) which place filters on emails containing certain words in the subject line, such as by sending them directly to spam folders. Or it could mean downloading specialist third party software packages (e.g., CyberPatrol's CyberList, K9, DansGuardian) which, once activated, will deny access to any websites that fall within the software providers' proscribed categories. This would be equivalent to informing the post office that one does not wish to receive mail from a particular list of senders who are known for sending hate mail, for instance.

However, these technologies suffer from three main problems. First, they can have a tendency to underfilter. Online hate speakers can disguise email messages and websites to look harmless. And they may be able to bypass filters by changing the subject lines or titles of the emails or websites, or even by changing words within the emails or websites themselves. In short, there are forms of covert hate speech that filtering

104. *See*, *e.g.*, ABRAHAM H. FOXMAN & CHRISTOPHER WOLF, VIRAL HATE: CONTAINING ITS SPREAD ON THE INTERNET 189 (2013); Johnson, *supra* note 19, at 85–90; Balkin, *Free Speech and Hostile Environments*, *supra* note 7, at 2311; James Banks, *Regulating Hate Speech Online*, 24 INT'L REV. L. COMPUTERS & TECH. 223, 238 (2010); Zick, *supra* note 27, at 28–30.

technologies are inadequately equipped to filter out.[105] More generally, these technologies are susceptible to the constantly changing forms of hate speech transmitted through the Internet and found on the web, and are liable to lag behind the changing techniques used by online hate speakers. This seems to have been the fate of the Anti-Defamation League's HateFilter software.[106] If levels of underfiltering are significant, then "it could be argued that the inability to filter out undesirable speech creates an unacceptable dilemma for a would-be user: use the Internet and subject yourself to the risk of encountering [unwelcome] speech, or abstain altogether from using the medium."[107] Second, these technologies can have a tendency to overfilter. Imprecise software might block access to websites on which hateful words are *mentioned* but not *used*.[108] Consider websites that are devoted to discussing the problem of hatred; that provide support for the victims of hate speech; that monitor and report the use of cyberhate; that provide platforms for counterspeech; and so on.[109] Is it reasonable to expect people to risk inadvertently cutting themselves off from helpful public discourse about hate speech in order to avoid hate speech itself? I shall return to the general issue of unreasonable burdens in a moment. Third, user-oriented software solutions, which require users to set up, manage, and periodically adjust or customize content filters, could be *de facto* accessible only to computer and mobile device users who feel sufficiently knowledgeable and competent.[110] People who are, or perceive themselves to be, technologically illiterate may end up being practically helpless to

105. *See*, *e.g.*, Jessie Daniels, *Race, Civil Rights, and Hate Speech in the Digital Era*, *in* LEARNING RACE AND ETHNICITY: YOUTH AND DIGITAL MEDIA 148 (Anna Everett ed., 2008).

106. Jessica S. Henry, *Beyond Free Speech: Novel Approaches to Hate on the Internet in the United States*, 18 INFO. & COMM. TECH. L. 235, 247–48 (2009).

107. Araiza, *supra* note 73, at 404.

108. The general distinction between mentioning and using hate speech is due to KEVIN W. SAUNDERS, DEGRADATION: WHAT THE HISTORY OF OBSCENITY TELLS US ABOUT HATE SPEECH 3, 156–57 (2011).

109. *See* Timofeeva, *supra* note 22, at 280; BRETT A. BARNETT, UNTANGLING THE WEB OF HATE: ARE ONLINE "HATE SITES" DESERVING OF FIRST AMENDMENT PROTECTION? 188 (2007).

110. *See* DAVID S. HOFFMAN, HIGH-TECH HATE: EXTREMIST USE OF THE INTERNET 11 (1997).

avoid online hate speech except by going offline entirely.

Now it might be pointed out at this stage that in *Ashcroft v. ACLU*[111] and *ACLU v. Gonzales*[112] the courts considered expert testimony and surveyed evidence from governmental and other studies relating to both the effectiveness and ease of use of Internet filtering software that can be used by parents to prevent children from accessing pornographic material over the Internet. They concluded that underfiltering, overfilting, and difficulties of use were not significant problems. However, hitherto courts in the United States have not investigated whether or not the same can be said for Internet filtering software that can be used by potential victims of cyberhate to prevent themselves from receiving or accidentally accessing hate speech through the Internet or on the web.

Another method that users of the Internet and the web can employ in order to avoid receiving unwelcome content is to avail themselves of the blocking and unfollowing functions provided by social networking websites, like Facebook, and Internet messaging services and microblogging platforms, like Twitter. Consider again *Cassidy*, a case involving an indictment for the federal crime of intentionally causing substantial emotional distress to another person by way of an interactive computer service.[113] Cassidy's incessant Tweets mentioned Zeoli in foul and abusive terms and eventually caused her to deactivate her Twitter account.[114] Before that time, Zeoli addressed Internet users on a frequent basis from her own verified Twitter account, which had 17,221 followers.[115] Major argues that "normatively, the victims of cyberstalking should not have to quit Twitter to avoid that message."[116] But the practical question is whether Zeoli really needed to quit Twitter in order to avoid Cassidy's unwelcome tweets. Could she not have continued to use Twitter and chosen to unfollow and block Cassidy? In this way, Cassidy's tweets would not have appeared on Zeoli's home timeline, he

111. 542 U.S. 656 (2004).
112. 478 F. Supp. 2d 775 (E.D. Pa. 2007).
113. United States v. Cassidy, 814 F. Supp. 2d 574 (D. Md. 2011).
114. *Id.* at 588-91.
115. *Id.* 586 n.14.
116. Major, *supra* note 34, at 151.

could not have sent her direct messages, and he could not have seen Zeoli's tweets or other account information. Major, however, rebuffs this reasoning, "The captive audience doctrine is especially applicable to cyberstalking when cyberstalkers go out of their way to make their speech unavoidable, such as through the use of multiple usernames."[117] Perhaps Major has in mind the fact that when other Twitter users, including Cassidy, referred to Zeoli in their tweets, these mentions would have appeared in her mentions tab, whether or not she followed them. However, this response overlooks an important privacy function available to all Twitter users. Users now have the ability to click the "People You Follow" function on the mentions tab.[118] This enables a mentions filter such that the mentions tab will only display mentions from followed users.[119] By enabling this privacy function and only following people they know, Twitter users can now avoid unwelcome mentions.[120]

Twitter, then, does provide some privacy controls that allow people to prevent unwelcome speech from being sent directly to them.[121] It is not alone. Recall the discussion of Case 1: Tagging, in Part III.B above. Here I pointed to Facebook's advanced privacy controls that allow users to review tags that people add to their profile page timelines before the tags actually appear or go live. Arguably, the existence of these types of privacy controls weakens the claim that people can be captive audiences to unwelcome tags and mentions on social networking websites and Internet messaging services. Then again, perhaps the captive audience claim will not be fatally weakened by the existence of these controls so long as Facebook and Twitter continue to make

117. *Id*. at 150.

118. Using Twitter Search, TWITTER, https://perma.cc/AT4E-23D2.

119. *Id*.

120. *Id*.

121. Of course, if Zeoli chose to do a Twitter search of her own Twitter username ("handle"), then this would throw up Cassidy's tweets in the results page. Major believes that this is important. "The captive audience nature of Twitter is buttressed by the importance of search functions on Twitter and the growing importance of self-Googling and monitoring of one's online reputation. Because of these aspects, an individual may confront speech intended to cause them substantial emotional distress without seeking it out." Major*, supra* note 34 at 151.

the default settings for these controls disenabled rather than enabled. The crucial point is that Facebook and Twitter users must take active steps to enable the "Reviewing" and "People You Follow" controls, which means they must first come to know that these controls exist and then make efforts to locate and enable them. The problem is that many Facebook and Twitter users may be simply unaware of these advanced privacy controls until after they have had a bad experience with unwelcome tags or mentions. And it is not clear whether users of Facebook and Twitter users can be reasonably expected to go through every setting and control at the point at which they begin using these services and websites. Therefore, if it *is* unreasonable to expect this level of due diligence, then these sorts of cases *do* involve captive audiences, other things remaining equal. If it is *not* unreasonable, then it seems likely that they do *not* involve captive audiences, other things remaining equal.

I have investigated some of the technical difficulties associated with filtering, blocking, unfollowing, and reviewing. Yet these do not exhaust the types of practical impediments that could be faced by people when it comes to avoiding online hate speech. Some necessary avoidance behaviors might cut against the grain of psychological facts about ordinary human beings. There is judicial precedent for this sort of analysis. In *Erznoznik v. Jacksonville*,[122] for example, the Court held as invalid an ordinance making it unlawful for a drive-in movie theatre to exhibit films containing nudity when the screen is visible from a public street, on the grounds that "the offended viewer readily can avert his eyes."[123] In his dissent, however, Chief Justice Burger observed that "when films are projected on such screens the combination of color and animation against a necessarily dark background is designed to, and results in, attracting and holding the attention of all observers."[124] Indeed, it has recently been suggested that the human psychological propensity to be drawn to bright shiny objects stems from an innate need for

122. 422 U.S. 205 (1975).
123. *Id*. at 212.
124. *Id*. at 221.

water.[125] Could a similar form of argument be made for hate speech on the Internet and the web?

To be sure, many parts of the web are not the sorts of places that a person can literally stumble upon by accident, that is, without any forewarning of the likely content.[126] Web search engines present titles and snippet views of the contents of the websites listed on the search results pages. As such, it might be supposed that people can be expected not to click on any results which suggest, hint at, or contain snippets of unwanted content. Surely, we can expect (so the argument goes) the parents of deceased military personnel who put their children's names into web search engines not to click on link to websites with names such as www.godhatesfags.com. However, could not the captive audience doctrine be applied even to cases where people had some foresight or warning that they were about to access hate speech if they nevertheless felt in some sense compelled, psychologically, to click on the link? The compulsion might stem from an instinct for self-preservation. People may feel driven to seek out information that may be useful for keeping them safe even if that information is emotionally painful. Thus, previous victims of racist abuse might feel compelled to educate themselves about online hate speech so that they are less likely to be blindsided by an escalation of hatred. Alternatively, looking at emails, websites, or other online content even when one has an inkling that it might contain cyberhate could reflect a human tendency toward morbid curiosity. Many people experience an irresistible urge to enquire further about subjects that they suspect, or know, will be upsetting to them and which on one level they really do not want to know more about. And it is possible that people who suspect that they are the subject matter of online hate speech could experience a morbid curiosity to read or see the harmful content in its entirety and not just in snippet view. This tendency might reflect personality type. But it might also conceivably be heightened by experience. If someone has

125. Katrien Meert et al., *Taking a Shine to It: How the Preference for Glossy Stems from an Innate Need for Water*, 24 J. CONSUMER PSYCHOL. 195 (2014).

126. *See* ACLU v. Reno, 929 F. Supp. 824, 844 (E.D. Pa. 1996).

been the victim of hate speech in the past, for instance, ironically this trauma could make him or her less able to control, resist, or override an instinctive morbid curiosity to click on website links which he or she suspects may contain yet more hate speech. Bereavement is another traumatic experience that could conceivably reduce a person's ability to control, resist, or override an impulse toward morbid curiosity. Of course, hate speech is not only found online, and a general impulse toward morbid curiosity might also drive a person to listen to hate speakers in person or in offline situations. Then again, it is possible that the impulse toward morbid curiosity is harder to resist in the case of online hate sites precisely because of the relative convenience, anonymity, and physical safety provided by the Internet and the web, not to mention to shimmer of the computer screen.[127] Whilst I offer no empirical evidence about morbid curiosity and online hate speech, it is not beyond the realms of possibility that such evidence will begin to emerge as we learn more about human usage of the Internet and the web.

Does this mean, however, that there is a sense in which someone is not being held captive by the online hate speech *as such* but by his or her psychological compulsions? Perhaps, but I do not think that this makes it any less the case that someone could be a captive audience *to* online hate speech, even if this is *partly because* of morbid curiosity. Nor do I think that this link in the causal chain alters the fact that hate speakers can be held accountable both for the speech and for the captivity. By analogy to the eggshell skull rule in tort law, maybe hate speakers must take their audiences as they find them—neuroses and all.

B. Unreasonable Burdens

It is one thing to highlight the practical difficulties of avoiding online hate speech; it is quite another thing to show that the audience is for all intents and purposes held captive by those difficulties. After all, people always have the last resort of simply deactivating their Facebook or Twitter accounts or stepping away from the Internet and the web indefinitely.

127. I thank Sonu Bedi for this suggestion.

Consequently, some people might be tempted to say that an audience is unwilling only if it is unable to avoid or escape the speech, and that Internet and web users are never unable to avoid or escape online hate speech because they can always choose to live offline.

On closer inspection, however, this conceptualization of the captive audience doctrine is too restrictive. It implies a type of physical imprisonment. Instead, the captive audience doctrine is best understood as the claim that an audience cannot be reasonably expected to avoid or escape exposure because of the significant interests that would be harmed, sacrificed, or forfeited if it did. The operative test of captivity is whether an ordinary person can take practical steps to avoid or escape receiving the speech in question whilst at the same time not incurring an unreasonable burden, that is, a burden they ought not to have to incur. Only if someone can take practical steps to avoid online hate speech without bearing an unreasonable burden—without suffering harm to significant interests—can it be said that he or she is not a captive audience.[128]

Take Case 3: Trolling. If friends and family of the deceased interracial couple, including people who might themselves be in interracial relationships, want to access a full repository of pictures, videos, and messages about the deceased couple, and to partake in communal processes and rituals of grieving, remembrance, and memorialization, and if they lack reasonable alternatives to accessing this content and these activities either on other websites which are not subject to trolling or in offline spaces which are not subject to similar intrusions, then they have no option but to visit the website in question. Insofar as they lack reasonable alternatives to accessing this content and these activities, and it would be unreasonable to expect them to forego this content given their equal right to enjoy a normal opportunity range in life, including grieving and memorializing, then surely this makes them a captive audience.

128. *See also* JOEL FEINBERG, THE MORAL LIMITS OF CRIMINAL LAW, VOL. 2: OFFENSE TO OTHERS 32, 308 n.2 (1985); Major, *supra* note 34, at 127, 151; Balkin, *Free Speech and Hostile Environments*, *supra* note 7, at 2312**;** Sadurski, *supra* note 8, at 186; Strauss, *Redefining the Captive Audience Doctrine*, *supra* note 7, at 89 .

More generally, it is not hard to think of unreasonable burdens associated with a general abstinence from the Internet and the web. For one thing, people are increasingly using the Internet and the web as the central loci of their economic activities. I have in mind people who use the web to truck, barter, and exchange. But, more generally, most people nowadays use the Internet and the web to perform at least some of the essential duties associated with their jobs or professions. I am not necessarily speaking here of people whose profession is intimately connected with their being online and encouraging online comments. Perhaps it would stretch the idea of a captive audience too far to say that a public blogger, for instance, is a captive audience to the comments section on her own blog. Rather, I mean people who find themselves having to use the Internet and the web simply to do their jobs, as much as they might find themselves having to travel on public transport simply in order to get to their place of work. Consider the judge in Case 6: Cyberharassment. The point is that "[i]n today's interconnected world [avoidance] is not a viable option, as people who are forced offline forgo important personal and professional opportunities."[129]

Several scholars of the existing captive audience doctrine have already observed that sometimes avoiding unwelcome speech may involve sacrificing economic interests.[130] Indeed, courts in the United States have frequently accepted that employees can be captive audiences in virtue of (a) the practical necessities of earning a living in order to make ends meet, (b) the reality that employees are required by their contracts and/or directed by their employers to be at work, within the confines of certain spaces, at certain times, and so forth, and (c) the uncertainties, stresses, and financial costs associated with seeking alternative forms of employment if one quits one's job in

129. Jacqueline D. Lipton, *Combating Cyber-Victimization*, 26 BERKELEY TECH. L.J. 1103, 1113 (2011).

130. *See* Balkin, *Free Speech and Hostile Environments*, *supra* note 7, at 2312–13; Balkin, *Some Realism About Pluralism*, *supra* note 7, at 423–24; Jessica M. Karner, *Political Speech, Sexual Harassment, and a Captive Workforce*, 83 CALIF. L. REV. 637, 682–83 (1995).

order to avoid or escape unwelcome speech.[131] Courts have also held that employees using public transit vehicles in order to travel to work can be captive audiences because (d) the nature of living and moving in built-up areas means public transport is the only viable option.[132]

Of course, not every free speech scholar agrees on these issues. According to Eugene Volokh, for instance, the argument that employees can be a captive audience is flawed because if it is true, it proves too much.[133] Put simply:

> [I]f captivity consists of an inability to avoid offensive speech, in today's society we are all "captive" to profanity. We may walk away from someone who is using it, but we cannot avoid it altogether—we hear it wherever we go. This is, regrettably, also true of bigoted abuse. In many places, blacks will be called names wherever they go; obese or disfigured people may be insulted wherever they go. Even in public, they may be able to avoid an individual insulter (though not without being insulted first), but they cannot avoid the insults altogether.[134]

But I would argue that we are only drawn to the implausible conclusion that all audiences are captive audiences by oversimplifying the concept of captivity. Captivity is not a matter of being unable to avoid unwelcome speech altogether. Rather, it is a matter of being unable to both avoid unwelcome speech and avoid unreasonable burdens whilst doing so.

Returning to Case 6: Cyberharassment, it may be perfectly true to say that as a woman the judge risks being targeted by misogynistic speech in various contexts, but the question is whether or not it would be reasonable to expect her to avoid those contexts. In terms of her professional duties, the question is whether it would be reasonable to expect a woman who suffers cyberharassment whilst working as a judge to quit her job, given (e) the difficulty of finding an alternative line of work that would

131. *See* Resident Advisory Bd. v. Rizzo, 503 F. Supp. 383, 402 (E.D. Pa. 1980).

132. *See* Lehman v. Shaker Heights, 418 U.S. 298, 302 (1974); Pub. Util. Comm'n v. Pollak, 343 U.S. 451, 468 (1952).

133. Eugene Volokh, *Freedom of Speech and Workplace Harassment*, 39 UCLA L. REV. 1791, 1839–40 (1992).

134. *Id.*

provide an equivalent sense of professional fulfillment or achievement of life ambitions, (f) the problem of finding an alternative line of work in which other forms of discriminatory harassment will not re-emerge, (g) the lack of equal opportunities faced by all women in accessing well-paid professional employment, and (h) the risks for women in terms of domestic violence and other forms of oppression of being financially dependent on male breadwinners during periods of job-seeking. Of course, the court authorities could, in the light of the problem of misogynistic cyberharassment, opt to refrain from posting online any legal judgments made by female judges. Then again, it might be unreasonable to expect authorities to take this step, given (i) the unwanted symbolism of publishing only the judgments of male judges. Of course, they could refrain from posting any judgments period. But then this might be unreasonable given (j) the negative effect on public confidence in the justice system, and (k) the detriment to the educative function of the law.

More generally, other unreasonable burdens associated with logging off in the Information Age might include: diminished access to information about important public services, community events, news, and current affairs; loss of the ability to contribute to the formation of public opinion not simply about current affairs but also in relation to popular culture in general (public opinion which may undergird formal processes of democratic decision-making); loss of access to friendship and social life; loss of access to communal games or play; loss of access to potential life partners, such as through social networking websites or online dating apps; loss of access to processes of grieving, remembrance, and memorialization of deceased friends and loved ones; and so on. The suggestion here is *not* that people who log off have no other means of accessing these opportunities. What I am describing is *not* akin to the sense in which welfare recipients waiting in line are a captive audience to unwelcome speech because they "have no choice but to come to the Local Office for the basic necessities of life."[135] Rather, my

135. Families Achieving Independence & Respect v. Neb. Dep't of Soc. Servs., 111 F.3d 1408, 1412 n.5 (8th Cir. 1997).

suggestion is that if a great many people are predominantly doing these things online rather than offline, and if a vast majority of people are at least sometimes doings things online rather than offline, then persons who live entirely offline will not merely be the odd ones out but may be significantly disadvantaged.

Consider once again Case 3: Trolling. Facebook alone contains vast numbers of pages dedicated to processes of grieving, remembrance, and memorializing. Given the demands of work and the problems of geographical distance, not all family members, friends, and colleagues of the deceased may have an opportunity to attend a formal ceremony. Unsurprisingly, therefore, a good deal of the memorializing which takes place online is about expressions of feelings and the sending of condolences that could not be expressed or sent on the day of the funeral. Of course, it might be argued that strictly speaking nobody is compelled to look at, much less contribute to, these websites. But that observation is not merely callous, but ignores the possibility that partaking in processes of grieving, remembrance, and memorialization may be partly constitutive of a flourishing human life. In theory, users could avoid the hate speech *ex ante* by installing filtering technologies that would block them from accessing websites that have fallen pray to trolling. But this would harm a significant interest in partaking in the aforementioned processes. If they do not install filters, then even once they know that the hate speech is present on the sites, they cannot take steps to avoid it *ex post* without again sacrificing significant interests. In short, trolling forces some people to make a tragic choice that other people do not have to make.[136]

What I am suggesting, in other words, is that using the Internet and the web can be compelled by the facts of life in the Information Age.[137] I have also been assuming that in extreme cases averting your eyes in the Information Age may involve

136. Similar ideas were expressed by the court in Phelps-Roper v. Strickland, 539 F.3d 356, 366 (6th Cir. 2008).

137. This phrasing is a conscious reference to Justice Douglas' Concurring opinion in Int'l Ass'n of Machinists v. Street, 367 U.S. 740, 776 (1961).

deactivating, logging out, and switching off. But are there *degrees* of averting your eyes that are worth considering here? What if people find themselves only having to avoid some features and parts of the Internet and the web where there are plenty of other features and parts to choose from?[138] In *United States v. Sayer*,[139] the District Court (D. Me.) judge rejected the idea that victims of cyberstalking can be reasonably expected "not to open mail or [to] ignore electronic communications such as email, Facebook postings, tweets, and text messages" because "[t]he First Amendment does not give stalkers a license to place special conditions on how their victims use modern forms of communication as the price of avoiding hateful attention."[140] The Court's decision was affirmed by the First Circuit Court of Appeals.[141] When it comes to hate speech on the Internet and the web, however, might it be reasonable to expect people to abstain from using at least some features and parts of the Internet and the web in order to avoid hate speech? Provided that users' options are not severely limited, (some people might think) maybe the interest in autonomy in using the Internet and the web need not be significantly damaged by some degree of abstinence.

What is likely to matter, it seems to me, is whether or not there is a decent range of options available to Internet and web users. And this must surely have to do with quality and not merely quantity. Most people use email accounts for work and personal usage. In order to avoid receiving messages containing hate speech would users need to steer clear of only some email account providers or all? Many people use social networking websites and Internet messaging services. Would it be reasonable to expect these users to stay away from whichever websites and services offered limited *de facto* protection against hate speech, even if they are the most widely used websites and services, in favor of specialist websites and services providing much better

138. *Cf.* Joseph Raz, *Free Expression and Personal Identification*, 11 OXFORD J. LEG. STUD. 303, 321 n.36 (1991).

139. Nos. 2:11–CR–113–DBH, 2:11–CR–47–DBH, 2012 WL 1714746 (D. Me. May 15, 2012).

140. *Id*. at n.17.

141. United States v. Sayer, 748 F.3d 425 (1st Cir. 2014).

protection (assuming they existed)? Virtually everyone uses Internet browsers and web search engines of one form or another. Could we reasonably expect certain groups of people to pass up on any web search engines that did not aggressively filter out websites containing hate speech? Large numbers of people use video-sharing websites. Might potential victims of hate speech be reasonably expected to choose their preferred video-sharing websites based on the relative amounts of hateful content to be found (assuming people could find reliable data on this)? More generally, if it turns out that in order to avoid hate speech certain groups of people will need to abstain from using the most widely used email account providers, social networking websites, Internet messaging services, web search engines, and video-sharing websites, would this leave people with decent access to the Internet and the web?

How could this decent access be defined? One possibility is to ask whether or not people enjoy a normal opportunity range on the Internet and the web, as defined by the array of online activities that ordinary people are likely to want and need to partake in given reasonable life plans or reasonable conceptions of a fulfilling life. This might include: maintaining and creating friendships; pursuing a career; accessing information on current affairs; consuming and contributing to popular culture; play; participation in civic and political engagement and activism. If avoiding hate speech means abstinence from the most widely used websites and services, and if this abstinence in turn means not being able to enjoy a normal opportunity range of online activities, then the degree of avoidance expected would be unreasonable.

Take Case 5: Copycat Parodies. If checking one's reputation as a model (professional opportunities), confirming that one is a member of society in good standing (dignitary opportunities), and contributing to the discussion of issues surrounding transgender identity (public discourse opportunities), *are* components of a normal opportunity range on the Internet and the web, then this case *does* involve a captive audience, other things remaining equal. If these activities are *not* components of a normal opportunity range on the Internet and the web, then this case does *not* involve a captive audience, other things remaining

equal.

To come at the issue of partial avoidance from another angle, would it be reasonable to expect people to continue to avail themselves of the most widely used websites and services, but at the same time adapt the ways they use them? Online hate speakers often pray on people who opt to reveal aspects of their identities or on people whose identities are revealed without their consent or somehow assumed or inferred. Consequently, would it be reasonable to expect potential victims of hate speech to use social networking websites and Internet messaging services, say, but not to reveal or to actively disguise aspects of their offline identities such as their race, ethnicity, nationality, citizenship, origin of birth, war record, religion, sexual orientation, gender or transgender identity, disability, age, or physical appearance?

Now it is certainly true that one of the supposed advantages of the Internet and the web as a medium for sending and receiving information and as a virtual public square is that people are not compelled to reveal aspects of their offline identities unless they wish to do so. Their beliefs, ideas, and attitudes can be presented to other users in anonymous ways if they so desire.[142] But at this juncture, we are talking about

142. It has also been suggested, of course, that the anonymity of the Internet and the web can provide opportunities for freer speech because people can say what they think without fear that other people will react or respond unfavourably simply because of the colour of their skin or their gender, say. *See*, *e.g.*, GORDON GRAHAM, THE INTERNET: A PHILOSOPHICAL EXAMINATION 143 (1999). Nevertheless, this same anonymity can also embolden people to be more outrageous, obnoxious, or hateful in what they say than would be the case in real life. *See*, *e.g.*, Anne Wells Branscomb, *Anonymity, Autonomy, and Accountability: Challenges to the First Amendment in Cyberspace*, 104 YALE L.J. 1639, 1642–43 (1995); John Suler, *The Online Disinhibition Effect*, 7 CYBERPSYCHOLOGY & BEHAV. 321, 321–26 (2004). The anonymity of the Internet and the web may remove fear of being held accountable for speech acts. *See* Jerry Kang, *Cyber-Race*, 113 HARV. L. REV. 1130, 1135 n.16 (2000). Moreover, the non-face-to-face aspect of online communication means that the consequences of actions are unseen by the perpetrator thus potentially creating deficits in human empathy and sympathy. If one cannot see the emotional hurt wrought by one's own online hate speech, one may be more likely to downplay its significance ("It's only harmless flaming; people shouldn't take it so seriously"). *Cf.* Jacqueline Taylor, *Humean Humanity Versus Hate*, *in* THE PRACTICE OF VIRTUE: CLASSIC AND CONTEMPORARY READINGS IN VIRTUE ETHICS (J.

people effectively being distressed, scared, intimidated, or cowed by online hate speech into a kind of forced anonymity. Anonymity may be a bad thing, not merely because it is enforced, but also because of the cost in free speech it may impose. To be forced into using websites and services in anonymous ways might inhibit users' speaking their minds authentically, forming deep relationships, and sending and/or receiving genuinely personal messages. In other words, it might limit what they can get out of online experiences and interpersonal interactions in terms of self-realization, information gathering, contributing to public discourse, and so on.

Thus, the threat that one might become the victim of cyberhate may foreclose a range of ways of speaking and expressing oneself, as well as certain topics for discussion, and even particular spaces or places on the Internet and the web, for fear of giving away telltale signs of one's identity and opening oneself up to a deluge of cyberhate as a consequence. But can we reasonably expect members of groups or classes of persons who are likely to be subjected to cyberhate to have less meaningful friendships or other relationships online, gain a poorer understanding of who they are through self-expression online, pursue fewer economic opportunities online, contribute in diminished ways to the formation of popular culture and public opinion online, and participate to a lesser extent in various forms of civic, political, and religious life online, as the price they must pay to avoid hate speech? Surely not.

V. IMPLICATIONS

Finally, what, if anything, follows from showing that online hate speech can create or exploit captive audiences? One possible source of significance concerns the free speech doctrine of content and viewpoint neutrality. For several decades, this doctrine has stood as a constitutional barrier to hate speech law in the United States. Most notably, in *R.A.V.* the U.S. Supreme Court struck down the City of St. Paul Bias-Motivated Crime Ordinance *inter alia* because it involved content and viewpoint discrimination.

Welchman ed., 2006); Brown, *supra* note 86, at 132–37.

The Court accepted the City of St. Paul's argument that it had intended the ordinance as prohibiting fighting words (a category of proscribable speech), but the Court extended the requirements of neutrality even to laws which seek to draw distinctions between subcategories of proscribable speech. The majority ruled that the ordinance violated content neutrality in virtue of the fact that on its face it proscribed only fighting words with certain sorts of content (i.e., messages relating to race, color, creed, religion, or gender) and violated viewpoint neutrality because as-applied it prohibited only particular kinds of viewpoint (i.e., people holding or advocating a particular position on issues of race, color, creed, religion, or gender).[143] Much has been written about *R.A.V.* and about whether or not content- and viewpoint-based regulations are always impermissible under the First Amendment and, furthermore, whether hate speech regulations might be permissible exceptions under a nuanced reading of the content and viewpoint neutrality doctrine, in virtue of their sometimes fitting into the broader class of harm-preventing speech regulations or even "militant" democracy-protecting speech regulations.[144] Nevertheless, what interests me here is the fact that in *R.A.V.* the Court chose not to consider whether or not this case involved a captive audience. But if it could be shown that certain instances of hate speech, including but not limited to cross burning, involve captive audiences, then surely there is at least potential for the captive audience doctrine to defeat, or fall

143. R.A.V. v. City of St. Paul, 505 U.S. 377, 391–92 (1992).

144. For discussion of these issues, *see*, *e.g.*, Owen Fiss, *Free Speech and Social Structure*, 71 IOWA L. REV. 1405, 1415–21 (1986); David Kretzmer, *Freedom of Speech and Racism*, 8 CARDOZO L. REV. 445, 458 (1987); Lawrence, *supra* note 9, at 458–59; Akhil Reed Amar, *The Case of the Missing Amendments: R.A.V. v. City of St. Paul*, 106 HARV. L. REV. 124, 125–26 (1992); Joshua Cohen, *Freedom of Expression*, 22 PHIL. & PUB. AFF. 207, 255–56 (1993); Shiffrin, *supra* note 37, at 50–64; James Weinstein, *Hate Speech, Viewpoint Neutrality, and the American Concept of Democracy*, *in* THE BOUNDARIES OF FREEDOM OF EXPRESSION AND ORDER IN A DEMOCRATIC SOCIETY (T. Hensley ed., 2001); Heyman, *supra* note 37, at 690–91, 713–74; Heyman, *supra* note 4, at 273 n.26; Cori Brettschneider, *Value Democracy as the Basis for Viewpoint Neutrality: A Theory of Free Speech and Its Implications for the State Speech and Limited Public Forum Doctrines*, 107 NW. U. L. REV. 603, 609–10 (2013); Brown, *supra* note 86, at 269–271, 287–97; Alexander Brown, *Hate Speech Laws, Legitimacy, and Precaution: A Reply to James Weinstein*, 32 CONST. COMMENT. 599.

within recognized exceptions to, countervailing free speech doctrines, including the content and viewpoint neutrality doctrine, in the relevant cases. And if that is possible, then perhaps the same could be said of laws banning online hate speech involving captive audiences.

Clearly, these are contested issues. In the context of discussing other forms of speech (besides hate speech), Strauss has suggested that "the captive audience doctrine allows courts to ignore the traditional requirement of content neutrality; courts inevitably engage in viewpoint- or content-based discrimination when applying the doctrine."[145] Following on from this, Major has tentatively proposed that Strauss' thesis might also hold for cases of cyberstalking[146] including *Cassidy*, albeit, as I have already pointed out, the court in that case did not explicitly address itself to the captive audience doctrine. Yet there is no doubting that some scholars of American free speech doctrine would insist that cases involving content- and viewpoint-based hate speech regulations are a different matter.[147] Indeed, when it comes to discriminatory harassment in the workplace, several scholars have argued that the captive audience dimension does *not* allow courts in the United States to ignore the traditional requirements of content and viewpoint neutrality.[148] Therefore, what I am proposing in terms of changes to current free speech doctrine are likely, at best, to seem plausible to only a minority of free speech scholars within the wider debate on, and controversies surrounding, hate speech regulations in the United States.

A second possible source of significance has to do with principles of net neutrality, namely, that Internet providers may not pick and choose what content to make available to their

145. Strauss, *Redefining the Captive Audience Doctrine*, *supra* note 7, at 86–87.

146. *See* Major, *supra* note 34, at 151.

147. *See* Weinstein, *supra* note 144.

148. *See* Lawrence Tribe, AMERICAN CONSTITUTIONAL LAW 949–50 n.24 (2d ed., 1988); Volokh, *supra* note 133, at 1833, 1840–43; Eugene Volokh, *One-to-One Speech vs. One-to-Many Speech, Criminal Harassment Laws, and "Cyberstalking,"* 107 NW. U. L. REV. 731, 768–69 (2013); Nadine Strossen, *The Tensions Between Regulating Workplace Harassment and the First Amendment: No Trump*, 71 CHI.-KENT L. REV. 701, 706–07, 709–10 (1995).

users—principles highly relevant to the regulation of the Internet and the web. In 2015, the Federal Communications Commission (FCC) published new rules on the open Internet,[149] including the following net neutrality rule: "A person engaged in the provision of broadband Internet access service, insofar as such person is so engaged, shall not block lawful content, applications, services, or non-harmful devices, subject to reasonable network management."[150] In December 2017, however, the FCC voted to repeal this "heavy-handed framework".[151] The purpose of my article is not to try to second-guess what the courts in the United States are likely to decide in cases challenging the FCC's decision to repeal net neutrality rules, or what the FCC and the courts will henceforth decide in cases which implicate both net neutrality rules and the captive audience doctrine—or even if they would in fact acknowledge the existence of such cases. Instead, I want to make it clear (based on the foregoing arguments) what I believe they *should* say.

For one thing, they should say that the principle of net neutrality is not lexically prior to the captive audience doctrine and *vice versa*.[152] No doubt, some people believe that regulating online hate speech can be justified in spite of the principle of net neutrality, and perhaps also that the question of whether or not online hate speech involves captive audiences makes no difference one way or the other to this justification.[153] Conversely, others believe that regulating online hate speech *cannot* be justified precisely because of the values that underpin the principle of net neutrality, and that the question of whether or not online hate speech involves captive audiences also makes no difference to this failure of justification. My own view is that

149. FCC Open Internet Order, 47 C.F.R. § 8.18 (2015).

150. *Id* at § 15.

151. FCC News Release, *FCC Acts To Restore Internet Freedom* (Dec. 14, 2017), https://perma.cc/8UJ3-B33A.

152. Elsewhere I have argued that it is appropriate to view the resolution of these types of dilemmas not as a matter of establishing lexical priority among principles but as a matter of forging principled compromises—not merely compromise between principles but also compromise, which is itself regulated by principles (moral duties or civic virtues) such as equality, reciprocity, mutual respect. See Brown, *supra* note 86, at ch. 10.

153. *Cf.* Cohen-Almagor, *supra* note 13.

existing arguments for regulating online hate speech *can* be made stronger by appealing to the captive audience doctrine. I do not mean to say, however, that this appeal is a necessary or a sufficient condition for making a decisive argument for the regulation of online hate speech. Instead, I simply say that appealing to the captive audience doctrine can make a positive difference to the justificatory score.

Furthermore, I believe that Internet regulators and courts—and Internet companies themselves for that matter—should accept that persons can be held as captive audiences in public forums on the Internet and the web, as well as in nonpublic forums and in hybrid public/non-public forums on the Internet and the web. Once again, in making this argument I do not seek to deny the importance of the Internet and the web *qua* tools of mass communication and servants of free speech values, such as the pursuit of truth, self-realization, and the formation of democratic public opinion. Instead, I mainly want to challenge the simplistic assumption that because communicative content received through or viewed on the Internet and the web never appears on users' screens without users taking a series of affirmative steps, there can be no captive audiences on the Internet and on the web.[154]

Finally, in this contribution I have only made suggestions as to how American free speech doctrine should be changed so as to provide constitutional breathing space, so to speak, for restrictions on online hate speech in circumstances that this type of speech creates or exploits captive audiences. I have not sought to propose or defend any particular hate speech laws or regulations that might be applicable to online hate speech involving captive audiences, nor any precise scheme of liability for online hate speakers in such cases, be it through Internet regulation, criminal law, civil law, or human rights law.[155] Nor have I tried to defend any specific regimes of Internet regulation—whether that is the statutory regulation of Internet companies or a scheme of self-regulation by Internet companies—

154. *See* Weintraub-Reiter, *supra* note 18, at 161–65.

155. For an overview of the possibilities, *see* Brown, *supra* note 86, at ch. 2; Brown, *supra* note 13; Brown, *supra* note 90.

that could be used to tackle instances of online hate speech involving captive audiences. These issues implicate much larger and vitally important debates about liability, responsibility, and Internet regulation addressed elsewhere in the literature.[156] Moreover, in highlighting the applicability of the captive audience doctrine, I do not mean to exclude other avenues for redress among victims of online hate speech. Various writers have argued, for instance, that the tort of intentional infliction of emotional distress can be, and should be, used as a remedy for victims of hate speech,[157] and some have extending these proposals so as to include online hate speech.[158] Another writer has examined the prospects for using the tort of invasion privacy (false light) in cases involving hate speech, including online hate speech.[159] I do not seek to pretend that the captive audience doctrine would be strictly necessary for these other extension projects.[160] Rather I have merely tried to motivate the applicability of the captive audience doctrine to cases of online hate speech, whilst noting the varied nature not merely of cyberhate itself but also of spaces and places on the Internet and the web.[161]

156. For more on the practicalities of, and ethical responsibilities for, regulating online hate speech, *see*, *e.g.*, Tsesis, *supra* note 13; Banks, *supra* note 104; Cohen-Almagor, *supra* note 13; Brown, *supra* note 13.

157. *See* James Jay Brown & Carl L. Stern, *Group Defamation in the U.S.A.*, 13 CLEV. ST. L. REV. 7, 29 (1964); Dean M. Richardson, *Racism: A Tort of Outrage*, 61 OR. L. REV. 267 (1982); Matsuda, *supra* note 9, at 2336 n.83. *See also* Brown, *supra* note 90.

158. Catherine E. Smith, *Intentional Infliction of Emotional Distress: An Old Arrow Targets the New Head of the Hate Hydra*, 80 DENV. U. L. REV. 1, 29–31 (2002).

159. *Cf.* Anita L. Allen, *Privacy Torts: Unreliable Remedies for LGBT Plaintiffs*, 98 CALIF. L. REV. 1711, 1757–64 (2010).

160. *See also* Brown *supra* note 90.

161. *See also* Brown *supra* note 96.