MEDIA AND LAW

Between Free Speech and Censorship

Edited by Mathieu Deflem and Derek M.D. Silva

SOCIOLOGY OF CRIME, LAW AND DEVIANCE

VOLUME 26

MEDIA AND LAW

SOCIOLOGY OF CRIME, LAW AND DEVIANCE

Series Editors: Mathieu Deflem and Derek M.D. Silva

Previous Volumes:

Volumes 1–5:	Jeffrey T. Ulmer
Volume 6:	Ethnographies of Law and Social Control, edited by Stacey Lee Burns, 2005
Volume 7:	Sociological Theory and Criminological Research, Views from Europe and United States, edited by Mathieu Deflem, 2006
Volume 8:	Police Occupational Culture: New Debates and Directions, edited by Megan O'Neill, Monique Marks and Anne-Marie Singh, 2007
Volume 9:	Crime and Human Rights, edited by Stephan Parmentier and Elmar Weitekamp, 2007
Volume 10:	Surveillance and Governance: Crime Control and Beyond, edited by Mathieu Deflem, 2008
Volume 11:	Restorative Justice: From Theory to Practice, edited by Holly Ventura Miller, 2008
Volume 12:	Access to Justice, edited by Rebecca Sandefur, 2009
Volume 13:	Immigration, Crime and Justice, edited by William F. McDonald, 2009
Volume 14:	Popular Culture, Crime and Social Control, edited by Mathieu Deflem, 2010
Volume 15:	Social Control: Informal, Legal and Medical, edited by James J. Chriss, 2010
Volume 16:	Economic Crisis and Crime, edited by Mathieu Deflem, 2011
Volume 17:	Disasters, Hazards and Law, edited by Mathieu Deflem, 2012
Volume 18:	Music and Law, edited by Mathieu Deflem, 2013
Volume 19:	Punishment and Incarceration: A Global Perspective, edited by Mathieu Deflem, 2013
Volume 20:	Terrorism and Counterterrorism Today, edited by Mathieu Deflem, 2015
Volume 21:	The Politics of Policing: Between Force and Legitimacy, edited by Mathieu Deflem, 2016
Volume 22:	Race, Ethnicity and Law, edited by Mathieu Deflem, 2017
Volume 23:	Homicide and Violent Crime, edited by Mathieu Deflem, 2018
Volume 24:	Methods of Criminology and Criminal Justice Research, edited by Mathieu Deflem and Derek M. D. Silva, 2019
Volume 25:	Radicalization and Counter-Radicalization, edited by Derek M. D. Silva and Mathieu Deflem, 2020

SOCIOLOGY OF CRIME, LAW, AND
DEVIANCE VOLUME 26

MEDIA AND LAW: BETWEEN FREE SPEECH AND CENSORSHIP

EDITED BY

MATHIEU DEFLEM
University of South Carolina, USA

and

DEREK M. D. SILVA
King's University College, Canada

United Kingdom – North America – Japan
India – Malaysia – China

Emerald Publishing Limited
Howard House, Wagon Lane, Bingley BD16 1WA, UK

First edition 2021

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

ISBN: 978-1-80071-730-5 (Print)
ISBN: 978-1-80071-729-9 (Online)
ISBN: 978-1-80071-731-2 (Epub)

ISSN: 1521-6136 (Series)

CONTENTS

PART III
REGULATING SPEECH ACROSS NATIONS

ABOUT THE CONTRIBUTORS

Lucia Bellucci, Department of Law "Cesare Beccaria", Università degli Studi di Milano, Italy, specializes in EU and international law and politics, media law and policy, and law and anthropology. She has published four books, including *La Sindrome ungherese in Europa* (2018), and wrote a chapter in *Culture in the Domains of Law* (2017).

Gabriela Capurro, Department of Community Health Sciences, University of Manitoba, Canada, specializes in health and risk communication, science communication, lived experiences of risk, and discourses of health and illness. Her work has appeared in peer reviewed journals such as *Canadian Journal of Sociology, Canadian Journal of Communication,* and *Science Communication*.

Mathieu Deflem, Department of Sociology, University of South Carolina, USA, specializes in the sociology of social control, terrorism, policing, sociology of law, popular culture and celebrity, and sociological theory. He is the Author of four books, including *The Policing of Terrorism* (2010) and *Sociology of Law* (2008).

Anne-Marie Gingras, Department of Political Science, Université du Québec à Montréal, Canada, specializes in political sociology, political communication, freedom of expression, and media representation of women in politics. She has published four books, including *Histoires de communication politique*: *Pratiques et état des savoirs* (Presses de l'Université du Québec, 2018).

Josh Greenberg, School of Journalism and Communication, Carleton University, Canada, is Professor and Director of the School. His research expertise is in the area of health and risk communication, crisis communication, and public engagement.

Kateryna Kaplun, School of Criminal Justice, Rutgers University-Newark, USA, specializes in cybercrime, surveillance, technology, white-collar crime, and mixed methods analyses. Her doctoral research examines the intersection of technology and gender, focusing on cyberstalking and cyberharassment statutes in the United States.

Sarah Lageson, School of Criminal Justice, Rutgers University-Newark, USA, studies criminal legal systems, privacy, surveillance, and technology. Her book, *Digital Punishment: Privacy, Stigma, and the Harms of Data-Driven Criminal Justice,* was published in 2020 by Oxford University Press.

Anthony Löwstedt, Department of Media Communications, Webster Vienna Private University, Austria, specializes in communication ethics, media theory,

and communication history and development. His recent work has appeared in the peer-reviewed journals *Communication Theory*, the *International Journal of Media and Cultural Politics*, and the *International Communication Gazette*.

Allyson M. Lunny, Department of Social Science, York University, Canada, specializes in hate studies, interdisciplinary discursive analysis, and socio-legal studies. She is the Author of *Debating Hate Crime: Language, Legislatures, and the Law in Canada* (2017) and articles in *Social & Legal Studies* and *GLQ: A Journal of Lesbian and Gay Studies*.

Tanner Mirrlees, Communication and Digital Media Studies, Ontario Tech University, Canada, specializes in the political economy of communication and digital media industries. He is the Author or Co-editor of five books, including *The Television Reader* (2012), *Global Entertainment Media* (2013), *Hearts and Mines* (2016), *Media Imperialism* (2019), and *EdTech Inc*. (2019).

Kimberly W. O'Connor, Department of Organizational Leadership, Purdue University Fort Wayne, USA, is an Associate Professor and Attorney whose research areas include social media and the law, cybersecurity, employment law, and data privacy. She serves as a consultant for K-12 digital safety programing and for social media-related matters in organizations.

Gordon B. Schmidt, Department of Organizational Leadership, Purdue University Fort Wayne, USA, researches the future of work including the gig economy, virtual leadership, and how social media impacts employee relations. He teaches courses in training, organizational behavior, and leadership. He has published and presented on his teaching innovations.

Devika Sethi, School of Humanities and Social Sciences, Indian Institute of Technology Mandi, India, specializes in Modern Indian history, specifically the history of decolonization in twentieth century India, book history, and the history of free speech and censorship. She is the Author of *War over Words: Censorship in India, 1930–60* (2019).

Derek M. D. Silva, Department of Sociology, King's University College at Western University, Canada, specializes in radicalization and violent extremism, policing, sport, and social control. His most recent work has appeared in the peer-reviewed journals *Punishment & Society*, *Sociological Forum, Race & Class*, and the *Sociology of Sport Journal*.

Natasha Tusikov, Department of Social Science, York University, Canada, studies the interactions among crime, regulation, and technology, platform and internet governance, and data governance. She is the Author of *Chokepoints: Global Private Regulation on the Internet* (2017) and she has been published in *Surveillance & Society* and *Internet Policy Review*.

LIST OF THE CONTRIBUTORS

Lucia Bellucci	Department of Law "Cesare Beccaria", Università degli Studi di Milano, Italy
Gabriela Capurro	Department of Community Health Sciences, University of Manitoba, Canada
Mathieu Deflem	Department of Sociology, University of South Carolina, USA
Anne-Marie Gingras	Department of Political Science, Université du Québec à Montréal, Canada
Josh Greenberg	School of Journalism and Communication, Carleton University, Canada
Kateryna Kaplun	School of Criminal Justice, Rutgers University-Newark, USA
Sarah Lageson	School of Criminal Justice, Rutgers University-Newark, USA
Anthony Löwstedt	Department of Media Communications, Webster Vienna Private University, Austria
Allyson M. Lunny	Department of Social Science, York University, Canada
Tanner Mirrlees	Communication and Digital Media Studies, Ontario Tech University, Canada
Kimberly W. O'Connor	Department of Organizational Leadership, Purdue University Fort Wayne, USA
Gordon B. Schmidt	Department of Organizational Leadership, Purdue University Fort Wayne, USA
Devika Sethi	School of Humanities and Social Sciences, Indian Institute of Technology Mandi, India
Derek M.D. Silva	Department of Sociology, King's University College at Western University, Canada
Natasha Tusikov	Department of Social Science, York University, Canada

INTRODUCTION: MEDIATING BETWEEN LIBERTY AND LAW

Mathieu Deflem and Derek M. D. Silva

A famous dictum by classical sociologist Émile Durkheim, now widely accepted as foundational to the study of society, is that social facts are coercive over the individual (Durkheim, 1982). The coercive nature of social facts cannot be misunderstood to mean that society precludes free will, but as applied in organic societies such as our own today, on the contrary, presupposes that individuals are free to do and say as they choose. The continued though altered co-existence of society and individual was so peculiar at the dawn of the twentieth century despite a growing culture of individualism that it is virtually synonymous with the birth of sociology. However, the coerciveness of social facts, even under conditions of a cult of the self, does mean that various modes of individual conduct of action and speech are not without consequences, including both more and less formal sanctions when things are said and done that violate expectations of the cultural and normative order. Human agency allows people to act freely, but not without cost.

With respect to the ideal and actualized balance between free speech and various forms of its control by law and other regulatory frameworks, we find ourselves in strange times. For ours is increasingly a society in which the vast majority of its members continually express a commitment to openness, diversity, and tolerance at the same time when limits and discouragements to speak as one wishes flourish widely as well. Such preventions and curtailments of speech most typically take on the form of calls to secure protections from harm. In the name of personal safety and well-being, these sanctions can go down to the finest detail such as in the case of microaggressions (Campbell & Manning, 2018).

Thus, it appears that right when diversity is celebrated and when a wide array of means of self-expression, especially by means of social media on the internet, are readily at one's hand-held disposal, the range of tolerated opinions seems to have narrowed in scope. Among the many examples, mention can be made of legal

Media and Law: Between Free Speech and Censorship
Sociology of Crime, Law and Deviance, Volume 26, 1–5

ISSN: 1521-6136/doi:10.1108/S1521-613620210000026001

and economic sanctions imposed on unpopular tweets and opinions expressed on the internet; the irony intrinsic to the mandated or at least expected use of pronouns others prefer; the refusal of comedians to perform on college campuses for fear they are not safe in the one space where the notion of safe spaces is preached and practiced arguably more than anywhere else; and the strange and seemingly growing cases of university professors who have been reprimanded, removed from their teaching, or even lost tenure when they spoke out about some matter others deemed to be objectionable.

Perhaps it is only paradoxical, but also expected, that the freedoms awarded in our age also invoke a renewed measure of repression. After all, only in a society dedicated to the liberty of all can one be offended or upset when some individuals act and speak in a manner that was not expected nor appreciated. Nonetheless, some noteworthy ironies persist. For one, many of the calls to suppress speech in the name of protection from harm have been taking place in our universities and on college campuses, in that institution and in those spaces, in other words, that are functionally devoted to thinking and expressing ideas. Additionally striking is that many of these calls to suppress or repress speech have come from students rather than from the higher university administration, when young people were traditionally among the first and most vocal to critique any censorship of speech and unconventional words and expressions, such as in the realm of popular music (Deflem, 2020).

All too ironically, moreover, a climate encouraging free inquiry and free speech today, more than ever, is a critical challenge in the university setting, which in the American context is expressed as liberal arts (Whittington, 2019). As a result, instead of academic freedom as a primary right to be protected and exercised, professional academics presently working in institutions of higher learning at times engage in avoidance and self-censorship (Chamlee-Wright, 2019). This rise of self-censorship among academics coincides with, and indeed may not be unconnected to, the rise of a public concern with diversity that is based on ascribed identity characteristics (especially race and gender) rather than achieved accomplishments and relevant perspectives related to expertise (Wood, 2019).

Today's concerns over free speech and censorship have found expression in a range of contemporary phenomena expressed in such terms as safe spaces and cancel culture. As sociologist Frank Furedi (2020) has recently argued, the practice of safe space has blended uncomfortably well with the social distancing that was called for because of the COVID-19 pandemic. As such, the quarantining practices that were enacted because of a dangerous virus mirrored the quarantining that was already advocated as a kind of self-isolation from the harm that might result from dangerous opinions. In this social climate that also promotes diversity and tolerance, the mere utterance of an unpopular opinion can lead someone to be reprimanded, suspended, fired from employment, and indeed canceled from social life itself.

That the discussed phenomena take place in university settings is all the more remarkable because they are squarely at odds with the university as a site of debate geared toward the formation of knowledge and wisdom (Hill, 2020). As such, we may be witnessing, in Durkheim's terms, a return to mechanical times in

which the mere expression of an opinion deemed harmful is sufficient to have one exiled from the social group. Perhaps even more problematic than an expressed opinion facing attempts at suppression and retaliation after they have occurred is that some opinions are ignored because they are too marginally position to be noticed or, arguably worst of all, unheard because they are not even thought or expressed. Paraphrasing musician and free speech advocate Frank Zappa, the fundamentalism of today's cancel culture undermines what could be learned by the development of "an unborn idea" (Eisner, 2016).

At no time in history have the means to speak been so expansive and the right to speak without fear been so curtailed as today. Technology has advanced to give everybody with access to a computer or phone a voice to speak. But when anything can be said by more readily available means than ever before, not everything can be said without fear. And, at the same time, as the media landscape has diversified and expanded in unprecedented ways, there rages a war on science (Silva, 2018) and a questioning of the traditional media under the heading of fake news. Strikingly, also, cancel culture is a phenomenon often addressed in popular media and news outlets (e.g., Henderson, 2019; Kato, 2020), but it has as yet failed to draw any dedicated social science attention, having to date only attracted one academic article specifically addressing the issue (Ng, 2020).

It is within the context of the changing understanding of freedom, rights, and responsibilities that the chapters in this volume of *Sociology of Crime, Law, and Deviance* examine some of the socio-legal issues involved, especially in connection with a diverse media landscape, ranging from traditional mass media to the increasingly important social networking sites. Covering a range of themes in terms of mediated forms of the right to speak and the limits that are or should be imposed thereon, the chapters are theoretically and methodologically diverse.

Part I of this book addresses several fundamental and conceptual issues with respect to the control of speech in a variety of spaces and institutions. Anthony Löwstedt seeks to develop a new conception of censorship from the standpoint of a duty to safeguard and promote various forms of diversity. Löwstedt argues that censorship so conceived cannot be eliminated but can and should be minimized when appropriate. Next, Kimberly W. O'Connor and Gordon B. Schmidt likewise focus on censorship, examining free speech protections in terms of the social media use in academia. They argue that the legal questions posed in this context are complex and constantly developing in an uncertain environment. Turning to the medical field, Gabriela Capurro and Josh Greenberg research news coverage about health risks with a special focus on (anti-)vaccination, antimicrobial resistance, and COVID-19. They find that risks are framed differently in connection with public health guidelines in terms of their character as being known, emerging, and novel.

Part II focuses on various aspects of the internet as a public sphere. Natasha Tusikov investigates the role of online payment platforms in cases of sex censorship and the regulation of lawful sexual content. On the basis of relevant corporate policies, the authors find that the censorship enacted by payment platforms is driven by market concerns of the online payment industry. Next, Tanner Mirrlees examined hate-content moderation by the companies Google, Apple, Facebook,

Apple, and Microsoft (GAFAM) with respect to the online presence and activities of the so-called alt-right. With these powerful companies defining hate speech, Mirrlees intriguingly suggests that the alt-right can be effectively controlled when hate speech is involved. Likewise concentrated on the online environment, Sarah Lageson and Kateryna Kaplun examine accusations publicly uttered on the internet by various state and non-state actors. They find that the expression of an online accusation itself is seen as a valuable objective, irrespective of additional consequences to search for a fair and just resolution.

The chapters of Part III, finally, discuss relevant issues of media and law in different national settings. Focused on the Canadian context, Anne-Marie Gingras offers a fascinating examination of freedom of expression and humor, with special attention to a case involving a stand-up comic in Canada. Gingras shows how the cases were dealt with a matter of human rights, displacing the issue from one of defamation to a matter of dignity and honor. Also pertaining to Canada, Allyson M. Lunny investigates the development of Canadian hate speech laws and the role of mass communications media therein. Influenced by technological advanced, Lunny argues that hate speech laws need to be attuned to changes that propel various forms of online hate. Turning to the European continent, Lucia Bellucci explores how media law in Hungary developed in an illiberal manner to encourage self-censorship and the prevention of freedom of expression. Bellucci argues that the conditions of the COVID-19 pandemic have been used to reinforce these speech-suppressing powers. Finally, Devika Sethi reports from research on the censorship of communist publications in setting of late-colonial India. With a focus on the General Communist Notification of 1932 that was enacted in British India, Sethi argues in favor of a more nuanced and multidimensional understanding of the actual workings of censorship laws, irrespective of their stated intent.

Collectively, the chapters in this volume present a useful mix of perspectives and themes to demonstrate the relevance of socio-legal issues involved with today's media landscape. Both politically as well as culturally, free speech and the control thereof in the media have in recent years clearly moved center stage in discussions on what must, should, and should not be said in the public sphere of ideas, tastes, and opinions, and are therefore deserving of our scholarly attention. In a world of cancel culture, alternative facts, fake news, (self-)censorship, edited art, hate speech, and career-ending tweets, this volume hopes to make a timely contribution useful for scholars and students in sociology, public policy, socio-legal studies, criminal justice, law, and other relevant social science disciplines.

REFERENCES

Campbell, B., & Manning, J. (2018). *The rise of victimhood culture: Microaggressions, safe spaces, and the new culture wars*. New York, NY: Palgrave Macmillan.

Chamlee-Wright, E. (2019). Self-censorship and associational life in the liberal academy. *Society*, *56*(6), 538–549.

Deflem, M. (2020). Popular culture and social control: The moral panic on music labeling. *American Journal of Criminal Justice*, *45*(1), 2–24.

Durkheim, E. (1982 [1895]). *The rules of sociological method*. New York, NY: The Free Press.

Eisner, K. (2016, July 6). 'Eat That Question: Frank Zappa in His Own Words' reveals a tireless foe of censorship. *The Georgia Straight*. Retrieved from https://www.straight.com/movies/732301/eat-question-frank-zappa-his-own-words-reveals-tireless-foe-censorship. Accessed October 5, 2020.

Furedi, F. (2020). Social distancing, safe spaces and the demand for quarantine. *Society*, *57*(4), 392–397.

Henderson, R. (2019, September 30). The atavism of cancel culture. Retrieved from https://www.city-journal.org/cancel-culture. Accessed October 5, 2020.

Hill, D. W. (2020). Communication as a moral vocation: Safe space and freedom of speech. *The Sociological Review*, *68*(1), 3–16.

Kato, B. (2020, July 10). What is cancel culture? Everything to know about the toxic online trend. *The New York Post*. Retrieved from https://nypost.com/article/what-is-cancel-culture-breaking-down-the-toxic-online-trend/. Accessed on October 5, 2020.

Ng, E. (2020). No grand pronouncements here…: Reflections on cancel culture and digital media participation. *Television & New Media*, *21*(6), 621–627.

Silva, D. M. D. (2018). Committing sociology in an age of 'fake news': Mass media, twitter, and the struggle for knowledge in the public sphere. In J. Michalski (Ed.), *Sociological theory in scientific perspective* (pp. 87–120). Hauppauge, NY: Nova Science Publishers.

Whittington, K. E. (2019). The value of pluralism in the academy. *Society*, *56*(6), 607–610.

Wood, P. (2019). Self-censorship and freedom. *Society*, *56*(6), 601–606.

PART I

SPACES AND INSTITUTIONS OF FREE SPEECH

CHAPTER 1

FIGHTING CENSORSHIP: A SHIFT FROM FREEDOM TO DIVERSITY

Anthony Löwstedt

ABSTRACT

Purpose – *It may be time to reformulate the rejection of censorship. Freedom is not the only opposite of, strategic resource against, or antidote to, censorship.*

Methodology/Approach – *This chapter argues against censorship with a Kantian-normative approach (the deontological position of the categorical imperative), using conceptual analysis, constructivism, and international legal scholarship, from the standpoint of a humanity-wide duty to safeguard and promote cultural diversity and biodiversity. Increasingly visible weaknesses of the argument against censorship from the utilitarian standpoint of freedom, a negative argument, can be avoided in this way.*

Findings – *Especially the neoliberal approach to freedom has no provisions against corporate and only little against copyright censorship, which are both becoming increasingly acute. Diversity, on the other hand, both biological and cultural, is argued to be instrumentally good, and intrinsically good, but the latter only if balanced by equality of basic rights.*

Originality/Value – *The resulting moral and legal imperatives are to support, safeguard, and promote diversity, and thus to minimize both censorship in culture and selection/elimination in nature, but only to minimize them, simply because they cannot themselves be eliminated. It is impossible to eliminate elimination. This becomes clear when one considers self- and soft censorship. At least in the wide sense, censorship is inevitable – but sustainable development is impossible without strict minimization of censorship.*

Media and Law: Between Free Speech and Censorship
Sociology of Crime, Law and Deviance, Volume 26, 9–23

ISSN: 1521-6136/doi:10.1108/S1521-613620210000026002

Keywords: Diversity; freedom; media; censorship; neoliberalism; sustainable development

INTRODUCTION

In the "post-truth" era, the liberal, utilitarian, and consequentialist justification of freedom of speech and expression is not (yet) crumbling, but it is weakened. Free markets and representative nation-state democracies have spread across most of the world since the end of the Cold War, but semi- or fully fascist governments and leaders are now again being elected around the world, as in Europe after World War I, with the public sphere as a site of entertainment, cults of personality, and conspiracy theories scapegoating the innocent. John Stuart Mill's classic argument against censorship, that truth and progress will prosper from fair and healthy competition in which all voices are given a chance to be heard still holds (Mill, 1989 [1859], pp. 20–22). But there are now more than six times as many people alive as in Mill's time, so there is even less time for all to be heard. Some inequalities have also increased since then. In any case, the media are not neutral arenas, but uneven playing fields, practicing their own novel forms of censorship, filtering, and marginalization. In the age of electronic media especially filtering and marginalization have taken over roles played by hard censorship in previous societies and civilizations. Media users often do not care much for the truth, but rather for receiving and providing entertainment, making money, rewarding exposure and vanity, peddling influence, and confirming bias and prejudice.

Mill's argument against censorship may only hold for an abstract, ideal society, and a utopia that can never be realized (Danaher, 2018; Stanley, 2018), so what to do with the here and now? In this chapter, a complex concept of diversity, including biodiversity, cultural diversity, and media diversity, will be presented as an ideal that can alternate with and complement the concept of freedom, including the freedoms of opinion, information, worship, speech, and expression, with which censorship has been contrasted traditionally.

We are used to opposing censorship with freedom, as, for instance, in the text of the First Amendment of the U.S. Constitution (1791): "Congress shall make no law respecting an establishment of religion, or prohibiting the free exercise thereof; or abridging the freedom of speech, or of the press…" or §19 of the Universal Declaration of Human Rights (1948):

> Everyone has the right to freedom of opinion and expression; this right includes freedom to hold opinions without interference and to seek, receive and impart information and ideas through any media and regardless of frontiers.

Originally, however, the concept of freedom was not necessarily an inclusive one. The land of the free was there for the free; but the unfree were just as essential. The slaves and the dispossessed, however, were not meant to enjoy that land or its provisions. Similarly, the very first laws banning censorship, in Sweden (1766) and Denmark–Norway (1770), were not introduced for inclusivist reasons. Rather,

"[i]n all three countries, leaders realized that they were going to have to share power and govern by manipulating and steering public opinion rather than suppressing it" (Laursen, 2005, p. 100).

The modern Western concept of freedom is to a large extent an elitist concept, one that has roots in a Graeco-Roman concept of freedom that also takes for granted slavery, as well as gender inequality and other forms of oppression indeed considering them as inevitable. There was a widespread agonistic, zero-sum competition approach to freedom in Classical Greece, with a progress-less belief that only a fixed quantum of freedom is possible and only a small number of people, at best only Greek men, can have some (Brunkhorst, 2000; Green, 1998, pp. 183–184).

There was, however, also a humanistic and egalitarian current, for instance in Stoicist ethics, which contradicted the more popular Aristotelian–Platonist–Epicurean elitism (Bloch, 1986, pp. 12–13; Werner, 1992, p. 11). The radical egalitarianism of early Christianity also contributed to a humanist current in later Western liberalism (Brunkhorst, 2000).

Thus, there are two ideals of freedom in the history of Western civilization: one exclusive – freedom for the chosen few; and another inclusive – freedom for all. The same word was used by idealists and activists, often with contradictory meanings and sometimes with mutually supporting ones, such as in the American and French revolutions of the eighteenth century, which contributed to the extraordinary success and popularity, and even a quasi-religious status of the term, especially in societies based on slave–labor economies, such as ancient Athens and Rome, western European colonies in the Americas, South Africa, and elsewhere, the independent United States, Brazil, Cuba, and western Europe itself. Until now, "freedom" has dominated as the opposite term and opposite concept of censorship, among many other evil or bad things, such as dictatorship, injustice, arbitrary rule, and oppression (De Dijn, 2020).

But earlier civilizations also developed arguments and attitudes against censorship, among them ancient Egypt. A case in point is *The Teachings of Ptahhotep,* a secular, ethical text composed around 1850 BCE (at the latest), which among other things suggests finding a balance between the right to privacy with the duty to find out and tell the truth, rather than with the principle of freedom of expression (Löwstedt, 2019, pp. 501–503). Ptahhotep also negated suppression of speech with what appears to be a right to speak truth to power:

> If you are a man who leads,
>
> Listen calmly to the speech of one who pleads;
>
> Don't stop him from purging his body
>
> Of that which he planned to tell. (Ptahhotep, 17; in Löwstedt, 2019, pp. 498–501)

Thus, there are several possible alternatives to the concept and ideal(s) of freedom that have been and can be used in the fight against censorship. The principle presented here is yet another one: the core concept of media diversity, situated within a wider context of cultural diversity, which, in its turn, is situated within an even wider context of biodiversity.

VARIETIES OF CENSORSHIP

Censorship may be analyzed as existing in three forms. Hard censorship comprises prohibition, law enforcement, destruction of messages, of means of message production, dissemination and reception, violence against communicators, including killing, torture, detention, threats, fines, and extortion. It is usually carried out by states and other attempted monopolies on violence, such as vigilante, terrorist, and organized crime groups.

There are, however, some aspects of it that do not appear as hard as others, for example, selective placements of government advertisements in media that are less critical of the government. Still, critical media can be bled out of existence slowly in this way. I would refer to this as "indirect hard censorship," rather than "soft censorship" as other authors have done (e.g., Lansner, 2014), since the long-term consequences can be the same as direct hard censorship. Forbidding the use of certain languages, such as the various bans on Kurdish in Turkish broadcasting, print, and education until 1991 and beyond (Associated Press, 2002), are also seemingly softer forms of hard censorship. They seem to be only censoring form and not content. But they can be just as destructive. Corporate censorship can also be hard, especially through threats, extortion, firing, and chilling effects, as for example in several media outlets throughout News Corp, the Murdoch empire (Jukes, 2012). In this sense, corporations can act or function as laws unto themselves, as "monopolies of silence," and as effective censors equivalent to assumed state, terrorist, or mafia monopolies on violence.

Soft censorship proper consists of editorial rejection, deletion of posts, accounts, and content editing, but no penalties beyond that, no fines, no prison sentences, and no direct physical harm. It is now typically carried out by privately-owned, profit-seeking companies, such as book and magazine publishers, newspapers, broadcasting stations, cable companies, internet service providers, search engines, and social media companies.

Self-censorship is perhaps the most pervasive of all. We stop ourselves from sharing, writing, saying, remembering, or even thinking things, we correct ourselves, largely because of learned mechanisms. Many of these mechanisms may be unconscious, taken over from previous generations, from authorities, employers, peers, and the media. At least morally, however, the author herself or himself is ultimately responsible.

According to Skjerdal (2010) self-censorship is prompted by commercial and economic pressures as well as political ones. Clark and Grech (2017) reported increasing unwarranted interference, fear, and self-censorship among a sample of 940 journalists in 47 Council of Europe member states and Belarus. Whereas a majority of journalists experienced severe unwarranted pressures as successful in gagging their own reporting, some 36 percent of respondents replied "…that the experience made them more committed to not engage in self-censorship" (Clark & Grech, 2017, p. 14). It is thus possible to resist self-censorship due to outside pressure at any time, but the price to be paid for that can also be very high. And then there is also self-censorship due to inside pressure. (Some researchers, however, define self-censorship as necessarily intentional, e.g. Bar-Tal, 2017.)

In this wide sense, censorship as a whole, including self-censorship, is pervasive and fundamental. It is as inevitable as selection/elimination in the evolution of life on this planet; and it relates to diversity similarly. Natural selection is like censorship. And media diversity is like genetic and behavioral variation. Only a few make it in nature, and only a few, apparently, make it in culture. Media contents, media organizations, media personalities, and ideas are like genes, organisms, and species. Most are ruthlessly weeded out. Only few are chosen.

In both cases, it is often impossible, and usually difficult, but human beings can have a say in what gets selected and saved, and what gets eliminated. And, we can have a say as to whether there will be more or less selection/elimination as a whole, whether there will be more or less diversity, both with regard to biodiversity and cultural diversity, including media diversity.

FORMS OF CENSORSHIP NOT READILY COUNTERED BY FREEDOM

The first "news filter," acting as partial or selective (and extremely successful) censor, according to Herman and Chomsky's "propaganda model" (1988), is "size, ownership, and profit orientation" of the disseminating news organizations. More than 99 percent of the population is filtered out because they do not have enough money to own such an organization. The owner has the ultimate power in the organization. If a single or just a few owners, whether private, state, or even public, own too many media or otherwise control too much of the media market, then healthy competition will be compromised. Therefore, antitrust, ownership, or pro-competition laws are enacted to ensure media pluralism.

But the owners of the largest media organizations argue against these laws on the premise that such laws disturb the purely free market and thus free competition and therefore amount to censorship. They argue that they should not be punished for being successful. The general experience has been, however, that tabloid commercialism and sensationalism led to the greatest ratings, and thus laws to end or prevent ownership concentration also serve to preserve quality media, especially critical, unbiased quality journalism. Due to the two arguments against concentration of media ownership – that it leads to censorship (whether state, public, or corporate) and deterioration of content quality, this form of censorship, legalized preventive measures against media ownership concentration, is widely accepted, as long as it is consistent and applies fairly and equally to all. Indeed, because it prevents the autocratic censorship by a state or corporate or any other form of monopoly (or oligopoly), this is possibly the most important and acceptable kind of censorship (Löwstedt & Palac, 2019, pp. 20–21), especially when public media ownership is weak. And its justification has much more to do with diversity than with freedom. Indeed, freedom can be used by both sides of the argument; both sides accuse the other of censorship.

Another far-reaching variety of hard censorship is the protection and enforcement of intellectual property rights (IPR): copyrights, patents, and trademarks.

Fines and prison sentences are meted out to individuals attempting to flout this form of censorship, especially when they attempt to profit from doing so, but sometimes copyright infringement is also used as an excuse to silence critical voices. Exceptions are also far-reaching and usually formulated in terms of education, research, and public interest. Justice and fairness, however, seem very far off, as the ones profiting from the protection and enforcement of IPR regimes are often not the authors, creative artists, and inventors themselves, but rather self-interested IPR managers, publishers, agents, and additional rights holders, who sometimes acquired those rights through deception or downright theft (Downes, 2006; EDRi, 2019; Humphreys & Simpson, 2018, p. 105; Voon, 2006). Again, by making a multitude of media contents and perspectives available, education, research, and the public are more likely to benefit. And again, diversity is a more helpful conceptual tool against this form of censorship than freedom is. The argument for freedom of information (as opposed to expression), however, may play an equal role to diversity in justifying the lifting of IPR censorship.

WHERE IT GETS MESSY: NEWS FILTERS AND VARIETIES OF MEDIA DIVERSITY

The other four news filters, according to Herman and Chomsky (1988), are advertising, sourcing, flak, and ideology. Together with size, ownership, and profit orientation, they constitute such a fine-meshed combination of filters that the effect is as destructive to communication as heavy-handed censorship can be. In fact, it is even more destructive, since the subjects of liberal capitalist democracies believe they are free and that their communication is free, when they really are not.

The influential propaganda model goes a long toward explaining censorship without state involvement or with only partial state involvement. However, it does have some problematic aspects, perhaps especially regarding what it is that is actually being filtered. The first filter is apparently of individuals (ownership) and media organizations (size and profit orientation); the second of news proper, that is, messages and content (advertising); the third of news sources, experts, pundits, and witnesses (sourcing); the fourth of media organizations again (flak); and the fifth of ideologies (ideology).

In view of this conceptual uncertainty, and of the fact that state censorship is not directly included among the news filters, except very partially through flak, Löwstedt and Palac (2019) suggest a different approach with a preliminary typology of "media diversities" which attempts to identify the seven most important forms and areas of media diversity.

- *Ownership and Control* – also including media that are not profit-oriented, both governmental and non-governmental.
- *Sponsorship* – mainly commercial advertising nowadays, but also including, for example, state sponsorship, or donations to charity or rights organizations with their own media outlets, such as the Red Cross/Red Crescent and

Amnesty International, or sponsorship for media paid for illegally and used by terrorist groups, such as Al-Qaeda, Islamic State, or Nationalsozialistischer Untergrund.

- *Type* – sensory impact on audiences (audio, visual, audiovisual, and haptic (sculpture and reliefs, Braille writing, and now also 5G Internet) media); levels of physical production and distribution; types of code and functions; electronic and non-electronic; analog and digital; decimal and sexagesimal; picto-, etho-, phono-, and prographic; mass, social, and interpersonal; fictional and non-fictional; quality and tabloid; single and multi-, one-way, and interactive media, etc. (Löwstedt, 2020).
- *Content* – can also be analyzed in various ways, for instance, with regard to choices of language, script, genre, style, norms, and values.
- *Employment* – workplace diversity is beneficial to organizations, but not always (total skin-colour-blindness, e.g., may be detrimental in case historical injustices and their after-effects are ignored or suppressed, yet it is still controversial to what extent affirmative action can be said to serve justice). In general, organizations and complex systems that fail are mostly ones that reject, ignore, and/or lack diversity (Page, 2011, pp. 167–195; Roberge & Van Dick, 2010).
- *Perspective and Opinion* – diversity of perspectives distinguishes good science, journalism, and art (Bertalanffy, 1968, pp. 239–242); Yet, certain informative perspectives are systematically avoided, ignored, and prevented by the media, and by those who control, regulate, and own them. Some subjects are presented from the same perspective repeatedly, such as the royals of Saudi Arabia, who have been predominantly portrayed as "reformers" for over 70 years by their mostly sympathetic observers at the *New York Times* (Al-Arian, 2017). Another unhelpful or unhealthy mechanism, equally harmful to the subtle bias provided by "quality media" (see also Kramer, 2016), is polarization. Tabloids and tabloid formats use extreme opinions to raise controversy and provide media spectacles, thus burying potentially crucial information. Deniers of anthropogenic climate change, for example, are often provided equal time as expert sources on TV news, or equal space in the print media, to the 97 percent of climate change scientists who do not live in denial. The winners are not only the climate change deniers and their sponsors and supporters. The tabloid media, especially, may thus gain a false appearance of being balanced and neutral. Rather than being compared to the fist fights between scorned and betrayed lovers on the Jerry Springer Show, their debate platforms on climate change may even seem balanced and fair (Collomb, 2014).
- *Censorship* – may be a form of media diversity like ownership or sponsorship; it may be seen as inevitable, especially if it includes soft and self-censorship, but in all complex societies so far, there is also hard censorship. If censorship is minimized, fine-tuned, and well-balanced, to counter incitement to violence and hate speech (with criminal law), invasion of privacy, concentration of ownership, copyright theft, defamation (with civil law), including also relevant aspects of soft and self-censorship, it will (at least in the long run) contribute to more diversity in the media overall. Unwise, insensitive, overbearing, and

oppressive censorship, the vast majority of instances of censorship in history so far, however, will lead to less diversity in the media.

In general, the more censorship there is, the less media diversity there will be. But there are exceptions. This is where it gets complicated. If radical hate speech is allowed, this can, in the long term, lead to less diversity, through the establishment of a government led by an emboldened radical hate group, such as in the early 1930s liberal Germany (C, 2018). In view of that particular experience, it is widely agreed today that intolerance must not be tolerated. Forms of "marginal" censorship, in order to safeguard diversities, especially the most vulnerable targets, against attacks by the hateful, the intolerant, and the reckless, are apparently necessary. Therefore, not even hard censorship can be abolished. But, just like selection/elimination, it can be reduced and minimized. Long-term media diversity can be used as equivalent to long-term freedom to justify censorship and even criminalization of hate speech and incitement to violence.

WHEN TO CENSOR?

Even in the most liberal countries, incitement to violence, hate speech, and other contents, such as child pornography or advertisement for forbidden products and services, are subject to legal censorship. The level of interference in communications, however, is kept low, in the interest of cultural diversity. The severity of the crime or threat presented by the content also plays a role, at least in the sphere of communication ethics. It is more important to ban ads for the sale of weapons of mass destruction than for the sale of marijuana.

The threat may consist of "clear and present danger" or its legal replacement in the United States, the 1969 *Brandenburg* v. *Ohio's* "imminent lawless action" test. But it has to be circumspect, and also take possible future trajectories into account, for instance, with regard to explosive xenophobia as in Europe in 2015, or to the destructive force of repressed historical injustice, as contested by the Black Lives Matter movement. The kinds of content also need to be taken into account. Especially political speech, and especially criticism of those most powerful, needs to be protected from censorship, as demanded repeatedly by the United States Supreme Court (Trager, Russomanno, Dente Ross, & Reynolds, 2014, pp. 76–78), and indeed already by Ptahhotep (see above).

To consider the video footage of the killing of George Floyd, the smartphone recording of the Minneapolis incident that led to worldwide demonstrations and uprisings against racism and police brutality, may be instructive for the need to be circumspect and flexible regarding whether to censor or not. Films that contain footage of real murders with impunity may reach clientele who get pleasure from watching such footage. It is obvious that by forbidding and banning dissemination, production, and possession of such material, lives may be saved, and murderers apprehended. People watching the footage of Floyd's slow and painful death for pleasure, keeping it as a souvenir or even as a trophy, similar to the postcards of lynchings of Black men a century ago, have been reported. Sharing

and showing it further could contribute to the entrenchment, encouragement, and even spread of murderous racism. Yet, most would argue that this drawback would still not make it worthwhile to ban the footage. It was certainly not originally recorded, produced, or disseminated to appeal to audiences (Brackett, 2020; Lewis, 2020; Löwstedt & Mboti, 2017), and in this case, the killers were apprehended, and the extent of their responsibility for the death of Floyd became apparent, because of the footage.

FURTHER PARALLELS BETWEEN BIODIVERSITY, CULTURAL, AND MEDIA DIVERSITY

Only sometimes, then, censorship seems necessary, least of all with regard to hard censorship, but still: even there. The conditions and requirements for justification, however, seem to differ, depending on circumstances. Now the ethics of biodiversity can come in handy. There too, circumstances can be decisive. The way in which life takes place and procreates in our biosphere makes selection/elimination a necessity. Every living being cannot be guaranteed a right to life. But as many different ones as possible can.

For Darwinist evolutionary theory during the nineteenth and twentieth centuries, although genetic mutations provide the primary source of evolutionary change, their evolutionary importance was considered less significant than that of the mechanism of natural selection. The basis for such reasoning was the assumption that mutations (regardless of their outcome) are unavoidably subject to natural selection. Similar to the industrial workings of the nineteenth and twentieth centuries factory, the mutations (or genetic variation) were imagined like the raw material and natural selection like the treatment or refinement or even just the marketing of that material. Mutations could also be illustrated as physical and productive labor, whereas selection and elimination were like managerial or editorial tasks. Furthermore, by steering and directing the random mutations toward more offspring, higher chances of survival, longevity, etc., the managerial work was and is imagined and ideologically constructed as intellectual and the most important, the most valuable work.

Natural selection, however, is more like Facebook, mutations and the resultant diversity like the uploads, the entirely user-generated contents of Facebook. Like Facebook, selection is neither creative nor productive at all, although it shapes and determines much of evolution and has the power to delete, both individual messages (genes) and entire accounts and groups (organisms and species). But, crucially, diversity, creativity, or productivity are not *necessarily* shaped by it. Everything does not have to be uploaded to Facebook, not even all digital media contents.

Kimura's theory of neutral mutations suggests that genetic mutations without so-called selective value, that is, without (directly or immediately) providing its organism carriers with enhanced or depreciated survival or procreation chances, are also the most commonly occurring kind of mutations (Kimura, 1985, 1994). Something similar seems to hold for cultural creativity and productivity. They are neither mainly about survival nor about spreading selfish genes.

Often, biodiversity and cultural diversity overlap. In many equatorial regions, where both biodiversity and cultural diversity are threatened the most today, the two are not just allies, but united. The disappearing indigenous Amazon societies, for example, are desperately trying to safeguard the diversity of plants and animals against brutally encroaching monocultures of various kinds and save themselves at the same time (Kolirin, 2020).

However, the two might also contradict rather than complement each other. There is not always a smooth transition from biodiversity to cultural diversity. Under certain conditions, a gain in cultural diversity may result in a net loss of biodiversity, or vice versa. If we should allow more different kinds of sushi to include whale-meat, for example, there would most likely be more (sushi) culture but less whale diversity, considering present levels of demand. If all human beings were now to leave Europe, for another example, cultural diversity would decrease but biodiversity would increase there (Löwstedt & Igropoulou, 2021).

Globalization, scientific agriculture, and technological progress are allowing ever-larger human populations, and biodiversity suffers immensely under the strain, but so does cultural diversity. Technological media development is currently dominated by two powerful institutions, the nation-state and the for-profit corporation, both of which, on balance, tend to suffocate cultural diversity. Since the 1970s the global population has doubled, but our rapidly globalizing society has simultaneously seen a possibly historically unparalleled deterioration of linguistic diversity (Austin & Sallabank, 2011, p. 2; Philippson, 2009, pp. 47–49). Although technological progress is undeniable, the media revolution has had very destructive effects on linguistic and cultural diversities.

According to Turin (2013), neither globalization nor any of its "tools" (including the Internet), but monolingualism, especially prevalent in the West or the G8 countries, is to blame for the current, catastrophic loss of living languages. Any language spoken by more than 10 million people is so because of imperialism. It is chiefly a political question, not a dictate of technological development, if a language will live or die. According to Kornai (2013), on the other hand, the Internet is to blame after all, since only less than five percent of the world's languages, so far, have achieved "digital ascent," and not many more will make it for much more than a century or so from now. "… [I]t is the final act of the Neolithic Revolution, with the urban agriculturalists moving on to a different, digital plane of existence, leaving the hunter-gatherers and nomad pastoralists behind" (Kornai, 2013, p. 10). Yet, I fail to see how the Internet will be *unable* to take in and allow and generate use of more than around five percent of all languages in the longer run. Wise political decisions could still reverse the trend of declining diversity. With this argument, I tend to agree more with Turin's anti-monolingualist approach, which may, however, underestimate the culturally destructive power of the unregulated global marketplace along with that of the nation-state.

In either scenario, not only linguistic, but also cultural diversity in general is now, paradoxically, under serious threat, although there are many more human beings, more levels and varieties of media, more means and channels of communication and expression at our disposal, and arguably more formal freedom and incentives to communicate, than ever before. Apparently, this paradox can

only be solved with a multidimensional approach to cultural diversity, that is, to diversities. There were corresponding losses of linguistic diversity due to the spread of phonographic writing, starting 5,000 years ago already, also along with outgrowths of military, political and economic power, and a simultaneous rapid increase of vocabulary of the written languages, meaning simultaneous increase of (another form of) linguistic diversity, in another dimension, so to speak. The current growth of artificial, formal, and constructed languages may also partially countervail destructive consequences of further impoverishment of "natural" languages (Petzold, 2015, pp. 164–169).

Diversities are not (yet) totally commensurable, that is, mutually comparable in quantitative terms. The multidimensional approach to diversity also contains dangers, such as the possibility of justification of cultural imperialism: "if English replaces 50 aboriginal languages, that's OK, because English is the richer language." (Therefore, linguistic diversity should never be reduced to or be dominated by vocabulary size; Philippson, 2009.) Uncritical progressivism, moreover, may threaten to de-politicize and turn media technology and media devices into "fetishes" with supernatural powers (Budka, 2013, pp. 29–31).

Another disturbing aspect of an uncritical approach to cultural diversity is that inequality may seem to enhance diversity. The argument goes like this: The more inequality a society manifests, for instance in income levels or life expectancies or levels of legal rights and privileges, the more diverse it is, and therefore the better it is (Wischermann & Thomas, 2008, pp. 7–8). Related potential or actual disadvantages of cultural diversity are lack of understanding, more misunderstandings, and lack of communication. These could lead to lack of efficiency, less trust, more suspicion, and loss of economic integration. The latter, however, are mainly short-term disadvantages. Once people who do not know each other's languages or cultures start working together, mutual understanding and trust will usually develop (Alesina, Devleeschauwer, Easterly, Kurlat, & Wacziarg, 2003; Patsiurko, Campbell, & Hall, 2012).

On the other hand, there are two kinds of positive aspects of cultural diversity. It can be *instrumentally good*: provide business opportunities and general enrichment. It can contribute to general tolerance, respect, and peace. Cultural diversity also makes us more intelligent, since a multiplicity of perspectives leads to more and better scientific knowledge (Phillips, 2014). Secondly, cultural diversity is arguably *intrinsically good*: "Variety is the spice of life," the saying goes. But it is actually much more than that. It is the flavor, the nutrition, and sometimes even a meaning of life, especially if cultural diversity is joined by biodiversity in a single concept of diversity. In current evolutionary theory, selection/elimination and variation are the two most basic principles. Still, diversity cannot be *the* meaning of life. Wischermann and Thomas (2008) demonstrate that additional values are necessary. Diversity is central, not only for humans. But so are other values, other meanings of life, such as rights and duties, freedom, empathy, solidarity, love, progress, or absence of suffering.

Nevertheless, diversity is often essential for survival, enhancing both adaptation and innovation, as well as intrinsically good. It has many more advantages than disadvantages (Page, 2011). Therefore, it seems not just ethical, but also intelligent

and strategically smart to minimize selection/elimination, especially of human individuals, families, communities, and also grand bio-units, such as habitats, species, and large populations, and to increase variation or diversity at all levels.

But there are some interesting, marginal limits here, too, for instance large and destructive parasite populations. One may argue that these cause deterioration of general bio- or cultural diversity in the long run, for example, the influenza virus of 1918 that wiped out 2 percent of the global human population (far more than World War I did) or HIV/AIDS or Covid-19. Again, one may argue similarly that keeping Nazi culture alive would not be good for cultural or media diversity because Nazis are opposed to media, political, cultural, and human diversity, in general, like the killer viruses are opposed to biodiversity.

This may be more than just an analogy. There might be an essential continuity between cultural diversity and biodiversity, despite occasional contrariness. Media freedom activists and anti-extinction campaigners may have more in common than they realize. For this and other reasons, there are concerted efforts to develop a reliable biocultural diversity index (see Harmon & Loh, 2005; Maffi & Woodley, 2010). If we can measure biocultural diversity in a fairly accurate way, then we can develop policies and strategies to achieve it and increase it and also find indications when it becomes urgent to reverse destructive developments. And the media will no doubt be both necessary and important parts of it.

In principle, nevertheless, diversity needs to be balanced by dignity and equal rights. In liberal, capitalist societies, especially the ownership and sponsorship filters greatly limit media diversity. Therefore, restrictions on concentration of media ownership and support for messages without commercial sponsorship are also necessary. Dogmatic, insensitive, and oppressive censorship, however, lead to less cultural and media diversity, and it is often difficult to strike a balance (Löwstedt & Palac, 2019). But there seems to be at least a theoretical possibility to develop a unified, co-evolutionary, and developmental diversity ethics that will help us navigate culturally as well as naturally.

FREEDOM AND DIVERSITY

One of the most celebrated formulations of Kant's categorical imperative, known as the "Formula of Humanity," is the following: "So act that you treat humanity, whether in your own person or in the person of any other, always at the same time as an end, never merely as a means" (Kant, 1785, p. 429, italics removed).

It may be interpreted as implying that each and every human being is a meaning of life proper, in an individual way. It rejects wholesale instrumentalizations of human beings, such as enslavement, and embraces the idea of diversity within humanity. Kant did not seem to see intrinsic value in non-human life developing in individual ways, but he did not deny it, either. We are now at an historical juncture where we need to allow nature to develop sustainably with or without us. It is therefore time to gently take a step beyond Kant, but also to understand sustainable development as implying development without loss of either biodiversity or cultural diversity (Löwstedt & Igropoulou, 2021).

As long as slavery, forced labor, human trafficking, exploitation (also of nature), and other gross rights violations, continue to exist, the ideal of freedom will be necessary. Inclusive freedom is not utopian; it happens every day. But current challenges to humanity demand additional ideals, and diversity may not only complement freedom. It may also take over some of the arguments for which freedom is now used, including arguments against censorship, ideally not displacing freedom – but relieving it, supporting it, and giving it a partly new trajectory. The rejection and minimization of censorship is a possible instance of this. Freedom and diversity will both help save humanity and the planet.

ACKNOWLEDGEMENTS

For inspiration and useful comments, I thank Kim Cooper, Patricia Löwstedt and Cornelius Zehetner.

REFERENCES

Al-Arian, A. (2017, November 27). Seventy years of the New York Times describing Saudi royals as reformers. *Jadaliyya*. Retrieved from https://www.jadaliyya.com/Details/34727

Alesina, A., Devleeschauwer, A., Easterly, W., Kurlat, S., & Wacziarg, R. (2003). Fractionalization. *Journal of Economic Growth*, *8*(2), 155–194.

Associated Press. (2002, November 21). Turkey allows broadcasting of Kurdish-language shows. *The New York Times*. Retrieved from https://www.nytimes.com/2002/11/21/world/turkey-allows-broadcasting-of-kurdish-language-shows.html

Austin, P. K., & Sallabank, J. (2011). Introduction. In P. K. Austin & J. Sallabank (Eds.), *Cambridge handbook of endangered languages* (pp. 1–24). Cambridge: Cambridge University Press.

Bar-Tal, D. (2017). Self-censorship as a socio-political-psychological phenomenon: Conception and research. *Advances in Political Psychology*, *38*, 37–65. doi:10.1111/pops.12391

Bertalanffy, L. (1968). *General system theory: Foundations, development, applications*. London: Braziller.

Bloch, E. (1986). *Natural law and human dignity*. (D. J. Schmidt, Trans.). Cambridge, MA: MIT Press.

Brackett, L. T. (2020, June 20). American lynching souvenirs make ugly return. *Tacoma News Tribune*. Retrieved from https://www.thenewstribune.com/opinion/op-ed/article243657907.html

Brunkhorst, H. (2000). Equality and elitism in Arendt. In D. Villa (Ed.), *The Cambridge companion to Hannah Arendt* (Cambridge Companions to Philosophy, pp. 178–198). Cambridge: Cambridge University Press. doi:10.1017/CCOL0521641985.010

Budka, P. (2013). Digitale Medientechnologien aus kultur – und sozialanthropologischer Perspektive: Überlegungen zu Technologie als materielle Kultur und Fetisch. *Medien und Zeit*, *28*(1), 22–34.

C, S. (2018, July 19). How the Nazi party came to power in a democratic country. *Medium*. Retrieved from https://medium.com/@QSE/the-time-for-never-again-has-passed-americans-are-complicit-in-genocide-as-defined-by-the-un-940b368e63e5

Clark, M., & Grech, A. (2017). *Journalists under pressure: Unwarranted interference, fear and self-censorship in Europe*. Strasbourg: Council of Europe. Retrieved from https://rm.coe.int/168070ad5d

Collomb, J.-D. (2014). The ideology of climate change denial in the United States. *European Journal of American Studies*, *9*(1). doi:10.4000/ejas.10305

Danaher, J. (2018, November 14). Mill's argument for free speech: A guide. *Philosophical Disquisitions*. Retrieved from https://philosophicaldisquisitions.blogspot.com/2018/11/mills-argument-for-free-speech-guide.html

De Dijn, A. (2020). *Freedom: An unruly history*. Cambridge, MA: Harvard University Press.

Downes, D. M. (2006). New media economy: Intellectual property and cultural insurrection. *The Journal of Electronic Publishing*, *9*(1). doi:10.3998/3336451.0009.103

EDRi (2019, January 15). Copyright as a tool of censorship. *European Digital Rights*. Retrieved from https://edri.org/our-work/copyright-week-2019-copyright-as-tool-of-censorship/

First Amendment (1791). Bill of Rights of the United States Constitution. Washington, D.C.: National Archives.

Green, P. (1998). *Classical bearings: Interpreting ancient history and culture*. Berkeley, CA: University of California Press.

Harmon, D., & Loh, J. (2005). A global index of biocultural diversity. *Ecological Indicators*, *5*, 231–241. Retrieved from http://www.csin-rcid.ca/downloads/loh_harmon_ei.pdf

Herman, E. S., & Chomsky, N. (1988). *Manufacturing consent: The political economy of the mass media*. New York: Pantheon.

Humphreys, P., & Simpson, S. (2018). *Regulation, governance and convergence in the media*. Cheltenham: Elgar.

Jukes, P. (2012). *The fall of the house of Murdoch*. London: Unbound.

Kant, I. (1785). Grundlegung zur Metaphysik der Sitten [Groundwork of the metaphysics of morals]. In I. Kant (Ed.), *Practical philosophy* (M. Gregor, Trans. and Ed., pp. 37–108). Cambridge: Cambridge University Press.

Kimura, M. (1985). *The neutral theory of molecular evolution*. Cambridge: Cambridge University Press.

Kimura, M. (1994). *Population genetics, molecular evolution, and the neutral theory: Selected papers*. Chicago, IL: University of Chicago Press.

Kolirin, L. (2020). 'The trees are my grandparents': The Ecuador tribe trying to save its culture. *The Guardian*, April 16. Retrieved from https://www.theguardian.com/world/2020/apr/16/the-trees-are-my-grandparents-the-ecuador-tribe-trying-to-save-its-culture

Kornai, A. (2013, October). Digital language death. *PLoS ONE*, *8*(10), e77056.

Kramer, M. (2016, July 26). To get a different perspective on the media, we went beyond New York. Way, way beyond, *Poynter*. Retrieved from: https://www.poynter.org/tech-tools/2016/to-get-a-different-perspective-on-the-media-we-went-beyond-new-york-way-way-beyond/

Lansner, T. (2014). *Soft censorship, hard impact: A global review*. Paris: WAN-IFRA World Association of Newspapers and News Publishers.

Laursen, J. (2005). Censorship in the Nordic countries, ca. 1750–1890: Transformations in law, theory, and practice. *Journal of Modern European History*, *3*(1), 100–121.

Lewis, C. (2020, June 6). George Floyd was lynched. *Open Democracy*. Retrieved from https://www.opendemocracy.net/en/opendemocracyuk/george-floyd-was-lynched/

Löwstedt, A. (2019). Do we still adhere to the norms of ancient Egypt? A comparison of Ptahhotep's communication ethics with current regulatory principles. *International Communication Gazette*, *81*(6–8), 493–517. doi:10.1177/1748048518802241

Löwstedt, A. (2020). Developmental stages of information and communication technology. *Communication Theory*. doi:10.1093/ct/qtaa015

Löwstedt, A., & Igropoulou, D. (2021, forthcoming). Between rights and diversities: Can the regulation of communication help prevent climate change and promote sustainable development? In J. Servaes & M. J. Yusha'u (Eds.), *Global handbook of communication and sustainable development*. Basingstoke: Palgrave-MacMillan.

Löwstedt, A., & Mboti, N. (2017). Media racism: Beyond modernity and postmodernity. *International Journal of Media and Cultural Politics*, *13*(1–2), 111–130. doi:10.1386/macp.13.1-2.111_1

Löwstedt, A., & Palac, J. (2019). Media and other biocultural diversities. In K. Holtzmann, T. Hug, & G. Pallaver (Eds.), *Das Ende der Vielfalt? Zur Diversität der Medien* (pp. 19–34). Innsbruck: Innsbruck University Press.

Maffi, L., & Woodley, E. (2010). *Biocultural diversity conservation: A global sourcebook*. London: Earthscan.

Mill, J. S. (1989 [1859]). *On liberty and other writings*. Cambridge: Cambridge University Press.

Page, S. E. (2011). *Diversity and complexity*. Princeton, NJ: Princeton University Press.

Patsiurko, N., Campbell, J. L., & Hall, J. A. (2012). Measuring cultural diversity: Ethnic, linguistic and religious fractionalization in the OECD. *Ethnic and Racial Studies*, *35*(2), 195–217.

Petzold, T. (2015). Human-algorithmic scaffolding. In W. de Been, P. Arora, & M. Hildebrandt (Eds.), *Crossroads in new media, identity and law: The shape of diversity to come* (pp. 156–171). Basingstoke: Palgrave Macmillan.

Philippson, R. (2009). *Linguistic imperialism continued*. Hyderabad: Orient BlackSwan.

Phillips, K. W. (2014, October 1). How diversity makes us smarter. *Scientific American*. Retrieved from https://www.scientificamerican.com/article/how-diversity-makes-us-smarter/?WT.mc_id=SA_FB_ARTC_OSNP

Roberge, M.-É., & Van Dick, R. (2010). Recognizing the benefits of diversity: When and how does diversity increase group performance? *Human Resource Management Review*, *20*(4), 295–308.

Skjerdal, T. (2010). Justifying self-censorship: A perspective from Ethiopia. *Westminster Papers in Communication and Culture*, *7*(2), 98–121.

Stanley, J. (2018, September 4). What Mill got wrong about freedom of speech. *Boston Review*, *44*. Retrieved from http://bostonreview.net/politics-philosophy-religion/jason-stanley-what-mill-got-wrong-about-freedom-of-speech

Trager, R., Russomanno, J., Dente Ross, S., & Reynolds, A. L. (2014). *The law of journalism and mass communication* (4th ed.). Los Angeles, CA: Sage.

Turin, M. (2013, December 3). Globalization helps preserve endangered languages. *YaleGlobal*. Retrieved from http://yaleglobal.yale.edu/content/globalization-helps-preserve-endangered-languages

Voon, T. (2006). UNESCO and the WTO: A clash of cultures? *International and Comparative Law Quarterly*, *55*(3), 635–652. Retrieved from http://www.worldtradelaw.net/articles/voonunesco.pdf

Werner, J. (1992). Zur Fremdsprachenproblematik in der griechish-römischen Antike. In K. W. Müller, K. Sier, & J. Werner (Eds.), *Zum Umgang mit fremden Sprachen in der griechisch-römischen Antike* (Palingenesia, Vol. 36, pp. 1–20). Stuttgart: Steiner.

Wischermann, U., & Thomas, T. (Eds.). (2008). *Medien – Diversität – Ungleichheit: Ausgangspunkte* (pp. 7–20). Wiesbaden: VS Verlag für Sozialwissenschaften. doi:10.1007/978-3-531-90860-1_1

CHAPTER 2

FREE SPEECH AND SOCIAL MEDIA IN ACADEMIA

Kimberly W. O'Connor and Gordon B. Schmidt

ABSTRACT

Purpose – *This chapter explores the topic of free speech protections and social media use in academia through an examination of the current legal landscape as it applies to various stakeholders on university campuses in the United States. The authors focus this examination primarily on public universities.*

Methodology/Approach – *Legal research methods were utilized, including an analysis of relevant United States federal and state laws, case law, and secondary sources such as law reviews. Non-legal sources, such as academic journals, were also reviewed, with particular emphasis on topics such as university policies, tenure protections, academic freedom, as well as current events.*

Findings – *The law regarding personal social media communications in a university setting is a series of complex and interconnected legal questions. Courts are still flushing out how free speech protections, personal social media use, and other relevant legal protections (e.g., employment law) may interface in a university-related case. Outcomes of cases are highly fact driven, and legal precedent is still being established.*

Originality/Value – *This chapter offers a comprehensive examination of the topic of free speech and social media use in United States academia by (1) examining legal protections as applied to various stakeholders on a college campus and (2) analyzing the current legal landscape of social media cases involving universities.*

Keywords: Free speech; social media; employment law; universities; academic freedom; civility

Media and Law: Between Free Speech and Censorship
Sociology of Crime, Law and Deviance, Volume 26, 25–41

ISSN: 1521-6136/doi:10.1108/S1521-613620210000026003

INTRODUCTION

Tenured Indiana University (IU) professor Eric Rasmusen, PhD, came under fire in late 2019 for his tweet suggesting that men were more intelligent than women and also questioned the role of women in academia. Despite public outcry from university alumni and benefactors, as well as current students who painted the words "Fire Eric Rasmusen" on a university bridge, Rasmusen kept his job (Herron, 2019, para. 1). IU Provost Lauren Robel wrote in a letter, which was later shared on social media, that Rasmusen would not be fired "for his posts as a private citizen, as vile and stupid as they are, because the First Amendment of the United States Constitution forbids us to do so" (Herron, 2019, para. 5).

In early 2020, the University of Florida rescinded an offer of admission to high school senior Liberty Woodley. Two years prior to her admission, when Woodley was 16 years old, she posted on Instagram, "I really try so hard not to be a racist person, but I most definitely am … & it gets the best of me sometimes." In the same post, she complained that two black girls sitting behind her were "retarded" and that they "won't shut the fuck up" and "i want to punch them," and she wrote that they were "most definitely crackwhores" and that people like them "do nothing for society" (Postal & Martin, 2020, para. 18). In rescinding Woodley's admission, a University of Florida representative explained, "The first amendment gives you the right to speak freely but the first amendment does not give you the right of admission to the University of Florida" (Rogers, 2020, para. 12).

In June 2020, Purdue University sophomore, Maxwell Lawrence, was dismissed from the university for racist statements he posted on social media, including a Tik Tok video where he pretended to run over Black Lives Matter protesters (Wilkins, 2020; WTHR, 2020). Additionally, Lawrence posted a racist meme that referred to the lynching of black men, and he used racial slurs in an Instagram group chat (Smith & Weliever, 2020). Purdue's President, Mitch Daniels, rejected a determination made by some university officials that Lawrence's speech was protected by the First Amendment, and instead Daniels permanently expelled Lawrence. In making this decision, Daniels citied a university policy that permitted expulsion when students posed "a threat to safety or the well-being of the university community" (WTHR, 2020, para. 6).

Each of the aforementioned cases involves a public university, free speech protections, and personal social media use; however, in each case the free speech analysis resulted in very different disciplinary outcomes. These situations therefore serve to illustrate how highly fact-driven free speech and social media cases are and how cases which occur in a university setting present a series of complex and interconnected legal questions. In fact, courts of law are still flushing out how free speech protections, personal social media use, and other relevant legal protections (e.g., employment law) may interface in a given case. Universities, however, often have additional and unique considerations such as civility policies, tenure protections, public versus private status, and of course academic freedom concerns. Therefore, this chapter seeks to address the topic of free speech protections and social media use in academia by examining various stakeholders on a university campus (faculty, admissions offices, students, and student athletes).

We will examine current events, legislation, case law, and academic literature to analyze and clarify the current legal landscape within the United States. We focus our examination primarily on public universities, as the analysis of free speech protections and university policies can create the most challenging legal questions.

ACADEMIA MEETS THE FIRST AMENDMENT

In the age of social media, First Amendment free speech is being applied in new and different ways. This is especially true as social media use is a highly prevalent means of communication in the United States. In 2019, 246 million Americans used social media to regularly post pictures, opinions, newsworthy events, like or comment on others' posts, or send messages to one another (Clement, 2020). This represents approximately 80% of the entire United States population. Of that group, 90% of American adults between 18 and 29 years old regularly use social media (Cooper, 2020).

The First Amendment guarantees that Congress shall make no law abridging the freedom of speech (U.S. Constitution, First Amendment, n. d.). However, not all speech is protected. Fighting words, obscenity, child pornography, fraud, and defamation are all examples of language that does not receive First Amendment protection. In addition, the First Amendment governs only governmental limitations on speech (*Nyabwa* v. *Facebook*, 2018). Thus, private entities are not bound by free speech limitations, unless it's the result of state law or contract provisions (Sarabyn, 2010). With regard to private universities, the law is quite clear that they are largely unrestricted when censoring the speech of their faculty, students, and staff, since private universities are not governmental actors (FIRE, 2015; Sarabyn, 2010). For public universities, however, the legal landscape is quite different and in many cases is still evolving with regard to social media-related speech.

Tenured and Untenured Faculty

There is a long line of established court precedent which states that free speech protections exist for public sector employees who are speaking as private citizens about "matters of public concern" (defined as speech "relating to any matter of political, social, or other concern to the community") (*Connick* v. *Myers*, 1983, pp. 142 & 146; *Garcetti* v. *Ceballos*, 2006). Therefore, free speech protections for public sector employees are rather limited in scope. Moreover, even if the subject of the speech involves a matter of public concern, public universities can still regulate speech that hinders the "efficiency of the public services it performs through its employees" (*Leary* v. *Daeschner*, 2000; *Pickering* v. *Bd. of Educ.,* 1968, p. 568). In such circumstances, a balancing test is used to determine if the employee's free speech interests outweigh the efficiency interests of the government as the employer (*Gillis* v. *Miller*, 2017; *Pickering v. Bd. of Educ.,* 1968).

In *Higbee* v. *Eastern Michigan University* (EMU, 2019), a 2019 faculty social media case, the *Pickering* balancing test was utilized when a professor was suspended for his social media post. The post in question referred to a 2016 incident

involving racist graffiti on campus, which was followed by a student-organized peaceful protest. Though it took a year to identify and locate the offender, eventually a former student of the university was arrested for the vandalism. Following the student's criminal arraignment, Professor Higbee posted on "EMU Talk," a public Facebook group about the university, the following statement:

> EMU administrators, a small group of well paid white guys in suits (plus one woman and a few lower level "HN in C" [*read:* Head Negro In Charge] functionaries), lacked the insight to imagine that they could ever, possibly, be remotely seen as responsible for institutional racist practices. And so they continued to act as the aggrieved party, needlessly alienating students who objected to racism. Why EMU officials, earning six figures or more, took this stance can only be explained by a combination of 1. ignorance about what racism is, 2. overconfidence that they are the good guys, 3. a lack of knowledge of EMU specifically and of higher education generally. (*Higbee* v. *EMU*, 2019, p. 698)

Professor Higbee was suspended without pay for one semester for his post, and he subsequently filed a cause of action claiming retaliation. Prior to this case, courts had already recognized that the right to protest racial discrimination is indeed a matter of public concern. The court in Professor Higbee's case determined that, by using a public forum to comment about recent racial incidents, Professor Higbee was not acting within his official duties as a professor. He therefore had free speech protection for his post because, according to the court, the post addressed his perceptions of the university's "institutional racist practices," and Professor Higbee was speaking as a private citizen about them (*Higbee* v. *EMU*, 2019, p. 699). Moreover, because the protests the year before were peaceful, and no additional interruption of university operations could be tied to Professor Higbee's post, EMU could not meet its burden of showing that the university's efficiency interests were interrupted. Thus, they did not outweigh Professor Higbee's substantial interest in speaking about a matter of public concern, and Higbee's Facebook post amounted to protected speech (*Higbee* v. *EMU*, 2019).

Faculty Social Media Posts: Newsworthy Cases

Cases such as *Higbee* v. *EMU* are few and far between. To date, there is very little legal precedent that specifically deals with personal social media use by public university employees. However, many newsworthy cases have made headlines, even if they did not end up in court. For example, Charles Negy, a University of Central Florida (UCF) psychology professor, took to Twitter to post comments about black privilege during the George Floyd protests. He posted,

> Sincere question: If Afr. Americans as a group, had the same behavioral profile as Asian Americans (on average, performing the best academically, having the highest income, committing the lowest crime, etc.), would we still be proclaiming "systemic racism" exists? (Comas, 2020, para. 7)

Likewise, in response to a Navy ship yard shooting that killed 13 people, University of Kansas journalism professor, David Guth, posted on Twitter, "#NavyYardShooting The blood is on the hands of the #NRA. Next time, let it be YOUR sons and daughters. Shame on you. May God damn you" (Jaschik,

2013; Levy, 2014. p. 78). Additionally, Fresno State University English professor, Randa Jarrar, called former first lady Barbara Bush "an amazing racist" on social media and subsequently added, "I'm happy the witch is dead. can't wait for the rest of her family to fall to their demise the way 1.5 million iraqis have. byyy-eeeeeee" (LoMonte & Jadon, 2018; Woosten & Wong, 2018, para. 5).

In all of the aforementioned cases, the faculty members kept their jobs (Guth retired from the University of Kansas in December 2019), despite public outcry, university investigations, campus protests, and signed petitions calling for termination in each case (Appleton, 2018; Campus Speech Database, 2013; Dguth, n.d.; Harris, 2020). Each university instead chose to err on the side of free speech protection in making their decisions (Comas, 2020; LoMonte & Jadon, 2018) (UCF, however, recently launched an investigation of Charles Negy on "discriminatory harassment" charges stemming from separate comments he allegedly made in the classroom) (Harris, 2020, para. 6). Notably, Jarrar also subsequently flaunted her tenure online to her critics by posting, "I will never be fired" (Woosten & Wong, 2018, para. 14).

Statements such as Jarrar's lead to the question of whether some faculty members might misunderstand the scope of their tenure protections. From a legal perspective, tenure is simply a contractual right to due process. In faculty disciplinary cases, tenure protection amounts to a shift in the burden of proof to the university to demonstrate that the faculty member is no longer meeting the institution's standards (Workplace Diversity, 2019). Thus, tenure should never be interpreted by faculty to be a foolproof shield that protects them from university discipline in free speech cases or otherwise (O'Connor & Schmidt, 2019). At most, tenure presents a set of minimum procedural requirements (i.e., notice and an opportunity to be heard) before a faculty member can be dismissed (Robson, 2014). This is in contrast with untenured faculty, who are typically contract employees, and universities may simply decline to renew the faculty member's contract in disciplinary cases (Hawke, 1997).

Still, even in cases where egregious statements are posted, universities might be reluctant to discipline faculty (tenured or not). Given the evolving legal landscape, it is possible that some administrators and faculty members may not fully understand free speech standards as they apply to social media cases. This may have been the case for Professor Rasmusen at IU, as the provost's statement that the Constitution expressly "forbids" them from firing Professor Rasmusen is inherently flawed. To date, there is no known legal precedent to establish that Rasmusen's sexist speech would constitute a matter of public concern. Furthermore, it would be a question of fact for the court to decide as to whether the public outcry that followed his postings (from the university alumni and benefactors, as well as the students who painted the words "Fire Eric Rasmusen" on a university bridge) amounted to a disruption of university operations (Herron, 2019).

Consequently, IU may have stopped their free speech analysis short when they reached the conclusion that, because Rasmusen spoke as a private citizen in his 2019 tweets, he was protected from university discipline (Herron, 2019). In an interesting twist, however, IU may have changed their free speech analysis by the

start of the 2020–2021 academic year. Right before the fall 2020 semester classes began, Professor Rasmusen took to Twitter again and posted negative comments about homosexual men and George Floyd, and he also posted about female students "really showing off their legs" (Kress, 2020, para. 4). This time, IU immediately took disciplinary action and placed Professor Rasmusen on unpaid leave for the 2020–2021 academic year (Kress, 2020).

University Policy Considerations: Academic Freedom and Civility

In faculty-related social media cases, university policy considerations may also apply. For example, universities are naturally dedicated to the free exchange of ideas. Academic freedom therefore stands for the idea that research and classroom discussion are protected activities (*Buchanan* v. *Alexander*, 2019). However, academic freedom and First Amendment free speech are not coextensive (Bérubé, 2019; *Buchanan* v. *Alexander*, 2019). Courts of law have clearly articulated that protection for speech related to academic freedom has limits. "Students, teachers, and professors are not permitted to say anything and everything simply because the words are uttered in the classroom context" (*Buchanan* v. *Alexander*, 2019, p. 852). One quite prominent argument, therefore, is that academic freedom ought not to apply to faculty members' personal social media posts if they do not concern content that is academically related (LeRoy, 2015; O'Connor & Schmidt, 2019).

Most universities also have civility policies that address rude or disrespectful speech, including extramural online speech (LeRoy, 2015; O'Connor & Schmidt, 2019). The focus of civility policies is not to regulate the content of speech itself, but to focus on the manner in which the content is delivered. University civility policies therefore are designed to promote a respectful and inclusive campus environment, where all members of the campus community feel safe to express themselves (LeRoy, 2015). The intent is that civility and free speech compliment, not oppose, one another (Thorne, 2013). Faculty members are often specifically called upon by universities to be examples of civility, "making clear to their students that civility and tolerance are hallmarks of educated men and women" (AAUP, n.d., para. 12). However, sometimes the content of faculty social media posts is deemed by the university to lack civility, especially when the students themselves are the ones who object to the content (O'Connor & Schmidt, 2019).

One recent example of a public university that utilized the language of their civility policy to justify the termination of two faculty members for their social media posts – one instructor and one tenured faculty member – is the University of Louisiana at Monroe (ULM) (US News, 2020). Though the university itself has not released the names of the faculty members, students widely shared screenshots of nursing professor Mary Holmes' post "Thank God for our President. He takes no salary for his loyal commitment. Not like the monkey before him," an apparent reference to President Obama (Parker, 2020, para. 4). Students also shared biology instructor Dennis Bell's request for white people to document "senseless and racist attacks on white people by blacks," in an apparent reference to the George Floyd riots, as well as a post from 2015 regarding the word "n****r," where he posted,

> […] I use the word freely as I want to because this is America and you can't tell me what I can or cannot say. I don't like the word because it has become associated with hatred, so I don't use it very often…. (Parker, 2020, para. 17)

ULM's public response was that the faculty members' posts did not "reflect the values of this institution or civilized society" (Associated Press, 2020, para. 3). Therefore, even if they had protection for their posts under the First Amendment, it can certainly be said that the faculty members' opinions did not reflect ULM's values of mutual respect and inclusivity. Reportedly, Dennis Bell, the continuing lecturer, was immediately terminated from ULM. Likewise, ULM administration is purportedly (at the time of this writing) in the process of terminating tenured professor Mary Holmes (Bolden, 2020).

UNIVERSITY ADMISSIONS AND THE IMPACT OF CANCEL CULTURE

The policies of academic freedom and civility that govern faculty also govern the entire campus community. Regarding university admissions, institutional academic freedom principles allow a university to select the students that are best suited for joining its community of scholars (Abrahamson, 2020). And, although students who apply to public universities certainly have free speech rights (*Tinker* v. *Des Moines School Corp*., 1969 – holding that students do not shed their Constitutional rights at the schoolhouse gate), a public university may revoke an offer of admission if an incoming student's behavior or actions violate the institution's civility policy, or cause a substantial disruption to university operations (Jaschik, 2020; *Tinker* v. *Des Moines*, 1969). In fact, it has become a common practice for universities to include revocation language in the students' acceptance letters to put the students on notice of this possibility (Sundquist, 2017).

The University of Florida certainly exercised this right when they rescinded their offer of admission to Liberty Woodley because of her social media statement that she is "most definitely" a racist person (Postal & Martin, 2020, para. 12). The university's free speech analysis, as evidenced in the university's revocation letter, comports with prior court precedent regarding student free speech on college campuses (Papandrea, 2017). In other words, students likely have First Amendment protection over racist speech that does not threaten anyone or cause a substantial disruption of university operations (Hudson, 2018; Pierce, 2015). But, such utterances are offensive and most certainly can negatively impact their admission or enrollment status (Levin, 2020).

Notably, Liberty Woodley's case is part of a growing trend where students are targeted as part of the "cancel culture" movement. Cancel culture is a term that refers to group shaming on social media – or a new trend of publicly calling out others online for behaviors that are considered to be inappropriate. The behaviors in question often are racist or sexist, or they involve bullying or sexual harassment (Romano, 2020). Of course, since a person's values and perceptions related to these topics can differ considerably, there can be significant disagreement between people on what is a cancel-worthy offense (Romano, 2020).

In Woodley's case, 19-year-old actress and former Disney star Skai Jackson brought Woodley's posts to light. Jackson spent the spring and summer months of 2020 waging a campaign on Twitter, where she encouraged her 560,000 followers to expose racist social media posts made by other teens. Jackson framed her actions as "anti-bullying," as she retweeted the racist posts sent to her by her followers (Bowenbank, 2020). Jackson further encouraged her followers to publicly share the names and future universities of any graduating seniors whose posts they found offensive. Jackson also forwarded a copy of each students' offensive statements to their prospective universities' admissions offices, which is how the University of Florida first learned of Woodley's posts (Rogers, 2020; Spiked, 2020).

Woodley's case is therefore not an isolated incident. The Class of 2020 graduated around the same time as the nation-wide George Floyd protests, and some students felt compelled to post their opinions about it, whether positive or negative. As a result, dozens of incoming college students were called out online for their racist social media posts in early- to mid-2020 (Jaschik, 2020; Levin, 2020). As in Woodley's case, many universities also rescinded the admission of similarly situated incoming freshmen. However, what is distinct here is that Woodley's racist comments were two years old (Rogers, 2020). This points to the fact that there is no statute of limitations, or prescriptive period that sets the maximum amount of time that action can be taken on a matter, for social media posts (Suciu, 2020). Older posts absolutely may be used by universities to make or revoke admissions decisions, or to take disciplinary action against students or other members of the campus community. Moreover, precedent is clear that social media users should have no reasonable expectation of privacy for their personal social media posts (*Nucci* v. *Target Corp*., 2015). The University of Florida therefore acted within its legal rights to rescind Woodley's admission because of posts made when she was just 16 years old.

For Woodley, this situation also resulted in online threats and even death threats made against her (Piper, 2020). Cyberbullying that rises to the level of online threats is not speech that is protected by the First Amendment and also likely violates anti-bullying statutes (O'Connor, Drouin, Davis, & Thompson, 2018). It is possible then that these cancel culture cyberbullies could eventually be held accountable for their participation in the online harassment of Woodley and others. It is also quite possible that they too could have their own admissions rescinded, or be expelled from universities as well.

CURRENT STUDENTS

It does not appear that Purdue University's President was overly concerned about his administration's determination that Maxwell Lawrence's posts might have been protected by the First Amendment. Rather, Purdue President Mitch Daniels' swift action to expel Lawrence was because Daniels believed that Lawrence's racist posts, which referenced the lynching of black men and running over Black Lives Matter protesters with a car, posed a threat to public safety (WTHR, 2020). The Supreme Court has defined this concern as a "true threat," meaning "those

statements where the speaker means to communicate a serious expression of an intent to commit an act of unlawful violence to a particular individual or group of individuals" (*Virginia* v. *Black*, 2003, p. 359). The Court has further clarified that the speaker "need not actually intend to carry out the threat" in order for the speech at issue to lose First Amendment protection (*Virginia* v. *Black,* 2003, p. 360). Therefore, Purdue likely could expel Lawrence for his social media posts, if they indeed constituted a true threat.

Regulation of Off Campus Speech

Interestingly, much of Lawrence's questionable social media activity took place while Purdue students were off campus due to the COVID-19 pandemic. And, while Supreme Court precedent is clear that universities may regulate on campus speech that causes a substantial or material disruption to university operations, there is little guidance from the court at this time regarding off campus speech (*Tinker* v. *Des Moines,* 1969; Towslee, 2020). Though there is consensus that off campus speech involving threats may be regulated, the lack of guidance from the court has led to a split among the circuit courts regarding the legal test that should be used when there is non-threatening off campus speech, or whether it should even be regulated at all (e.g., *B.L. Levy* v. *Mahanoy Area School District,* 2020, holding that the *Tinker* standard did not apply to the non-threatening off campus social media speech of a high school student). The Supreme Court will ultimately have to weigh in on the subject of off campus speech for there to be clear guidance on this issue.

Currently, for circuit courts that do apply the *Tinker* standard to non-threatening off campus speech, most utilize the "nexus" test. The nexus test examines whether there is a significant connection between the speech in question and the campus or campus community (Towslee, 2020). If the off campus speech is significantly related to the university, and it could cause a substantial disruption to university operations, then the university will likely be able to regulate the speech. However, the 2nd, 8th, and 11th Circuit Courts of Appeal all utilize the "reasonably foreseeable" test. In other words, whether off campus student speech may be regulated depends upon whether it is reasonably foreseeable that the speech in question would reach school property and have disruptive consequences there (Towslee, 2020). Notably, the "reasonably foreseeable" test has been criticized for its potential overbreadth, since arguably anything posted online could reasonably reach school property (Towslee, 2020).

"Unprofessional" Social Media Speech

Interestingly, the 8th Circuit did not use the reasonably foreseeable test in a recent case when it held that a public university could expel a student for off campus "unprofessional" social media speech (*Keefe* v. *Adams*, 2016, p. 531). In this case, Craig Keefe was a nursing student at Central Lakes College in Minnesota, in his last semester of study, when he referenced "giving someone a hemopneumothorax" (also known as a collapsed lung) on Facebook (Hudson, 2016). Keefe also called a female student in his program a "stupid bitch" (Hudson, 2016; *Keefe* v.

Adams, 2016, p. 527). Following his expulsion from the nursing program, Keefe argued that his statements were meant to be jokes and that the university could not discipline him for off campus speech unless it was unprotected by the First Amendment (*Keefe* v. *Adams*, 2016). The 8th Circuit rejected this argument and held that a professional program, such as nursing, may adopt a profession's code of ethics as part of its curriculum, and that a determination of non-compliance could be based on a student's speech alone (Hall, 2017; *Keefe* v. *Adams*, 2016). Therefore, the 8th Circuit added new limitations on the free speech rights of students in professional programs, taking the restrictions well beyond the *Tinker* standard because in Keefe's case, no substantial disruption to the university took place (Hall, 2017; *Tinker* v. *Des Moines*, 1969).

STUDENT ATHLETES

Like professional program students, student athletes are highly regulated in their social media activities. As a condition of student athletes' offers to play sports and/or their athletic scholarships, most student athletes must agree to university monitoring of their accounts. Some universities even completely ban their student athletes from social media use while they are enrolled at the university, or at least during their athletic seasons each year (Heintzelman, 2017; Umar, 2015). This is because universities have a particular interest in ensuring that student athletes are careful about what they post, since they are often well-known in the community, or may even be recognized at the state or national level (Hernandez, 2013). In exchange, student athletes often have most, if not all, of their college tuition and expenses paid for by the university, which has led to the somewhat controversial question of whether student athletes should or should not be considered university employees (Heintzelman, 2017).

Monitoring Student Athlete Social Media Accounts

Universities participating in National Collegiate Athletic Association (NCAA) divisions agree to follow the NCAA's guidelines regarding monitoring of student athlete accounts. The NCAA, a private association, handles the monitoring of recruits' accounts. Once a student athlete matriculates, the member university is given a great deal of leeway by the NCAA regarding the monitoring method(s) that will be used (Umar, 2015). And, the methods do vary tremendously. For example, some universities require their coaching staff to monitor the accounts of each of their team members. However, this can lead to highly subjective and varied determinations of misconduct. Other universities utilize computer programs, such as Varsity Monitor, Centrix Social, Fieldhouse Media, etc., which are preprogramed to recognize offensive language and will alert the university and/or the coaching staff if something is found (NCAA, 2020).

One recent study highlighted student perceptions of university social media monitoring practices. O'Connor, Schmidt, and Drouin (2016) found that most university students (78%) oppose the monitoring of the general student population's social media accounts. Students were slightly more tolerant of the monitoring

of student athlete social media accounts, though 68% of surveyed students were still opposed to it. With regard to restrictions on usage, 85% of students thought that student athletes should not have their social media accounts restricted by the university, and 87% of students thought that student athletes should not have to relinquish their social media usage altogether (O'Connor et al., 2016).

Social Media Training, NCAA Guidelines, and the Tinker Standard

Currently, there is no standardized social media training for student athletes through the NCAA, though many universities have devised their own programs to teach student athletes about the pitfalls associated with posting unsuitable content (NCAA, 2020). Proper training should inform students about NCAA social media guidelines, what the university considers to be improper social media practices, as well as the consequences of violating the rules (Bentley, 2012). For most student athletes, the main risk they face in social media cases is suspension or dismissal from their teams. This is because participating in college sports at the university level is considered a privilege, not a right (Hernandez, 2013). However, as a college student, a student athlete would have to violate the *Tinker* "substantial disruption to university operations" test to be suspended or expelled from the university altogether (Bentley, 2012).

In 2018, student athlete Donald De La Haye was kicked off the UCF football team and also dismissed from the university. This occurred because De La Haye had a YouTube channel that was popular enough (several hundred thousand followers) to generate a modest amount of advertising revenue (Henneke & Riches, 2018). However, NCAA guidelines required that student athletes retain their amateur status and not profit off of their name, image, or likeness while they are a student athlete. De La Haye's YouTube channel therefore violated NCAA policy. Following his dismissal from the team and the university, De La Haye filed a lawsuit claiming that his free speech rights were violated. The district court held that De La Haye had a valid legal claim against the university since his YouTube channel did not cause a substantial disruption to university operations (*De La Haye* v. *Hitt*, 2018). UCF and De La Haye eventually settled the case and UCF allowed him to return as a student and finish his degree. However, De La Haye was not allowed back on the football team (Henneke & Riches, 2018). Interestingly, one year after this case settled, the NCAA Board of Governors voted unanimously to start allowing college athletes to be paid for the use of their name, image, and likeness (Almsay, Sterling, & Barajas, 2019). The NCAA specified, though, that student athletes still may not be compensated to play their sport and that they still should not be considered university employees (Almsay et al., 2019).

REGULATING FREE SPEECH IN THE UNIVERSITY

University Social Media Policies

The majority of universities have social media policies to govern the use and content of their stakeholder's posts. However, general knowledge and understanding

of these policies is not necessarily widespread. O'Connor et al. (2016) also found that a surprising 70% of students did not know whether or not their universities even had a social media policy. Moreover, the results of this study also showed that students are quite underinformed about social media and the law. For example, 43% of students thought that they had a right to privacy for their social media posts, and another 42% of students said that they did not know if they had a right to privacy. With regard to the First Amendment, 37% of students said that they believed they had a right to free speech for the things that they post, and 50% of students said they did not know. This is particularly troubling as the sample pool that was used for this study was made up solely of public university students (O'Connor et al., 2016). Universities must do a better job educating students on this topic.

The language of university social media policies varies tremendously from campus to campus. For public universities, the language used in their social media policies must not infringe on the free speech rights of their stakeholders. For private universities, they may restrict their students' speech in the way that they see fit (Ardinger, 2011). However, there is one limitation with regard to private universities' and their restriction of employee social media speech. Sections 7 and 8 of the National Labor Relations Act protect private entity employee speech that relates to terms and conditions of employment, whether the employees are unionized or not (NLRA, 1935). Therefore, the social media policies of private universities have to account for this type of protected speech. Otherwise, their social media policies will not withstand court scrutiny.

Password Protection Laws

Over time, there have been a number of state laws that have been passed to protect university employees and students who are social media users. Password protection laws protect employees and students from requests to disclose their social media user passwords to their universities. Some state laws extend the protection so that universities also may not send friend requests to students or enlist third-party vendors (such as those mentioned above) to monitor student accounts (Heintzelman, 2017). Currently, 16 states have password protection laws in place that apply to educational institutions (NCSL, 2020). University employees and students in these states therefore retain personal control over their social media accounts, but these laws do not offer any additional protection for the things they post.

Title IX in the University

Regarding online harassment, universities have another legal issue to consider: Title IX of the Education Amendments of 1972. Title IX states that

> no person in the United States shall, on the basis of sex, be excluded from participation in, be denied the benefits of, or be subjected to discrimination under any educational program or activity receiving Federal financial assistance. Title IX (Education Amendments Act of 1972, 2018).

Courts have interpreted Title IX to permit a private cause of action for hostile environment harassment (*Cannon* v. *Univ. of Chicago*, 1979). Universities may therefore be liable if a plaintiff shows that the speech or conduct at issue is

> so severe, pervasive, and objectively offensive, and that so undermines and detracts from the victims' educational experience, that the victim-students are effectively denied equal access to an institution's resources and opportunities. (*Davis* v. *Monroe County Bd. Of Ed.*, 1999, p. 651)

Also, a plaintiff must prove that the university had "actual knowledge" of this speech or acted in a manner suggesting "deliberate indifference" (*Kollaritsch* v. *Michigan State,* 2019, p. 623). In some social media cases, universities may be concerned that their failure to act may lead to liability. Of course, to violate Title IX's protections, the facts of a case would likely have to include more than an isolated incident or single social media post to meet the "severe and pervasive" standard.

CONCLUSION

In sum, the legal landscape that surrounds the topic of free speech and social media in academia is quite complex. Though First Amendment protections indeed exist for public university faculty and students, they are narrow, only protecting statements that amount to a public concern. Tenure, academic freedom, and civility policies may also have an impact on faculty social media speech. Although many believe these factors add additional layers of protection for extramural faculty social media speech, the reality is that they likely do not. Rather, it is imperative for faculty, tenured or not, to realize that not all speech is protected from university discipline. Even if a faculty member's social media post does amount to protected speech, the university may still discipline that faculty member if what they post negatively impacts the university's efficient operation.

With regard to students and student athletes, their free speech rights are also limited in scope. Students must learn early on that what they post can have a negative impact on their admissions status and/or might cost them scholarship money, athletic, or otherwise. Likewise, for current students, if what they post substantially interferes with university operations, or even will likely interfere with university operations, they too may face disciplinary action.

Though universities aspire to be places that support the free flow of thought, speech, and expression, this chapter illustrates that the expectation for university stakeholders (who are also social media users) should not be that the things they post will always be protected. Rather, the legal issues surrounding personal social media communications in a university setting are highly complex and fact-dependent, and precedent is still being established. What is clear, though, is that the free speech protection that faculty and students at public universities enjoy, indeed has some limitations. And, while faculty and students may have the right to express their thoughts and opinions via social media, their free speech may not always be free from university discipline.

REFERENCES

AAUP. (n.d.). On freedom of expression and campus speech codes. American Association of University Professors. Retrieved from https://www.aaup.org/report/freedom-expression-and-campus-speech-codes. Accessed on August 10, 2020.

Abrahamson, R. P. (2020, July 31). Colleges are revoking admissions offers due to racist posts. *Today*. Retrieved from https://www.today.com/parents/colleges-are-revoking-admissions-offers-due-racist-posts-t187632. Accessed on August 10, 2020.

Almsay, S., Sterling, W., & Barajas, A. (2019, October 29). NCAA says athletes may profit from name, image, and likeness. *CNN*. Retrieved from https://www.cnn.com/2019/10/29/us/ncaa-athletes-compensation/index.html. Accessed on April 26, 2020.

Appleton, A. (2018, September 5). Randa Jarrar back teaching for the first time since tweet celebrating Barbara Bush's death. *The Fresno Bee*. Retrieved from https://www.fresnobee.com/news/local/education-lab/article217867000.html. Accessed on June 20, 2020.

Ardinger, A. (2011). Private universities and freedom of expression. *The Elon Journal of Undergraduate Research in Communications*, *2*(1), 94–102.

Associated Press. (2020, June 9). Louisiana University acts against 'racist' posts by faculty. *US News & World Report*. Retrieved from https://www.usnews.com/news/best-states/louisiana/articles/2020-06-09/louisiana-university-acts-against-racist-posts-by-faculty. Accessed on June 30, 2020.

B.L. Levy v. *Mahanoy Area School District* (2020 WL 3526130 [2020]).

Bentley, E. D. (2012). He tweeted what? A First Amendment analysis of the use of social media by college athletes and recommended best practices for athletic departments. *Journal of College and University Law*, *38*, 451–479.

Bérubé, M. (2019). Talking out of school: Academic freedom and extramural speech. *MLA Profession*. Retrieved from https://profession.mla.org/talking-out-of-school-academic-freedom-and-extramural-speech/

Bolden, B. (2020, July 11). Nursing professor who used racial slur no longer at ULM. *Monroe News Star*. Retrieved from https://www.thenewsstar.com/story/news/education/2020/07/11/nursing-professor-who-used-racial-slur-no-longer-at-ulm/3193863001/. Accessed on September 9, 2020.

Bowenbank, S. (2020, June 10). Here's why Skai Jackson is using her platform to expose racist bullies on Twitter. *Cosmopolitan*. Retrieved from https://www.cosmopolitan.com/entertainment/celebs/a32823857/skai-jackson-expose-bullies-racists-twitter/

Buchanan v. *Alexander*, 919 F. 3d 847 (5th Cir. 2019).

Campus Speech Database. (2013, September). *David Guth* v. *University of Kansas*. Duke Law. Retrieved from https://campus-speech.law.duke.edu/campus-speech-incidents/david-guth-v-university-of-kansas/

Cannon v. *Univ. of Chicago*, 441 U.S. 669 (1979).

Clement, J. (2020). Social media usage in the United States: Statistics & facts. *Statista*. Retrieved from https://www.statista.com/topics/3196/socialmedia-usage-in-the-united-states/

Comas, M. (2020, June 14). UCF protestors demand professor be fired for racist tweets. *Orlando Sentinel*. Retrieved from https://www.orlandosentinel.com/news/seminole-county/os-ne-ucf-professor-negy-racist-tweets-20200614-pqznqgsafnhqbd36eb2pign4si-story.html. Accessed on August 15, 2020.

Connick v. *Myers*, 461 U.S. 138 (1983).

Cooper, P. (2020). 140+ Social media statistics that matter to marketers in 2020. *Hootsuite*. Retrieved from https://blog-assets.hootsuite.com/wp-content/uploads/2022/03/social-media-statistics-1.png

Davis v. *Monroe County Bd. Of Ed.*, 526 U.S. 629 (1999).

De La Haye v. Hitt, No. 6: 18-cv-135-Orl-22GJK (M.D. Fla. 2018).

Dguth. (n.d.). Retrieved from https://dguth-journalism.ku.edu/

FIRE. (2015). Fighting for free speech at private universities. Retrieved from https://www.thefire.org/fighting-for-free-speech-at-private-universities/

Garcetti v. *Ceballos*, 574 U.S. 410 (2006).

Gillis v. *Miller*, 845 F. 3d 677 (2017).

Hall, M. (2017). *Keefe* v. *Adams*: Overregulating off campus speech under professional codes of conduct. *Minnesota Law Review*. Retrieved from https://minnesotalawreview.org/2017/04/25/keefe-v-adams/

Harris, S. (2020). The Floridian inquisition. *Quilette*. Retrieved from https://quillette.com/2020/08/13/the-floridian-inquisition/

Hawke, C. (1997). Commentary: Tenure's tenacity in higher education. *Education Law Reporter*, *10*(621), 635–636.

Heintzelman, E. (2017). #*NCAA* v. *Student athletes*: An empirical analysis of NCAA social media policies. *Hastings Communication and Entertainment Law Journal*, *39*, 43–65.

Henneke, R., & Riches, J. (2018). Attorneys: UCF's De La Hayes settles for a bright future off the field. *Texas Public Policy Foundation*. Retrieved from https://www.texaspolicy.com/attorneys-ucfs-de-la-haye-settles-for-a-bright-future-off-the-field/

Hernandez, A. (2013). All quiet on the digital front: The NCAA's wide discretion in regulating social media. *Texas Review of Entertainment and Sports Law*, *15*, 53–66.

Herron, A. (2019, November 24). IU professor says university can't fire him. He's right. Here's why. *IndyStar*. Retrieved from https://www.indystar.com/story/news/education/2019/11/24/eric-rasmusen-says-iu-cant-fire-him-hes-right-heres-why/4272679002/. Accessed on December 11, 2019.

Higbee v. *Eastern Michigan University*, 399 F. Supp. 3d 694 (E.D. Mich. 2019).

Hudson, D. (2016). The First Amendment Encyclopedia. Retrieved from https://www.mtsu.edu/first-amendment/article/1565/keefe-v-adams-8th-cir

Hudson, D. (2018). Free speech on public college campuses overview. *Freedom Forum Institute*. Retrieved from https://www.freedomforuminstitute.org/first-amendment-center/topics/freedom-of-speech-2/free-speech-on-public-college-campuses-overview/

Jaschik, S. (2013). Fireable tweets. *Inside Higher Ed*. Retrieved from https://www.insidehighered.com/news/2013/12/19/kansas-regents-adopt-policy-when-social-media-use-can-get-faculty-fired

Jaschik, S. (2020). Admissions offers revoked over racist comments. *Inside Higher Ed*. Retrieved from https://www.insidehighered.com/admissions/article/2020/06/22/colleges-reverse-admissions-offers

Keefe v. *Adams*, 840 F. 3d 523 (2016).

Kollaritsch v. *Michigan Bd. of Trustees*, 944 F. 3d 613 (2019).

Kress, S. (2020, August 26). Kelley professor Eric Rasmusen on unpaid leave for the academic year. *Indiana Daily Student*. Retrieved from https://www.idsnews.com/article/2020/08/iu-kelley-professor-eric-rasmusen-on-unpaid-leave. Accessed on August 28, 2020.

Leary v. *Daeschner*, 228 F. 3d 729 (2000).

LeRoy, M. H. (2015). #AcademicFreedom: Twitter and First Amendment rights for professors. *Notre Dame Law Review Online*, *90*(3), 158–166.

Levin, D. (2020, July 3). Colleges are rescinding admissions offers as racist social media posts emerge. *The New York Times Company*. Retrieved from https://www.boston.com/news/education/2020/07/03/colleges-are-rescinding-admissions-offers-as-racist-social-media-posts-emerge. Accessed on August 10, 2020.

Levy, R. (2014). The tweet hereafter: Social media and the free speech rights of Kansas public university employees. *Kansas Journal of Law & Public Policy*, *24*, 78–135.

LoMonte, F., & Jadon, D. (2018, May 9). Don't expect professors to get fired when they say something you don't like. *Society & Culture*. Retrieved from https://www.salon.com/2018/05/09/do-not-expect-professors-to-get-fired-when-they-say-something-you-do-not-like_partner/. Accessed on September 13, 2020.

NCAA. (2020). Social media literacy for student athletes. Retrieved from http://www.ncaa.org/about/resources/research/social-media-literacy-student-athletes

NCSL. (2020). State social media privacy laws. *National Conference of State Legislatures*. Retrieved from https://www.ncsl.org/research/telecommunications-and-information-technology/state-laws-prohibiting-access-to-social-media-usernames-and-passwords.aspx

NLRA. (1935). National Labor Relations Act. § 7 and 8.

Nucci v. *Target Corp*., 162 So. 3d 146 (Fla. Dist. Ct. 2015).

Nyabwa v. *Facebook* (U.S. Dist. LEXIS 13981 [2018]).

O'Connor, K., Drouin, M., Davis, J., & Thompson, H. (2018). Cyberbullying, revenge porn and the mid-sized university: Victim characteristics, prevalence, and students' knowledge of university policy and reporting procedures. *Higher Education Quarterly*, *72*(4), 344–359.

O'Connor, K., & Schmidt, G. B. (2019). Civility 101: Free speech, social media, and university faculty. *Industrial and Organizational Psychology*, *12*, 395–399.

O'Connor, K., Schmidt, G. B., & Drouin, M. (2016). Suspended because of social media? Students' knowledge and opinions of university social media policies and practices. *Computers in Human Behavior*, *65*, 619–626.

Papandrea, M. (2017). The free speech rights of university students. 101 Minn. L. Rev. 1801.

Parker, Z. (2020, June 10). ULM faculty members under fire for racist remarks on social media. *The Ouchita Citizen*. Retrieved from https://www.hannapub.com/ouachitacitizen/news/local_state_headlines/ulm-faculty-members-under-fire-for-racist-remarks-on-social-media/article_3d8d2814-ab1b-11ea-bbbb-236aa52a81a6.htm. Accessed on September 9, 2020.

Pickering v. *Board of Education*, 391 U.S. 563 (1968).

Pierce, M. (2015, March 10). Is University of Oklahoma frat's racist chant protected by 1st Amendment? *Los Angeles Times*. Retrieved from https://www.latimes.com/nation/la-na-oklahoma-fraternity-explainer-20150310-story.html#page=1. Accessed on June 20, 2020.

Piper, G. (2020, June 9). University of Florida rescinds admission to student for old 'racist' posts; she's getting death threats. *The College Fix*. Retrieved from https://www.thecollegefix.com/university-of-florida-rescinds-admission-to-student-for-old-racist-posts-shes-getting-death-threats/. Accessed on June 12, 2020.

Postal, L., & Martin, A. (2020, June 13). Students' racist social media posts prompt university investigations, raise free speech questions. Orlando Sentinel. Retrieved from https://www.orlandosentinel.com/news/education/os-ne-admissions-universities-racist-posts-20200613-bnqgz5rzlvey5oxzgqgj745k5u-story.html. Accessed on September 13, 2020.

Robson, C. (2014, January 16). Coming to terms with tenure. *University Business*. Retrieved from https://universitybusiness.com/coming-to-terms-with-tenure/. Accessed on August 10, 2020.

Rogers, N. (2020). UF reverses admission decision after racist post resurfaces. Retrieved from https://www.wcjb.com/content/news/UF-reverses-admission-decision-after-racist-post-arises-571141451.html

Romano, A. (2020, January 2). Why we can't stop fighting about cancel culture. *Vox*. Retrieved from https://www.realclearpolitics.com/2020/01/02/why_we_cant_stop_fighting_about_cancel_culture_496906.html. Accessed on August 15, 2020.

Sarabyn, K. (2010). Free speech at private universities. *Journal of Law and Education*, *39*, 145–190.

Smith, J., & Weliever, A. (2020, June 24). Student leaders urge policy reform after expulsion. 'This is not an isolated incident.' *The Exponent*. Retrieved from https://www.purdueexponent.org/campus/article_df9c7b7a-b6ef-11ea-8796-e771d5ecd980.html. Accessed on June 25, 2020.

Spiked. (2020). Now cancel culture is coming for children. Retrieved from https://www.spiked-online.com/2020/06/22/now-cancel-culture-is-coming-for-children/

Suciu, P. (2020). Once again, we're re minded social media is forever. *Forbes*. Retrieved from https://www.forbes.com/sites/petersuciu/2020/07/07/once-again-were-reminded-social-media-is-forever/#2a6daf587470

Sundquist, K. (2017). Why would a college rescind an acceptance? *Collegevine*. Retrieved from https://blog.collegevine.com/why-would-a-college-rescind-an-acceptance/

Thorne, A. (2013). Can civility and academic freedom coexist? *National Association of Scholars*. Retrieved from https://www.nas.org/blogs/article/can_civility_and_academic_freedom_coexist

Tinker v. *Des Moines Independent Community School District*, 393 U.S. 503 (1969).

Title IX of the Education Amendments Act of 1972, 20 U.S.C. §§1681 - 1688 (2018).

Towslee, J. (2020). The "Nexus" test vs. the "reasonably foreseeable" test: How off campus student speech can cause on campus consequences. *Iowa Critical Legal Studies Journal*, *13*, 1–32.

U.S. Constitution, First Amendment. (n. d.). Retrieved from https://constitution.congress.gov/constitution/amendment-1/

Umar, T. (2015). Total eclipse of the tweet: How social media restrictions on student and professional athletes affect free speech. *Jeffrey S. Moorad Sports Law Journal*, *22*, 311–351.

US News. (2020, June 9). Louisiana University acts against 'racist' posts by faculty. *US News and World Report*. Retrieved from https://www.usnews.com/news/best-states/louisiana/articles/2020-06-09/louisiana-university-acts-against-racist-posts-by-faculty. Accessed on September 9, 2020.

Virginia v. *Black*, 548 U.S. 343 (2003).

Wilkins, R. (2020, June 25). Purdue student expelled for 'racist and despicable' social media posts. *Journal and Courier*. Retrieved from https://www.southbendtribune.com/news/local/

purdue-student-expelled-for-racist-and-despicable-posts-on-social-media/article_16d57446-b634-11ea-af57-438e328c5c70.html. Accessed on June 25, 2020.

Woosten, C., & Wong, H. (2018, April 19). After calling Barbara Bush an 'amazing racist,' a professor taunts critics: 'I will never be fired.' *The Washington Post*. Retrieved from https://www.washingtonpost.com/news/grade-point/wp/2018/04/18/after-calling-barbara-bush-an-amazing-racist-a-professor-taunts-critics-i-will-never-be-fired/. Accessed on August 12, 2020.

Workplace Diversity. (2019). The truth about tenure in higher education. Retrieved from https://diversityinhighereducation.com/articles/The-Truth-About-Tenure-in-Higher-Education

WTHR. (2020). Purdue University expels student after 'racist and despicable' post on social media. Retrieved from https://www.wthr.com/article/news/local/purdue-university-expels-student-after-racist-post-on-social-media/531-1a5c8f9e-f38a-4d33-8e5a-64797ff2b218?fbclid=IwAR05EFm-VCvQnoiF9AvM8lQi130Rhl3G_PE_xidC6kz8leb_BkiYh64EdGtI

CHAPTER 3

STORIES ABOUT RISK: MEDIA NARRATIVES OF KNOWN, EMERGING, AND NOVEL HEALTH THREATS

Gabriela Capurro and Josh Greenberg

ABSTRACT

Purpose – *The authors examine framing and narrativization in news coverage of health threats to assess variations in news discourse for known, emerging and novel health risks.*

Methodology/Approach – *Using the analytical categories of known, emerging, and novel risks the authors discuss media analyses of anti-vaccination, antimicrobial resistance (AMR), and Covid-19.*

Findings – *Known risks are framed within a biomedical discourse in which scientific evidence underpins public health guidelines, and following these directives prevent risk exposure while non-compliance is characterized as immoral and risky. News coverage of emerging risks highlights public health guidelines but fails to convey their importance as the risks seem too distant or abstract. Media coverage of novel risks is characterized by the ubiquity of uncertainty, which emerges as a "master frame" under which all incidents and events are subsumed. Stories about novel risks highlight the fluid and changing nature of scientific knowledge, which has the unintended effect of fueling uncertainty as studies and experts contradict each other.*

Media and Law: Between Free Speech and Censorship
Sociology of Crime, Law and Deviance, Volume 26, 43–59

ISSN: 1521-6136/doi:10.1108/S1521-613620210000026004

Originality/Value – *This chapter introduces a new analytical framework for examining how media stories represent public health risks, along with previously unpublished analysis of media coverage about AMR and Covid-19. This chapter provides insight about the nature of risk discourses involving media, public health officials, activists, and citizens.*

Keywords: Risk communication; frames; narratives; media; health risks; Covid-19

INTRODUCTION

The first human case of Covid-19 was identified in the last weeks of 2019 in China and by February 2020 the virus had quickly spread around the globe. Travel bans, border closures, and the shutting down of social and economic life in much of the world followed soon after. At the time of writing, this novel coronavirus pandemic, caused by the SARS-CoV-2 virus, has infected more than 30 million people around the world and caused over 1 million known deaths (Johns Hopkins University, 2020).

Health risks generate a pervasive sense of uncertainty (Lupton, 2012), particularly novel risks such as Covid-19 for which there is no vaccine or cure. News coverage of risk, from novel threats to emerging illnesses and known hazards, play a key role in shaping public understanding of health and disease (Briggs & Hallin, 2016; Gerlach & Hamilton, 2014), and influence risk behavior (Capurro, Greenberg, Dubé, & Driedger, 2018). In this chapter, we examine media frames and narratives of known, emerging and novel health risks, focusing specifically on the cases of vaccine hesitancy (known), antimicrobial resistance (AMR) (emerging), and Covid-19 (novel).

MEDIA AND RISK

News media are a dominant source of health information for the general public and healthcare issues, from outbreaks of disease to biomedical research discoveries, regulatory reforms (privatization, deregulation, etc.) and "health news you can use" that provide guidance on how to safely navigate the risks of everyday life are common staples of our daily news diets (Bomlitz & Brezis, 2008; Briggs & Hallin, 2016). Whether broadcast, print or online, news media are actively involved in the co-production of knowledge through their agenda-setting function and role (McCombs & Shaw, 1972). Media have the power to put some health risk issues into the public spotlight, while relegating others to the shadows or removing them from the public conversation altogether (Beck & Levy, 2013). In the middle of a public health emergency (e.g., SARS in 2003, H1N1 in 2009, Ebola virus in 2014, and the ongoing Covid-19 pandemic) stories of chronic illness and disease, for example, will be completely buried by the torrent of coverage about the more immediate threat.

Media and Science

Previous studies show that media coverage of scientific topics is often influenced by science research, yet it does not merely reproduce research findings but rather situates them in a context that makes the story more resonant with audiences (McInerney, Bird, & Nucci, 2004; Perencevich & Treise, 2010). Lupton (2012) argues that news coverage should be considered constitutive of medical knowledge, given that "the mass media portray aspects of medicine, health care, disease, illness and health risks in certain ways" that impact public understanding of them as "many people first learn about new medical technologies and therapies, or the latest research into the link between lifestyle factors and health status, via the mass media" (p. 15). Briggs and Hallin (2016) go further, arguing that media are directly implicated with various other social actors (scientists, corporations, universities, pharmaceuticals, etc.) in the co-production of biomedical knowledge.

Despite the declining influence of traditional news and the fragmentation of the contemporary mediascape, news organizations continue to publish their stories across media platforms, including websites, mobile applications, and social media. Although some news organizations produce specific content for online and mobile platforms, most repurpose news from their legacy platforms (print, radio, and television) and publish them online, reproducing the same definitions and narratives, which are later shared and distributed among other social media users (Russell, Hendricks, Choi, & Stephens, 2015). Regarding health and risk messages, health institutions, authorities, news organizations, and the public all share health information on websites, blogs, and social media (cf. Kata, 2010, 2012; Moorhead et al., 2013). In this sense, social media can be useful tools for risk communication allowing public health offices, governments, and journalists to disseminate personalized messages immediately thus making outreach more effective (Moorhead et al., 2013).

Telling Stories of Disease: Media Frames and Narratives

News media play an important role in shaping how we come to know and understand public health threats, establishing both scientific and colloquial definitions of a disease or risk, labeling, and stigmatizing the views and actions of individuals or groups implicated in their emergence, assigning blame for undesirable consequences, and framing ideal or preferred solutions and distribution of resources (Capurro et al., 2018; Caulfield et al., 2019; Greenberg, Capurro, Dubé, & Driedger, 2019; Nerlich & James, 2009). From outbreaks of vaccine-preventable diseases (e.g., measles), to emerging threats (e.g., AMR), and the spread of novel zoonoses that take the world by surprise (e.g., Covid-19), news about health risks abound. Health information saturates our media environments, contributing to the formation of what Gerlach and Hamilton (2014) call "pandemic culture." Our overexposure to discourses of disease – in the news, on television, in cinema, literature, and elsewhere – shapes our understanding of science and medicine (Gerlach & Hamilton, 2014), including how various risks are prioritized and ranked for public and policy attention.

News coverage of health threats also provides the language and symbolic resources to make sense of those issues (Briggs & Hallin, 2016). Health reporters have specific newsgathering practices that impact news selection, framing, and sourcing (Tanner, Friedman, & Zheng, 2015) that ultimately affect how health risks are communicated to broader audiences. Through the use of personal narratives, that is, storytelling that describes cause-and-effect relationships between events, news media can influence and shape risk behaviors (Capurro et al., 2018; Shelby & Ernst, 2013) generate public sympathy for patients and advocate and/or mobilize support for particular health policies (Rachul & Caulfield, 2015). News coverage of diseases also correlates with public perception of the severity and prevalence of the illnesses (Young, Norman, & Humphreys, 2008), and repeated exposure to misinformation and "fake news" can influence belief and perceptions of health threats in ways that puts the public at further risk of illness or infection (Jolley & Douglas, 2014).

METHODOLOGY

To understand how different types of health risks are portrayed, we organize them into three ideal-type categories: *known risks, emerging risks,* and *novel risks*. These analytical categories allow us to examine the framing and narrativization of health threats that are well-established and for which there is both public experience and familiarity and compare these to those risks that are newer. As principles of partiality and selectivity, media frames are tools of emphasis, interpretation, and presentation, and are used routinely by newsmakers to organize verbal and visual discourses into formats that are accessible to media users and audiences (Greenberg, 2000). Frames organize central ideas and define scientific issues by emphasizing some aspects of an issue over others to allow individuals to rapidly identify why it is important and how it affects them (Nisbet, Brossard, & Kroepsch, 2003). In this sense, the concept of framing captures the ways in which media set the discursive parameters within which individuals may come to "locate, perceive, identify and label" the events and happenings going on around them (Goffman, 1974, p. 21).

Media coverage goes beyond merely structuring a given field of social intelligibility. It also operates syntagmatically to determine the orders and components of a given story, capturing the movement of representation across time. Narratives help to choreograph into a whole and complete story a series of complexly related events such that they become meaningful to readers (Greenberg, 2000). In short, media frames and narratives have a profound impact in public understanding of biomedical science and public health risk by setting the terms by which we can understand them and then situating them into a story format that resonates with our own lives and experiences (cf. Caulfield et al., 2019; Gerlach & Hamilton, 2014; Nisbet, 2009).

Known Risks

Known health risks are those hazards that are most familiar to us because they've been around the longest. Their causes, the way they spread, the level of risk they

pose to individuals or the whole population, preventive measures, and how to treat them are generally well known, although in some cases they may pose dilemmas about how best to prevent them. In that sense, known risks typically generate a lower feeling of uncertainty then emerging or novel risks, as public health guidelines and treatment protocols are typically well defined and institutionalized. There is a well-established biomedical discourse about these risks and generally accepted norms of behavior for minimizing exposure to risk. Examples of known risks are vaccine hesitancy, HIV, heart disease, diabetes, and other common chronic illnesses.

Emerging Risks

Emerging health risks are lesser known, by comparison. Biomedical research is sparser or less established, and these risks contain elements of both scientific and policy uncertainty. While aspects of these hazards may be well known, for example, their causes and nature of spread, other elements remain unknown or unclear, such as which treatment to use or how to implement appropriate measures to reduce personal or population-wide exposure. While some emerging health risks appear abruptly and are felt immediately, others go unnoticed and grow slowly with deep but hidden repercussions. Additionally, these risks can present new and unforeseen challenges, for example, evolving resistance to therapy, or unanticipated longer-term effects associated with infection. They may also become endemic in new places where they have not previously been present. In this sense, there are health guidelines and treatment protocols, but these continue to change. Examples of emerging health risks in North America include AMR, Lyme Disease, and West Nile Virus.

Novel Risks

Novel risks appear unexpectedly and there is typically no prior experience or protocol for managing the threat they pose. Examples include zoonoses which were not seen before in humans or newer, deadlier strains of viruses with which we may have prior familiarity. These risks create high levels of uncertainty as there is no prior knowledge about the nature of the infection, how it spreads, morbidity, and long-term consequences. This type of risk takes the world by surprise as unprepared medical experts and public health officials try to get a handle on the crisis to minimize the threats they pose to populations. There are generally no specific public health guidelines or treatment protocols for novel risks as officials and healthcare providers navigate uncharted waters with little scientific evidence, although approaches for other health threats may be mobilized to guide a response. Examples of this type of novel risk in the last 20 years include SARS, MERS, Zika virus, and the current SARS-CoV-2 (Covid-19) pandemic.

KNOWN RISK: VACCINE HESITANCY

Vaccines are among the most important medical inventions in modern history and have played a fundamental role in reducing the incidence and mortality of

infectious diseases (World Health Organization (WHO), 2010). Prior to the era of mass immunizations, communicable diseases such as smallpox, chickenpox, measles, polio, and rubella devastated populations, particularly children, not only because of their high mortality rates but because of potential life-altering complications, such as blindness, deafness, and brain damage. Despite their effectiveness and the relatively low risk that vaccines pose, growing number of parents remain "vaccine-hesitant" (Dubé et al., 2013) and do not always comply in vaccinating themselves and their children. Vaccine-hesitant parents worry about the safety of vaccines, whether children should get so many vaccines in a short period of time, and the need for vaccination at all (Greenberg, Dubé, & Driedger, 2017).

The past decade has been marked by numerous outbreaks of communicable diseases in North America. A measles outbreak in Quebec infected more than 700 people in 2011; in 2014, another measles outbreak among under-immunized children in British Columbia reached 433 cases; a measles outbreak, which started at Disneyland in California, spread to half a dozen US states, and into Canada and Mexico in late 2014; an outbreak of 75 cases of chicken pox in an Orthodox Jewish community in New York in 2016; and an outbreak of mumps in Toronto surpassed 125 cases in 2017. Outbreaks of vaccine-preventable diseases have become a common occurrence in media discourse with particular emphasis on the figure of the threats posed by *anti-vaxxers*, a pejorative term to designate vaccine-hesitant parents.

Vaccines Save Lives

Although anti-vaxxer stories are pervasive, media coverage of vaccination and vaccine hesitancy is nevertheless dominated by a pro-vaccine discourse (Greenberg et al., 2019). Our analysis of news reporting on the 2015 Disneyland measles outbreak found that news media consistently framed measles as a serious threat, using emotive vocabulary (e.g., "killer virus," "killer disease," and "extremely contagious") and privileging the perspectives of public health and biomedical experts (Greenberg et al., 2019). Moreover, vaccines, in general, and the measles–mumps–rubella vaccine in particular, were described routinely as highly effective and safe, "miracles of modern medicine that save lives" (Greenberg et al., 2019).

This pro-vaccine frame was enacted in both hard news coverage and in columns, editorials, and op-eds, where news organizations establish the normative terms of a social issue by engaging directly in debate and supporting a particular perspective (Greenberg, 2000). In the hard news coverage, public health officers and medical experts were quoted far more frequently, and even penned some columns and op-eds of their own. News coverage, privileging the voices of public health officials, constantly reminded people to keep their immunizations up to date, while portraying stories of "deadly consequences" associated with the pre-mass vaccination era.

The Pleading Mother and the Vulnerable Child

Vaccines are discussed in news media as a moral obligation that one must accept to protect others as well as themselves (Capurro et al., 2018; Greenberg et al.,

2019). As health reporter Sarah Kliff (2015) wrote, vaccination "is not a personal decision. It's a social obligation." The "social responsibility" frame was grounded on the biomedical concept of "herd immunity" and the need to protect vulnerable populations and reduce the potential for disease outbreaks.

Studies have shown that both pro-vaccine and anti-vaccine advocates resort to the "good mother"/"bad mother" trope, while the father figure is mostly absent (Chivers Yochim & Silva, 2013; Hiltz, Christiansen, & Brady, 2020). This gendered discourse positions mothers as solely responsible for the well-being of their children, either by unequivocally following medical advice (pro-vaccine side), or by being critical of the medical establishment and rejecting vaccinations (anti-vaccine side). The *Ottawa Citizen*, for example, published during the 2015 measles outbreak in North America the story of an "anti-vaxxer" mother who became pro-vaccine after her children became seriously ill with whooping cough. She was described as having been "converted" to science by changing her mind and having her children immunized (Spears, 2015). These "redemption narratives" are common in news coverage of outbreaks of vaccine-preventable diseases, offering effective anecdotal strategies for vaccine promotion (Pucci, 2018).

The vulnerability of children is ever-present in news coverage of outbreaks of communicable diseases, particularly in the stories of those who can't be vaccinated due to illness, which increases considerably their risk if exposed to a vaccine-preventable disease. During the 2015 measles outbreak in North America, the *Toronto Star*, for example, published the testimony the father of a child living with cancer who expressed fear and anger at people who choose not to vaccinate their healthy children: "[O]ne of the greatest achievements of the twentieth century has been the elimination of these diseases of our kids," he explained. In the case of his son, being exposed to measles is "really life or death" (Porter, 2015). Furthermore, various news media reproduced a Facebook post by a mother whose newborn was exposed to measles in a doctor's waiting room received not only massive attention on social media but also the traditional news coverage (Clarke, 2015).

These narratives position parents, but particularly mothers, as responsible for protecting their children from vaccine-preventable diseases but also from anti-vaccine beliefs. Vaccine-hesitant mothers are portrayed as negligent, as "bad mothers," not only for putting their children at risk but also for risking vulnerable children who cannot get immunizations. In these narratives, both the mother who pleads for her vulnerable child and the "converted" mother who has shed her former identity as a "bad mother" are characters in an evolving media narrative that puts blame for outbreaks on vaccine-hesitant parents but also shows the possibility for moral redemption.

The Anti-vaxxer Threat

The "anti-vaxxer" was framed as the counter-source to public health officials. While medical experts were positioned as dutifully on guard to protect the population from harm, vaccine-hesitant parents were consistently described as reckless and selfish, and targets of moral approbrium. The anti-vaxxer has thus become a scapegoat for outbreaks of vaccine-preventable diseases, and a target of criticism,

frustration, and outrage, akin to modern "folk devils" (Capurro et al., 2018). A key aspect of this rhetorical construction of "anti-vaxxers" as folk devils are efforts to subject them to caricature and ridicule. As Globe and Mail columnist Tabatha Southey wrote, "Could we stop anti-vaxxers if measles contained gluten?" (Southey, 2015). Another opined that anti-vaxxers are "parents who claim that they will keep measles at bay by feeding their children organic food and channeling their chakras" (Lye, 2015), and a third described anti-vaxxers as worse than terrorists: "Even ISIS supports getting kids vaccinated" (Fisher, 2015).

However, precisely identifying who "anti-vaxxers" are is difficult because of the lack of clear socio-demographic characteristics common across this highly diffused group. In the coverage of the Disneyland outbreak, for example, many stories attempted to profile vaccine-hesitant parents, who were dismissed as irrational, opposed to science, and thus lacking in definitional credibility, especially in columns and op-eds. For example,

> [Jenny] McCarthy, a former Playboy model turned pop immunologist, has made it her mission to connect autism and childhood vaccinations. There is no medical evidence to back this connection, of course. This is a mission powered by junk science, anecdotal tales, gut feelings and widely debunked studies. (Menon, 2015)

A well-known Canadian pundit, Rex Murphy, also weighed into the anti-vaccination debate by eviscerating actress Jenny McCarthy and her "fellow neurosurgeons on 'The View'" depicting them as selfish and immoral for taking advantage of herd immunity and the selflessness of other families:

> they've confessed to have the intellectual power of a dead tree stump and many even living in one … not vaccinating a child amounts to taking a free ride on the good practices of others. Their good practice is the real protection against your immense carelessness. (Lum, 2015)

In reporting about vaccination and outbreaks of communicable diseases, reporters avoided falling into "false balance, that is, giving equal space to expert scientific knowledge and non-expert opinions and claims" (Dixon & Clarke, 2013). Critics of vaccination were not used as sources and their opinions and experiences were rarely published. Instead the news coverage offered more definitional authority to qualified sources, who spoke to the established science of vaccination, than to voices without technical or medical expertise. Despite the clear pro-vaccination narrative, media coverage also acknowledged that vaccines are "an imperfect net" that can be responsible for "rare and typically minor side-effects" (Greenberg et al., 2019). Notwithstanding such nuances in the vaccination debate (cf. Dubé et al., 2013; Greenberg et al., 2017), media reports unequivocally framed vaccine hesitancy as "selfish" and "irresponsible," and as the undisputed cause of the Disneyland measles outbreak.

EMERGING RISK: AMR

Alongside vaccines, antibiotics are recognized as one of the greatest lifesaving medical interventions ever conceived. Combined with proper sanitation measures, antibiotics were introduced to control the spread of dangerous bacteria in

the twentieth century, and largely curtailed its threat. But now antibiotics, and other antimicrobials, are losing their effectiveness against harmful organisms that are evolving AMR. AMR has been labeled as one of the most important contemporary threats to humanity (WHO, 2017). From gonorrhea, tuberculosis, and *C. difficile* to Carbapenem-resistant *Enterobacteriaceae* (CRE) and *Candida auris*, multidrug-resistant infections are spreading around the globe (CDC, 2019), causing an estimated 1.5 million deaths each year worldwide (O'Neill, 2016).

Communicating the risk of AMR in the news can be challenging because it is not a simple story. Unlike viral infections, antibiotic resistance is not an acute risk but rather a slow-growing one, and it refers to dozens of infections caused by a myriad of pathogens. In trying to make sense of the emergent risk of AMR, multiple narratives have been constructed by an assemblage of actors and institutions (biomedical experts, public health agencies, advocates, policy-makers, media, etc.) that define the hazard, identify its causes, distinguish who is at risk, who is responsible, and delineate preventive measures.

Studies have found that news media often use a catastrophe frame when covering AMR. Apocalyptic predictions are now a common trope in the coverage of AMR, including the anticipation of both natural devastation and personal tragedy (Brown & Crawford, 2009). The feeling of doom is emphasized by focusing on the ability of bacteria to evolve, or mutate, in a short period of time, thus quickly becoming resistant to new antibiotics (Brown & Crawford, 2009). Reporters resort to the reification of resistant bacteria using the colloquial term "superbugs" (Nerlich & James, 2009; Washer & Joffe, 2006) which they depict as having human-like qualities and being "clever" in their ability to outsmart the scientific community (Brown & Crawford, 2009).

The "catastrophe frame" is a narrative tool used in news stories of AMR to alert the public and policy-makers about the urgency of the situation and need for action (Nerlich & James, 2009). It is achieved through military metaphors and anthropomorphizing of bacteria (Bouchoucha, Whatman, & Johnstone, 2019), as well as inaccurate definitions and oversimplified accounts of the causes, populations at risk, and preventive measures to reduce risk (Capurro, 2020; De Silva, Muskavitch, & Roche, 2004). Therefore, the catastrophe frame can increase skepticism of scientific expertise (Boklage & Lehmkuhl, 2019).

Underrepresentation of a Complex Risk

AMR is almost always framed as inviting an apocalyptic scenario in which common medical treatments have been rendered useless against the power of these "nightmare bacteria." This catastrophe frame contributes to increased feelings of dread. For example:

> The CDC is recommending increased testing for the superbug gene. The gene spreads readily among bacteria, and the biggest fear is that the gene will spread to bacteria that are now susceptible only to colistin. That would result in a kind of super-superbug, invincible to every lifesaving antibiotic available. (Sun, 2016)

Despite the urgency of AMR coverage, the volume of coverage is surprisingly limited. We sampled AMR coverage in the North American press between 2016

and 2018 and found a surprisingly sparse number of news stories. AMR was described in the coverage as a multidimensional risk of global proportions and with multiple actors implicated in its spread. The risk of AMR was defined as an inevitable, albeit exacerbated, evolutionary process, thus highlighting its nature as a human-caused risk. AMR was mostly described as being caused by human activity such as antibiotic misuse or overuse in in humans and animals.

The complexity of AMR as a health risk is partly due to the need for collective action. As an emergent risk, the causes and consequences of AMR are well understood, as are some of the measures that must be taken to reduce its spread; conversely, how to actually implement and fund infection prevention and control measures at the local, national, and international level, is a major challenge facing nations around the globe. AMR was also framed as a tragedy of the commons, highlighting how individual actions driven by self-interest affect the common good, that is, from countries not regulating antimicrobial use to individuals taking antimicrobials when unnecessary to reduce the perception of immediate personal risk. As one columnist explained:

> Many Canadians feel – incorrectly, it turns out – that antibacterial soap is best for our family; when we buy it, we feel we're making that decision in a vacuum. The fact that using this soap might mean someone you've never met gets infected with a superbug may seem like a distant concern at best. But of course, if we all act this way, in ignorant self-interest, we all eventually lose. (Callaghan, 2016)

Reproduction of Expert Discourse

Another important aspect in news coverage of AMR are the sources used in each story to establish its central frames (Hallin, Manoff, & Weddle, 1993; Schudson, 2012). Not unlike coverage of vaccine hesitancy, public health officials and scientists are typically the "primary definers" in the coverage, usually in the context of reporting or commenting on a new scientific study or the release of a formal report on AMR (typically by the WHO). Where doctors and researchers are interviewed, they typically provide local context and, occasionally, an assessment of the risk posed by AMR to the population. Whereas news narratives about novel viruses focus on competing perspectives of risk (see more below), in the case of AMR the news coverage found all sources agreed on the causes and threats of AMR as a global health risk. The coverage reinforced the definitional power of biomedical experts, and in doing so it obscured cultural and embodied ways of experiencing AMR.

NOVEL RISK: SARS-COV-2

Since first reported in Wuhan, China, in late December 2019, the outbreak of the novel coronavirus now known as SARS-CoV-2 and the disease it causes, Covid-19, has spread globally causing large-scale mandatory quarantines around the world and the closure of international borders, all exacting a massive toll on the global economy. At the time of writing there have been more than 31 million cases of Covid-19 around the world, and over 1 million known deaths (Johns

Hopkins University, 2020). There is no vaccine against SARS-CoV-2 but over a dozen vaccine candidates are currently in various stages of clinical trials.

We examined news coverage of SARS-CoV-2 in seven national and regional Canadian newspapers from the beginning of March until the end of July 2020. The news coverage has focused on various topics, including the effectiveness of face masks, the possibility of airborne transmission, the need for social distancing and self-isolation, the economic impact of locking down social life, among others. However, one aspect that has remained constant in the coverage is uncertainty, caused by (a) lack of knowledge of the virus; (b) the constant change of public health guidelines which get revised when new evidence emerges; and (c) government and health officials not collecting and/or sharing data with the public. In this sense, uncertainty acts as a "master frame" (Benford, 2013) for making sense of Covid-19 and the various events related to it, that is, use of masks, airborne transmission, physical distancing, vaccine research and development, etc.

Politicization of Science

In the news coverage of Covid-19 many issues became politicized as experts and commentators disagreed with some actions taken by political leaders. The use of face masks to reduce the spread of infection provides an illustrative example. At the beginning of the pandemic, the Government of Canada, on the advice of its Chief Public Health Officer, Dr Theresa Tam, decided to follow guidance from the WHO to not recommend the use of cloth face masks for the general public as there was insufficient evidence of its effectiveness (Mendleson, 2020). As cases of Covid-19 began to increase, so did the pressure on Dr Tam and other public health officials to follow the example of some Asian and European countries. For example, *Globe and Mail* columnist Robyn Urback (2020) wrote:

> Canada's Chief Public Health Officer, Theresa Tam, needs to clarify and explain her opposition to the widespread use of masks, which is now mandatory in the Czech Republic and required for grocery store shopping in Austria. Disease control experts in South Korea, China and Taiwan cite them as essential to disease control.

The use of face masks also became a hot political issue as various unions started demanding that employers supply workers with personal protective equipment, but employers refused to do so because public health officials were not recommending it. For example:

> Employees of the Toronto Transit Commission and regional transit agency Metrolinx, which oversees GO, are not permitted to wear masks on the job. Spokespeople at the agencies argue basic masks are not effective at preventing transmission of the virus and can alarm the public. (Moore, 2020)

When Dr Tam later conceded that public use of face masks could add an additional layer of protection if physical distance and hand hygiene were also practiced, news stories framed her shift in guidance as evidence of incompetence and recklessness, rather than as flexibility and accommodation in the face of changes in science. Some commentators also suggested that after the 2003 SARS outbreak in Toronto, public health should have followed the precautionary principle and recommend masks even if scientific evidence was lacking.

Government and public health officials were also heavily criticized for not sharing epidemiological data with the public, which they alleged showed a lack of transparency and bad leadership. While authorities use epidemiological models as tools to prepare for different scenarios, the media coverage framed these models as essential public information and potential evidence that the severe lockdown had worked, so as a way to reduce anxiety. For example:

> Conservative MP Matt Jeneroux, the party's health critic, said Canadians should have all the information. "You have enough buy-in. Canadians are staying at home, they're practising social distancing measures, but the big question is how long do we have to do this for?" he said. "The more information that Canadians are provided with will drive home how important this really is." (Tumilty, 2020)

Eventually, under pressure from the media and political opponents, health officials released their models, but the criticism only shifted to focusing on their quality, which many claimed was deficient resulting from inadequate data collection. An editorial in the *Globe and Mail* states that "Canada has long-standing problems with data. Across the spectrum of life and business, our governments don't collect, compile or release enough of it." The newspaper joined those publicly criticizing the quality of the epidemiological models and the decisions made based on them: "Decisions are only as good as the information underpinning them. During this pandemic, all necessary data must be gathered, and there is no reason not to make it public" (Globe and Mail, 2020).

Expert Disagreements

Covid-19 is a novel health hazard about which there was no known scientific or biomedical knowledge prior to December 2019. Scientific research is now being undertaken at a frenetic pace, and new studies are being published on various aspects of the disease daily. Given this pressure and demand for speed, errors are not unexpected, particularly in the context of pre-print reports that haven't been fully vetted by scientific peer review. News coverage of the pandemic has thus been dominated by an "unsettled science" narrative balancing the views of experts, who may be comfortable working in a context of emerging and uncertain science, against those of the public who desire clarity and certainty in the face of serious illness and possible death.

Amid uncertainty regarding how to treat patients with Covid-19, a US television network reported on March 18, 2020, that the results of a small trial on hydroxychloroquine – a decades-old antimalarial drug – demonstrated a "100% cure rate against coronavirus" (Fox News, 2020). Days after this misleading claim was made, US President Donald Trump began publicly praising the drug (Grady & Thomas, 2020), leading to a surge in demand, supply shortages, and outrage from the medical the scientific communities.

Expert disagreement was also emphasized in the coverage of asymptomatic cases and their contribution to the spread of infection. For example:

> In the early days of COVID, experts were skeptical of initial reports of asymptomatic cases, which weren't thought to play a significant role in spreading the disease. But four months on, there is mounting evidence that "silent spreaders" are a major driver of the pandemic, with

> top scientists now speculating they could comprise between 25 and 50 per cent of all cases. (Yang, 2016)

Disagreement between scientists, experts, public health officers, and elected officials were a standard feature of Covid-19 media coverage. In the news stories, disagreements were depicted as both a part of the scientific process but also as a source of worry, of not knowing whether the measures taken by public health officers were correct, whether wearing masks increased the risk of contagion for Canadians, and even whether some potential treatments were effective or purely based on pseudoscience and wishful thinking.

Frustration over Shifting Guidelines

Management of Covid-19 in Canada has been characterized since the beginning by shifting public health guidelines as elected officials and public health officers have been adopting and amending policies based on other countries' experience with the virus and educated guesswork. The slow emergence of clear scientific evidence has been informing policy changes, which was explained in some of the news coverage, however many commentators have also criticized these shifts in health policy.

For example, at the beginning of the pandemic healthcare workers were instructed to use N95 respirators when treating suspected and confirmed cases on Covid-19, but the measure was later changed to the use of only surgical masks due to shortages of N95 respirators and no clear evidence of airborne transmission. Public health officials were routinely criticized for revising guidelines in light of new evidence, suggesting a clear clash in understanding about the nature of the scientific process. Instead of framing these changes as a normal part of scientific research and science policy, in a context of trying to control an unknown threat, some media coverage framed it as evidence of complacency and incompetence.

DISCUSSION AND CONCLUSION

In this chapter we have compared the framing and narrativization of *known*, *emerging*, and *novel* public health risks. Our findings show that well-known risks are framed within a biomedical discourse in which scientific evidence underpins public health guidelines, and following these directives prevent risk exposure, for example, vaccinating, while non-compliance is characterized as immoral and risky, in this case anti-vaccination. This seems to be a recurrent formula followed by reporters whenever they cover outbreaks, seeking the same sources, and using the same narrative and frames. Additionally, the risks are assumed to be fully known and there is no room for concerns; the science is settled, vaccines are safe, despite the growing concerns of many (Greenberg et al., 2019).

In the case of emerging risks, which tend to be slow-growing threats and are seldom spectacular, the news coverage highlights public health guidelines but fails to convey their importance as the risk seems too distant or abstract. In this sense, the relationship of causality between not following preventive measures

and becoming exposed to the risk remains weak as reporters struggle to make the risk relevant to their readers. While researchers highlight the importance of raising public awareness of AMR, low levels of media attention makes the risk of AMR seem abstract, its consequences distant, and risk prevention measures irrelevant (Capurro, 2020).

Finally, media coverage of novel risks is characterized by the ever-presence of uncertainty. Uncertainty has emerged as a "master frame" under which all related incidents, issues, and events are subsumed (to wear a mask or not; to close schools or not; to open bars and restaurants or not; to allow indoor gatherings or not; etc.). In the case of Covid-19, even expert guidelines and scientific evidence are contingent and complying with public health directives does not reduce the uncertainty either. Furthermore, any public health policies and decisions are quickly politicized. Reporters still sought medical and public health experts and researchers as sources but without conclusive scientific evidence, experts were frequently positioned in opposing camps.

While in the coverage of vaccine hesitancy scientific knowledge provides certainty and reassurance, and the vaccine hesitant are a source of anxiety, in the case of Covid-19, scientific knowledge does not provide the same reassurance, on the contrary, it fuels the feeling of uncertainty. Suddenly, reporters and citizens become part of discussions normally reserved for expert review (i.e., which study is more valid) and in the process, science becomes contingent, politically inflected, and anxiety-inducing.

REFERENCES

Beck, U., & Levy, D. (2013). Cosmopolitanized nations: Re-imagining collectivity in world risk society. *Theory, Culture & Society*, *30*(2), 3–31. https://doi.org/10.1177/0263276412457223

Benford, R. D. (2013). Master frame. In *The Wiley-Blackwell encyclopedia of social and political movements*. American Cancer Society. https://doi.org/10.1002/9780470674871.wbespm126

Benford, R.D. (2013). Master Frame. In The Wiley-Blackwell Encyclopedia of Social and Political Movements (eds D.A. Snow, D. Della Porta, B. Klandermans and D. McAdam). Wiley-Blackwell: Malden, MA and Oxford. https://doi.org/10.1002/9780470674871.wbespm126

Boklage, E., & Lehmkuhl, M. (2019). Coverage of antimicrobial resistance in the German Press: 1993–2013. *Health Communication*, *34*(9), 958–963. https://doi.org/10.1080/10410236.2018.1446250

Bomlitz, L. J., & Brezis, M. (2008). Misrepresentation of health risks by mass media. *Journal of Public Health*, *30*(2), 202–204. https://doi.org/10.1093/pubmed/fdn009

Bouchoucha, S. L., Whatman, E., & Johnstone, M.-J. (2019). Media representation of the antimicrobial resistance (AMR) crisis: An Australian perspective. *Infection, Disease & Health*, *24*(1), 23–31. https://doi.org/10.1016/j.idh.2018.09.084

Briggs, C. L., & Hallin, D. C. (2016). *Making health public: How news coverage is remaking media, medicine, and contemporary life*. London: Routledge.

Brown, B., & Crawford, P. (2009). 'Post antibiotic apocalypse': Discourses of mutation in narratives of MRSA. *Sociology of Health & Illness*, *31*(4), 508–524. https://doi.org/10.1111/j.1467-9566.2008.01147.x

Callaghan, M. (2016). Time to wash our hands of antibacterial soap. *Toronto Star*, September 7, A13.

Capurro, G. (2020). "Superbugs" in the risk society: Assessing the reflexive function of North American Newspaper coverage of antimicrobial resistance. *SAGE Open*, *10*(1), 1–13. https://doi.org/10.1177/2158244020901800

Capurro, G., Greenberg, J., Dubé, E., & Driedger, M. (2018). Measles, moral regulation and the social construction of risk: media narratives of "Anti-Vaxxers" and the 2015 Disneyland outbreak. *Canadian Journal of Sociology*, *43*(1), 25–48. https://doi.org/10.29173/cjs29301

Caulfield, T., Marcon, A. R., Murdoch, B., Brown, J. M., Perrault, S. T., Jarry, J., … Hyde-Lay, R. (2019). Health misinformation and the power of narrative messaging in the public sphere. *Canadian Journal of Bioethics*, *2*(2), 52–60. https://doi.org/10.7202/1060911ar

CDC. (2019, May 7). The biggest antibiotic-resistant threats in the U.S. Centers for Disease Control and Prevention. Retrieved from https://www.cdc.gov/drugresistance/biggest_threats.html

Chivers Yochim, E., & Silva, V. T. (2013). Everyday expertise, autism, and "good" mothering in the media discourse of Jenny McCarthy. *Communication and Critical/Cultural Studies*, *10*(4), 406–426. https://doi.org/10.1080/14791420.2013.841320

Clarke, K. (2015). Baby's measles scare goes viral on Facebook; Mother's angry post slamming 'anti-vaxxers' strikes a nerve. *Toronto Star*, February 12, GT1.

De Silva, M., Muskavitch, M. A. T., & Roche, J. P. (2004). Print media coverage of antibiotic resistance. *Science Communication*, *26*(1), 31–43. https://doi.org/10.1177/1075547004267026

Dixon, G. N., & Clarke, C. E. (2013). Heightening uncertainty around certain science: Media coverage, false balance, and the autism–vaccine controversy. *Science Communication*, *35*(3), 358–382. https://doi.org/10.1177/1075547012458290

Dubé, E., Laberge, C., Guay, M., Bramadat, P., Roy, R., & Bettinger, J. A. (2013). Vaccine hesitancy: an overview. *Human Vaccines & Immunotherapeutics*, *9*(8), 1763–1773. https://doi.org/10.4161/hv.24657

Fisher, M. (2015, February 2). Even ISIS supports getting kids vaccinated. *Vox*. Retrieved from http://www.vox.com/2015/2/2/7966421/vaccination-isis

Fox News. (2020, March 18). Interview. *Tucker Carlson Tonight*. https://www.mediamatters.org/media/3860681

Gerlach, N., & Hamilton, S. N. (2014). Trafficking in the zombie: The CDC zombie apocalypse campaign, diseaseability, and pandemic culture. *Refractory: A Journal of Entertainment Media*, *23*.

Globe and Mail. (2020, March 31). Globe editorial: Memo to the rest of Canada: How Dr. Bonnie Henry and B.C. are getting it right by laying all their cards on the table. *Globe and Mail*. Retrieved from https://www.theglobeandmail.com/opinion/editorials/article-memo-to-the-rest-of-canada-how-dr-bonnie-henry-and-bc-are-getting/

Goffman, E. (1974). *Frame analysis: An essay on the organization of experience* (pp. ix, 586). Cambridge, MA: Harvard University Press.

Grady, D., & Thomas, K. (2020, March 19). With Minimal Evidence, Trump Asks F.D.A. to Study Malaria Drugs for Coronavirus. The New York Times. https://www.nytimes.com/2020/03/19/health/coronavirus-drugs-chloroquine.html

Greenberg, J. (2000). Opinion discourse and Canadian newspapers: The case of the Chinese "Boat People." *Canadian Journal of Communication*, *25*(4), 517. https://doi.org/10.22230/cjc.2000v25n4a1178

Greenberg, J., Capurro, G., Dubé, E., & Driedger, S. M. (2019). Measles, Mickey and the media: News coverage of the 2015 Disneyland outbreak. *Canadian Journal of Communication*, *44*(2), 175–189. https://doi.org/10.22230/cjc.2019v44n2a3346

Greenberg, J., Dubé, E., & Driedger, M. (2017). Vaccine hesitancy: In search of the risk communication Comfort Zone. *PLoS Currents*, *9*. Retrieved from www.ecurrents.outbreaks.0561a01111 7a1d1f9596e24949e8690b; https://doi.org/10.1371/currents.outbreaks.0561a011117a1d1f9596e24949e8690b

Hallin, D. C., Manoff, R. K., & Weddle, J. K. (1993). Sourcing patterns of national security reporters. *Journalism Quarterly*, *70*(4), 753–766. https://doi.org/10.1177/107769909307000402

Hiltz, E., Christiansen, E., & Brady, M. (2020). *Good Karen, bad Karen: Visual culture and the anti-Vaxx Mom on Reddit*. Submitted for publication. Retrieved from https://www.academia.edu/42880950/DRAFT_FORTHCOMING_Good_Karen_Bad_Karen_Visual_Culture_and_the_Anti_Vaxx_Mom_on_Reddit_

Johns Hopkins University. (2020, September 7). *COVID-19 map*. Johns Hopkins Coronavirus Resource Center. Retrieved from https://coronavirus.jhu.edu/map.html

Jolley, D., & Douglas, K. M. (2014). The effects of anti-vaccine conspiracy theories on vaccination intentions. *PLoS ONE*, *9*(2), e89177. https://doi.org/10.1371/journal.pone.0089177

Kata, A. (2010). A postmodern Pandora's box: Anti-vaccination misinformation on the internet. *Vaccine*, *28*(7), 1709–1716. https://doi.org/10.1016/j.vaccine.2009.12.022

Lye, H. (2015, March 13). Dear Anti-Vaxxers: I Believe in Western Medicine Because it Saved My Life. Huffington Post Canada. http://www.huffingtonpost.ca/harriet-alida-lye/anti-vaccine_b_6863604.html

Kata, A. (2012). Anti-vaccine activists, Web 2.0, and the postmodern paradigm: An overview of tactics and tropes used online by the anti-vaccination movement. *Vaccine*, *30*(25), 3778–3789. https://doi.org/10.1016/j.vaccine.2011.11.112

Kliff, S. (2015, January 31). Christie is wrong. Vaccination is not a personal decision. It's a social obligation. *Vox*. Retrieved from https://www.vox.com/2015/1/31/7955947/vaccination-herd-immunity

Lum, Z.-A. (2015, February 6). Rex Murphy: Anti-Vaxxers have 'the intellectual power of a dead tree stump'. *Huffington Post Camada*. Retrieved from http://www.huffingtonpost.ca/2015/02/06/rex-murphy-anti-vaccine_n_6631184.html

Lupton, D. (2012). *Medicine as culture: Illness, disease, and the body* (3rd ed.). Thousand Oaks, CA: SAGE.

McCombs, M. E., & Shaw, D. L. (1972). The agenda-setting function of mass media. *The Public Opinion Quarterly*, *36*(2), 176–187.

McInerney, C., Bird, N., & Nucci, M. (2004). The flow of scientific knowledge from lab to the lay public: The case of genetically modified food. *Science Communication*, *26*(1), 44–74. Retrieved from https://doi.org/10.1177/1075547004267024

Mendleson, R. (2020, April 4). Trying to cut through mask confusion; Should we all wear facial protection to slow virus? Experts are shifting on advice. Toronto Star, A14.

Menon, V. (2015, February 5). Measles and an outbreak of stupidity. *Toronto Star*, E5.

Moore, O. (2020, March 5). Toronto's transit unions demand better health precautions as coronavirus spreads. *Globe and Mail*. Retrieved from https://www.theglobeandmail.com/canada/toronto/article-torontos-transit-unions-demand-better-health-precautions-as/

Moorhead, S. A., Hazlett, D. E., Harrison, L., Carroll, J. K., Irwin, A., & Hoving, C. (2013). A new dimension of health care: Systematic review of the uses, benefits, and limitations of social media for health communication. *Journal of Medical Internet Research*, *15*(4), e85. https://doi.org/10.2196/jmir.1933

Nerlich, B., & James, R. (2009). "The post-antibiotic apocalypse" and the "war on superbugs": Catastrophe discourse in microbiology, its rhetorical form and political function. *Public Understanding of Science*, *18*(5), 574–590. https://doi.org/10.1177/0963662507087974

Nisbet, M. C. (2009). Communicating climate change: Why frames matter for public engagement. *Environment: Science and Policy for Sustainable Development*, *51*(2), 12–23. https://doi.org/10.3200/ENVT.51.2.12-23

Nisbet, M. C., Brossard, D., & Kroepsch, A. (2003). Framing science: The stem cell controversy in an Age of Press/Politics. *Harvard International Journal of Press/Politics*, *8*(2), 36–70. https://doi.org/10.1177/1081180X02251047

O'Neill, J. (2016). *Tackling drug-resistant infections globally: Final report and recommendations* (p. 84). Review on Antimicrobial Resistance. Retrieved from https://amr-review.org/sites/default/files/160525_Final%20paper_with%20cover.pdf

Perencevich, E. N., & Treise, D. M. (2010). Methicillin-resistant *Staphylococcus aureus* and the media. *Infection Control and Hospital Epidemiology*, *31*(S1), S48–S50. https://doi.org/10.1086/656000

Porter, C. (2015). Measles vaccinations for greater good. *Toronto Star*, February 10, A10.

Pucci, A. (2018). *Forgive me child for I have sinned: Anti-Vaxxer Folk devil mothers and narratives of redemption*. Carleton University. Retrieved from https://curve.carleton.ca/2011da02-c6fa-4c84-ba77-d9b8b4623a06

Rachul, C., & Caulfield, T. (2015). The media and access issues: Content analysis of Canadian newspaper coverage of health policy decisions. *Orphanet Journal of Rare Diseases*, *10*(1). https://doi.org/10.1186/s13023-015-0320-z

Russell, F. M., Hendricks, M. A., Choi, H., & Stephens, E. C. (2015). Who sets the news agenda on Twitter? Journalists' posts during the 2013 US government shutdown. *Digital Journalism*, *3*(6), 925–943. https://doi.org/10.1080/21670811.2014.995918

Schudson, M. (2012). *The sociology of news* (2nd ed.). New York, NY: W.W. Norton & Company.

Shelby, A., & Ernst, K. (2013). Story and science: How providers and parents can utilize storytelling to combat anti-vaccine misinformation. *Human Vaccines & Immunotherapeutics*, *9*(8), 1795–1801.

Southey, T. (2015). Could we stop anti-vaxxers if measles contained gluten? *The Globe and Mail*, February 7, F3.

Spears, T. (2015). Converted antivaccine mom says stop yelling; Anger at those who won't vaccinate their children won't work, Tara Hills says. *Ottawa Citizen*, April 18, A6.

Sun, L. (2016). Newest superbug found in a Connecticut toddler. *Washington Post*, September 10, A2.

Tanner, A. H., Friedman, D. B., & Zheng, Y. (2015). Influences on the construction of health news: The reporting practices of local television news health journalists. *Journal of Broadcasting & Electronic Media*, *59*(2), 359–376. https://doi.org/10.1080/08838151.2015.1029123

Tumilty, R. (2020). Into the unknown; health-care workers brace for the pandemic onslaught, but governments aren't revealing how bad it could get. *National Post*, April 2, A1.

Urback, R. (2020). No more incremental measures. We need to enact tomorrow's policies today. *Globe and Mail*, March 31. Retrieved from https://www.theglobeandmail.com/opinion/article-no-more-incremental-measures-we-need-to-enact-tomorrows-policies/

Washer, P., & Joffe, H. (2006). The "hospital superbug": Social representations of MRSA. *Social Science & Medicine*, *63*(8), 2141–2152. https://doi.org/10.1016/j.socscimed.2006.05.018

World Health Organization (WHO). (2010). *WHO|The Smallpox Eradication Programme – SEP (1966–1980)*. WHO. Retrieved from http://www.who.int/features/2010/smallpox/en/

World Health Organization (WHO). (2017). *Global priority list of antibiotic-resistant bacteria to guide research, discovery, and development of new antibiotics*. WHO. Retrieved from http://www.who.int/mediacentre/news/releases/2017/bacteria-antibiotics-needed/en/

Yang, J. (2016). *The rise of the super gonorrhea. Toronto Star*, October 12.

Young, M. E., Norman, G. R., & Humphreys, K. R. (2008). Medicine in the Popular Press: The influence of the media on perceptions of disease. *PLoS ONE*, *3*(10), e3552. https://doi.org/10.1371/journal.pone.0003552

PART II

THE INTERNET AS PUBLIC SPHERE

CHAPTER 4

CENSORING SEX: PAYMENT PLATFORMS' REGULATION OF SEXUAL EXPRESSION

Natasha Tusikov

ABSTRACT

Purpose – *This chapter examines the role of payment platforms in the United States in sex censorship in which platforms have a pattern of denying financial services to people and businesses involved in publishing legal sexual content. It answers the following questions: what explains payment platforms' regulation of lawful sexual content and what are the consequences?*

Methodology/Approach – *Drawing from the platform governance literature, this chapter closely examines the corporate policies for PayPal and the credit card companies that prohibit certain types of sexual content and services.*

Findings – *This chapter argues that payment platforms' censorship of sexual expression is shaped by the distinctive nature of and market concentration within the online payment industry. Payment actors' systematic campaign of sexual censorship disproportionately affects small businesses and individual operators in the sex and adult entertainment industries and amounts to "digital redlining," a form of financial discrimination.*

Originality/Value – *Payment providers' role in regulating sex online has received considerably less scholarly attention than research on social media platforms. This gap in scholarship is notable as big payment actors have*

Media and Law: Between Free Speech and Censorship
Sociology of Crime, Law and Deviance, Volume 26, 63–79

ISSN: 1521-6136/doi:10.1108/S1521-613620210000026005

systematically denied services for about a decade relating to sexually oriented goods and services (see Blue, 2015a).

Keywords: Payment platforms; sexual expression; censorship; risk; surveillance; financial discrimination

INTRODUCTION

The internet's decade-long "war on sex" has raged across search engines, app stores, social media platforms, and marketplaces with internet platforms banning apps, stores, artistic and educational content, advertising, and products and services related, however tangentially, to sex (see Blue, 2015a). Platforms here refer to two-sided, programable architecture like PayPal that facilitate interactions between businesses and users (van Dijck, Poell, & de Waal, 2018, p. 9). In their anti-sex campaigns, platforms forbid content, product, or services that they variously define as "adult content" (Tumblr, 2020) or "sexually oriented" material (PayPal, 2020a), which often conflates illegal activities, such as communication for prostitution that varies in legality among jurisdictions, and lawful sexual speech. This type of speech is commonly termed adult or "Not Suitable for Work" (NSFW) content that includes videos, photos, artwork, fan fiction, poetry, and music undertaken by a diverse group of people, such as artists, performers, educators, and people working in the sex and adult entertainment industries. This chapter uses the terms adult and sexual content interchangeably.

Online anti-sex campaigns are evident in search engines that downrank search results for adult entertainment websites, app stores that ban apps where users exchange sexual or erotic content and content delivery networks that enable the fast delivery of content, and restrictions on a social network friendly to sex workers (see Blue, 2015b). Platforms' censorship of sexual speech is typically most visible on large, United States-based social media platforms, particularly Facebook, Instagram, Tumblr, Twitter, and YouTube where content and performers may suddenly vanish. Social media platforms' perception of inappropriate sexual content can extremely broad. Tumblr's definition of adult content, for example, forbids "real-life human genitals or female-presenting nipples" and "any content (including photos, videos, GIFs & illustrations) that depicts sex acts" (Tumblr, 2020), although certain types of artistic, educational, or newsworthy content are generally permitted at the platform's discretion. Platforms' regulation of "obscenity" and inappropriate sexual content particularly focuses on women's bodies, removing content relating to child birth, post-mastectomy scarring, breastfeeding, menstruation, pubic hair, female nipples, as well as banning many types of LGBTQI+ content (see Chemaly, 2015; Faust, 2017; Roberts, 2019). Critics of social media platforms' sex censorship characterize it as a "kind of cultural genocide on the open internet, one comprised of women, LGBTQI people, artists, educators, bloggers, filmmakers, sex workers, abuse survivors, [and] untold communities" (Blue, 2018).

In contrast to social media platforms, payment providers' role in regulating sex online has received considerably less scholarly attention (an exception is Swartz, 2020). This gap in scholarship is notable as big payment actors have for about a decade systematically denied services relating to sexually oriented goods and services (see Blue, 2015a). As journalist, author and activist Violet Blue has documented, banks (such as JP Morgan Chase), credit card associations (namely Visa, Mastercard, and American Express), and payment platforms (like PayPal and Square) have systematically targeted a wide range of people working in the sex trade and retail adult entertainment industry, including escorts, erotic masseuses, adult toy store employees, erotic dancers, cam girls, and performers and models (Blue, 2015a). Payment actors' denial of services for adult content and sexual services makes it difficult for people to get paid via credit or debit cards, bank accounts, or even via crowdfunding sites like Patreon. Proponents of platforms defunding the sex and adult entertainment industries, such as the US National Center on Sexual Exploitation, argue that these industries are inherently exploitative and enable criminal activities like human trafficking, revenge pornography, and child pornography (see McNamara, 2020), a discourse that problematically portrays people working in those industries as victims and lacking agency (see Musto, 2016).

Many of the activities that payment providers target, however, are legal, although the legality of prostitution and escort work varies widely. Even where prostitution is decriminalized, like some Australian states, payment actors continue to deny individuals merchant bank accounts, and, in some cases, personal accounts (see Simpkins, 2019). Australian politician David Limbrick, condemning Australian banks, credit card companies, and platforms for financially discriminating against sex workers and adult retailers, including where prostitution is legal, said their policies "effectively amount to corporate slut shaming" (Simpkins, 2019).

Alongside targeting sex workers and adult industry performers, payment actors also freeze funds or deny their services with regard to sexual expression that does not involve the sale of sexual services. Those in the adult entertainment industry argue that NSFW content is "tech's biggest shadow industry" given its ongoing struggles to secure reliable payment services (Heartpie, 2017). PayPal is particularly active, with Blue characterizing PayPal as "the king of denying service, seizing accounts and freezing funds for anyone discovered to be associated with sexual content online" (Blue, 2015a). Over the past several years, PayPal in the United States has frozen accounts and seized funds relating to the sale of a corset, a memoir written by former escort Vicki Gallas, a sex bloggers' calendar, and submission fees for the Seattle Erotic Art Festival (Blue, 2015a).

Being expelled from social media platforms has the potential to subdue or outright silence targeted individuals, although they may move to other platforms. Payment platforms, however, have the capacity to block payments and suspend accounts temporarily or permanently thereby strangling the ability of individuals or entities to raise donations or sell goods (see Tusikov, 2017, 2019). Losing trusted, commercially popular payment services can destroy businesses,

particularly small outfits that do not have the resources or technical capabilities to shift to alternative forms of payment like cryptocurrency (see Andrews, 2019). Payment providers' denial of services with regard to sexually oriented content takes on added significance during the Covid-19 pandemic with the increased vulnerability of already precarious sex workers facing economic downturns and public health restrictions on the types of in-person work they can perform (see DeLacey, 2020).

This chapter examines the role of payment platforms in the United States in sex censorship where platforms have a pattern of denying financial services to people and businesses involved in publishing legal sexual content. It answers the following questions: what explains payment platforms' regulation of lawful sexual content and what are the consequences? Regulation in this context refers to the process of setting, monitoring, and enforcing rules or standards, whether by state or non-state actors, that involve formal legal or informal mechanisms. The United States is a key site to examine given the global commercial dominance of its internet platforms, including payment providers, and its outsized influence on setting rules and standards relating to how the internet operates and is governed (Tusikov, 2017).

Drawing from the platform governance literature, the chapter closely examines the corporate policies for PayPal and the credit card companies that prohibit certain types of sexual content and services. The chapter argues that payment platforms' censorship of sexual expression is shaped by the distinctive nature of and market concentration within the online payment industry. Despite the growing number of payment methods, there are a handful of dominant payment actors, especially Visa, Mastercard, and PayPal. Further, many payment systems like PayPal, Square, or Apple Pay rely upon the credit card networks, meaning that if the credit card companies forbid a particular activity, that rule also governs the operation of the associated payment provider.

In other words, certain payment actors, particularly the major US credit card companies, exert structural power over the online payment system. They determine the rules and processes for what constitutes acceptable and prohibited transactions and govern who can be granted merchant accounts to accept payments via credit cards. Credit card companies rate most transactions for sexual content and services as high risk, making them pricey to undertake and leaving small merchants with few commercial options. Alongside these market dynamics, 2018 legislation in the United States – the *Fight Online Sex Trafficking Act*, known colloquially as FOSTA/SESTA – is entrenching payment platforms' regulation of sexual speech. The chapter finds that payment actors' systematic campaign of sexual censorship disproportionately affects small businesses and individual operators in the sex and adult entertainment industries and amounts to "digital redlining," a form of financial discrimination.

The chapter proceeds in six parts. First, it explains how payment platforms employ surveillance to govern through data and then explains payment platforms' operation as regulators, including how credit card companies wield power in the online payment ecosystem. Third, the chapter highlights the effect of FOSTA/SESTA on online sexual speech and then argues that the online

regulation of sexual speech constitutes digital redlining before providing a brief conclusion.

PLATFORM SURVEILLANCE

Platforms that prioritize the mass accumulation and interpretation of users' personal data as part of their business models operate according to a logic that scholars alternately term platform capitalism (Srnicek, 2017) or surveillance capitalism (Zuboff, 2019). It has become commonplace to note that social media platforms or marketplaces like Amazon employ surveillance-intensive practices with the aim of influencing and predicting consumer behavior through mass, often automated data collection and data analytics to generate revenue and exert market control within industry sectors (see Zuboff, 2019). Platforms are fueled by data, automated and organized by algorithms, and governed through user agreements (van Dijck et al., 2018, p. 9). Payment platforms follow a similar logic (see Martin, 2019; Westermeier, 2020), as is evident by the data-intensive business practices of PayPal and other smaller platforms like Stripe and Square.

Monitoring Bank Customers

While platforms' surveillance-intensive business models are relatively novel, having shifted to data accumulation and monetization in the early 2000s (see Zuboff, 2019), the financial industry in the United States has over a century of experience monitoring its customers. In fact, the roots of contemporary surveillance capitalism within the financial system can be traced back to the beginnings of commercial and consumer credit bureaus, with the latter emerging in the 1870s in the United States following the establishment of commercial credit ratings systems (Lauer, 2018). Josh Lauer's fascinating history of these credit bureaus documents their creation of huge archives of personal information, including the collection of data on customers' marital status, job performance, domestic arrangements, physical appearance, and health, legal and criminal histories. As Lauer (2018, p. 11) argues, corporate credit surveillance facilitated the growth of state surveillance, even assisting the Federal Bureau of Investigation and Internal Revenue Service in the early 1900s given the credit bureaus' vast repositories of personal data.

Surveillance thus is an intrinsic element of the financial system. The early credit bureaus evolved into modern credit giants like Experian that minutely track the spending behavior of over a billion people to determine consumers' financial status and their financial identity in terms of being creditworthy or a credit risk (Lauer, 2018). The credit card companies, which operate as an "invisible surveillance network" (Lauer, 2018, p. 1) pervasively monitor customers and their transactions, using proprietary algorithms to verify each purchase by checking the customer's spending profile, location, and credit limit and the merchant's location, and either approving the purchase or flagging it as suspicious.

Money is Data

Financial institutions' pervasive surveillance of their individual and corporate customers enables the continued growth of financial transactions and helps manage any risk to their businesses. Historically, financial institutions regarded this data collection as a means to delivering financial services; however, with the increasing incorporation of technology into the financial industry, data collection has taken on an important new dimension. Payment actors traditionally "thrived on money, not data" (Westermeier, 2020, p. 2) but have now adjusted their business models to exploit the economic value of the transactional data they collect. Simply put, "data is the new money, but also, money is data" (Pardes, 2019 cited in Westermeier, 2020, p. 2). More important than the purchase amount is transactional data, including information on the sellers, buyers, shopping histories, payment methods, and customer preferences (Maurer, 2012, p. 477, cited in Westermeier, 2020, p. 2). Data thus becomes an asset. Platforms' accumulation and aggregation of personal data are undertaken to assetize data with the understanding that financial value will be accrued (Birch, Chiappetta, & Artyushina, 2020, p. 11).

Payment platforms' conceptualization of money as data is evident in the granularity of their surveillance of customers' behavior and transactions. PayPal, for example, collects data including, users' mailing and email addresses, telephone numbers, account numbers, shopping history, and geolocation data like IP addresses, as well as users' personal characteristics like national origin, disability, citizenship, and military status (PayPal, 2020b). PayPal also makes inferences based on its collection of personal data with regard to its users' traits, habits, and interests to personalize users' preferences and to conduct its fraud and risk assessments (see PayPal, 2020b). Platforms' data collection practices and generation of inferential data can raise privacy concerns as transactional data may reveal highly sensitive information about people's political views, religion, health, and sexual habits. By combining different forms of data that until now have been separately collected – and, additionally, not been targeted by payment platforms – companies can conduct more comprehensive, personalized assessments of people's social or economic situations (Westermeier, 2020, p. 2).

Payment platforms also collect information not traditionally considered financial data. In 2013, PayPal acquired the Venmo mobile payment application with a social networking function. Designed to let people split payments, Venmo encourages its users to share details with their social networks, posting messages and emojis about their activities and transactions (Venmo, 2020). Unless users set their Venmo accounts to private, messages about transactions are publicly visible, revealing sometimes sensitive transaction details, such as drug sales where the communication between seller and customers are captioned with emojis describing the purchase and use of marijuana (see Do Thi Duc, 2017). By combining financial data with social media data about their users' contacts, interests, and habits, PayPal can have a granular level of detail about their users that goes far beyond what is traditionally understood as financial credit rating

(see Swartz, 2020). Venmo thus has similarities with the Chinese social credit program, designed by the Chinese government and carried out by platforms like Tencent and Alibaba, with the goal of monitoring, controlling, and predicting the trustworthiness – and creditworthiness – of people and companies within the country (Liang, Das, Kostyuk, & Hussain, 2018, p. 415).

PAYMENT PLATFORMS AS REGULATORS

Payment platforms are not new to policing their users for violations of their policies, and can be effective regulators because they have contractual relationships with their users and detailed data on their customers. They have blocked transactions and terminated accounts with regard to a range of illegal or unwanted activities, including copyright infringement (Bridy, 2015), child pornography (Laidlaw, 2015), counterfeit goods (Tusikov, 2017), hate speech (Tusikov, 2019), and sexually explicit content and services (Swartz, 2020).

Created in 1998, PayPal is an online payment company that performs money transfers and processes payments. In 2019, PayPal had 305 million active accounts and facilitated US$12.4 billion payment transactions with a total payment volume of US$712 billion, while Venmo processed over $102 billion in 2019 (PayPal, 2020c). PayPal's policies prohibit transactions involving "items that are considered obscene" and "certain sexually oriented materials or services" (PayPal, 2020a, 2(g), 2(i)). PayPal may also terminate its services to merchants if their transactions pose any "legal, reputational, or risk-based reason" (PayPal, 2020d, Section 9.2(3-4)). If the company determines that its seller's "transactions pose an unacceptable level of risk," the platform may refuse to process transactions, hold payments, or limit or suspend the seller's ability to use the payment services (PayPal, 2020d, Section 5.2). Like most platforms, the company reserves the right to suspend or terminate access to any of its services "for any reason and at any time upon notice to you" (PayPal, 2020e).

Online Payment Ecosystem

Given the number of actors potentially involved in a transaction, it can be difficult to determine which actor ordered a transaction to be blocked or account closed. Payment actors tend to blame each other when the disputes become public (see Malcolm, 2017), underscoring the complexity of the payment ecosystem and the opacity which enables actors to shift blame to others.

There is a wide variety of digital payment methods for merchants to accept payments, ranging from services by big US technology firms like Apple Pay and Google Pay to smaller firms like WePay, Stripe, Square, and the Dutch platform Adyen. Despite this diversity, many payment platforms rely upon the traditional brick-and-mortar banks to grant merchant bank accounts and the credit card system to enable customers to pay with credit cards. PayPal's services rely upon the existing credit card networks, as well as the automated clearinghouse, a utility partially operated by the US Federal Reserve, that enables customers to link

their PayPal account directly to their checking account to transfer funds, thereby creating "a private on-ramp to an existing semi-public infrastructure" (Swartz, 2020, p. 39).

As banks and the credit card system provide a vital, largely invisible infrastructure, they set the rules governing the use of their services. Visa and Mastercard are particularly powerful in this regard given their global duopoly on the credit card market outside China, with Visa controlling 60 percent of the credit and debit card market and Mastercard at 30 percent, while American Express trails far behind at about 8 percent (Krauskopf, 2020). Given their dominance, when Visa and Mastercard prohibit transactions relating to specific products or services, all payment providers using their credit card networks must abide by their rules.

Payment providers design their services to make the transfer of funds rapid and easy. Behind the simple interface, however, there is an often–complex network of contractual obligations and rules with multiple actors depending on the payment type. Stripe describes this as living "in a tangled thicket of regulations and restrictions" (Lyon, 2016). To transfer funds between bank accounts, for example, PayPal and Square must abide by rules set by banks. To facilitate payments via credit card, the platforms must obey rules set by credit card networks (i.e., Visa, Mastercard, and American Express), as must the banks that granted the credit card to cardholders (issuing institutions) and the banks that granted the merchant account that accepts credit card payments (acquiring institutions). Square, for example, in the United States offers payment through credit cards operated by the banks JPMorgan Chase and Wells Fargo. As a result, merchants using PayPal or Square must comply with the rules of the platforms, the credit card networks, and acquiring institutions.

As the card networks do not have direct contractual relationships with merchants, they cannot restrict or terminate access to merchant accounts since affiliated banks granted them. Instead, the card networks instruct acquiring institutions like Wells Fargo to investigate suspicious behavior and take appropriate action, which includes the termination of merchant accounts.

American Express has a blanket prohibition on transactions relating to escort services, massage parlors, online adult entertainment, and prostitution (American Express, 2020). Visa and Mastercard permit transactions regarding certain types of adult content and services, provided the transaction is legal in both the buyer's and seller's jurisdictions. They also grant themselves latitude to interpret the suitability of merchants' products and services with regard to how it may affect their brands, prohibiting any transactions that may bring their brands "into disrepute" (Visa, 2020, p. 79). Mastercard's policies are more detailed than Visa's, prohibiting "the sale of a product or service, including an image, which is patently offensive and lacks serious artistic value" and gives as examples of "offensive" non-consensual sexual behavior, sexual exploitation of a minor, or "or any other material" the company deems "unacceptable" (Mastercard, 2019, p. 109).

The credit card networks have interpretive flexibility given their broadly worded policies to determine, in Mastercard's terms, what counts as "offensive," "unacceptable," and "serious artistic value," thereby leading to sometimes

arbitrary determinations of acceptable transactions. In 2017, for example, Visa and Mastercard withdrew their services from the adult social network FetLife, which describes itself as "Facebook but run by kinksters," because, it appears, that the platform offered members a forum to discuss and post depictions of consensual bondage/discipline and sadism/masochism (BDSM) practices (Malcolm, 2017). While it appears that the credit card companies ordered the termination of FetLife's payment services, neither Visa nor Mastercard prohibit transactions relating to consensual BDSM. As of December 2020, FetLife still does not accept payment through credit cards.

As the FetLife case shows, given the lack of transparency and clarity around the credit card companies' rules and decisions, it is not always evident why a specific product or type of content might be banned other than a generic explanation that it is "high risk" or "unacceptable." The payment provider Stripe criticizes blanket prohibitions against certain types of legal content and services, stating that rather than evaluating businesses individually, "many financial institutions choose to make decisions around broad categories of business" in which "restrictions can function more like hammers than scalpels" (Lyon, 2016).

Fines are a key mechanism by which the credit card networks ensure compliance with rules. PayPal, for example, explains that it is subject to fines if the platform fails to detect merchants "engaging in activities that are illegal or considered 'high risk'" (PayPal, 2020f, p 13). PayPal can either prevent high-risk merchants from using its services or register them with the card networks and "conduct additional monitoring" as required (PayPal, 2020f, p 13).

Banned Merchants Database

A key part of the financial ecosystem that affects the regulation of adult content is a proprietary database used by credit card companies that lists terminated merchants, the "MATCH" list (the Member Alert to Control High Risk Compliance Program) (see Mastercard, 2020). This global database, developed and maintained by Mastercard and accessible to American Express, lists merchants worldwide who have violated the credit card companies' rules. Visa has its own Terminated Merchant List database. Credit card companies typically require acquiring institutions to check the databases before granting applicants merchant accounts to see if merchants have had their merchant accounts terminated by another acquirer and the reasons given.

Acquirers may terminate merchant accounts for specific violations, termed "reason codes," that include illegal transactions, violation of standards like engaging in prohibited transactions, and excessive chargebacks (Mastercard, 2020, pp. 123–124). Once terminated for a specific reason code, acquiring institutions must add terminated merchants to MATCH. Terminated merchants are often considered to be a higher risk and less likely to be granted new merchant accounts by other acquiring institutions. Even if an individual has been terminated, another financial institution can decide to grant a merchant account provided the institution requires the merchant to implement certain changes, such as instituting anti-fraud measures.

Problematically, Mastercard does not verify or confirm the accuracy of the information contained in the MATCH system as it relies upon submissions from acquiring institutions (Mastercard, 2020). Blacklisted merchants can only be removed from the database if the information was incorrect, or the acquiring institution verifies that the merchant addressed the underlying problem. Only the submitting acquiring institution can remove faulty information, meaning that inaccurate information persists until removal or until the five-year limit is reached and all information is deleted.

The credit card companies thus play major roles in creating and maintaining the financial infrastructure upon which much of the digital payment system operates. Visa, Mastercard, and American Express set the rules governing the use of their services, which their networks of affiliated financial institutions must obey if they want to offer their credit cards. The power of rule setting is particularly evident with regard to the MATCH database where acquiring institutions globally assess and terminate merchants in accordance with criteria and processes set by Mastercard. Visa and Mastercard, given their dominant market share, can effectively establish payment "chokepoints" where blacklisted entities often have few other payment options (Tusikov, 2017).

"Being Mainstream Is Cheap"

"Being mainstream is cheap while being NSFW is risky," explains Heartpie, a California-based social network where users can purchase individual subscriptions for or crowdsource the creation of original, custom-tailored adult content (Heartpie, 2017). Sexual content and services, along with other issues like gambling, are considered high risk because of the perceived likelihood of fraud or chargebacks, which expose companies to higher costs and potential liability. Chargebacks are a form of disputed transaction in which the customer asks the card issuer to reverse the transaction.

The payment ecosystem becomes more complicated for smaller players like FetLife that want to accept credit card payments. Large merchants like department stores connect directly to acquiring banks, but small merchants have their credit card payments paid through an independent sales organization (ISO) that acts as an intermediary between merchants and acquiring institutions (Swartz, 2020). Payment actors sort merchants into risk categories with those having similar probability of chargebacks priced similarly, and some ISOs specialize in high-risk payments for which they charge merchants higher fees (Swartz, 2020, p. 87). As a result, people selling sexually oriented goods and services typically must pay higher payment processing fees.

Sexual speech is caught in the clash between what Swartz (2020, p. 91) describes as the "traditional" and "platform" models of payment that has significantly changed the management of risk. In the traditional model, cardholders' interests are represented by issuing institutions, while merchants' interests are represented by acquiring institutions (Swartz, 2020, p. 91). There is a market for risk as high-risk transactions create market categories for pricing (Swartz, 2020, p. 96). High-risk transactions aren't banned, they are pricy for merchants.

In contrast, in the platform model, neither buyers nor sellers are directly represented as the client is the platform like PayPal (Swartz, 2020). Unlike traditional ISOs that have customizable contracts with their merchants to accommodate all levels of risk, payment platforms govern their users with blanket terms-of-service contract. Risk is not managed through the market, but banned in terms-of-service agreements (Swartz, 2020, p. 93). High-risk transactions are typically banned because payment platforms like PayPal must guarantee that all their transactions will be low risk for chargeback to qualify for the lowest rates (Swartz, 2020, p. 93).

An entity that can offer a certain scale of transactions can seek out "special adult-friendly payment processors who have an arrangement with their bank to provide merchant accounts to high-risk merchants" (Heartpie, 2017). Adult content merchants can still process payments, but they generally lack access to mainstream providers and low-risk transaction pricing. Individuals and small-scale businesses, however, have few options as neither the traditional nor platform models cater to "high-risk" categories of person-to-person transactions (Swartz, 2020, p. 94).

LAWS ENTRENCHING FINANCIAL DISCRIMINATION

Payment providers' anti-sex stance in the United States has become more entrenched with the passage of the *Fight Online Sex Trafficking Act* (FOSTA) and the *Stop Enabling Sex Traffickers Act* (SESTA). Signed into law in April 2018, SESTA, which was effectively subsumed into FOSTA, controversially expands US federal prostitution laws by making internet platforms legally responsible for third-party content that promote or facilitate prostitution with imprisonment of up to 10 years for platform operators. Prior to FOSTA/SESTA, Section 230 of the 1996 *Communications Decency Act*, referred to as the "safe harbor" law, and one of the most important pieces of legislation that built the internet, granted broad immunity to platforms for some types of speech by users, including sexual expression. These laws rolled back part of this immunity, meaning entities providing web hosting, social media, search, or payment services could be liable for their users' content, even for legal forms of sex work.

Legislators and advocates justified FOSTA/SESTA as necessary to combat human trafficking but, as a mounting body of research demonstrates, the laws have not reduced trafficking but made sex work more dangerous by eliminating protections (Chamberlain, 2019). As critics of the law note, the combination of the laws' overbroad language, strict penalties, and repeal of platforms' previous immunity for content posted by their users mean that platforms will categorically ban entire categories of sexual expression (Chamberlain, 2019, p. 2197).[1]

DIGITAL REDLINING

This power dynamic is evident in the case of Pornhub. In December 2020, Visa and Mastercard pulled their services from Pornhub in response to allegations

that the platform hosted child sexual abuse videos and videos of non-consensual sexual acts, following PayPal's similar move in 2019. Pornhub is a big business, as the largest free, adult-video streaming service, and it makes its money from paid subscriptions and advertising. With the loss of major payment providers, Pornhub has shifted to cryptocurrencies. While the platform will likely continue to generate significant advertising and subscription revenue, the move has hit hard its Model Program, which allows adult performers to upload and earn revenue from videos they create themselves (see Cole, 2019).[2] These performers seemingly operate like contractors and earn revenue independently from Pornhub and, as individuals, have no collective way to negotiate with an ISO specializing in high-risk transactions.

Digital redlining has been used to explain platforms' discriminatory advertising practices that enable real estate brokers and landlords to exclude certain buyers or renters from seeing ads (Karlis, 2019), and the predatory targeting of low-income and minority communities for harmful or inferior products like for-profit education (Gilman, 2019). In finance, digital redlining occurs at the nexus of finance and digital technologies, involves multiple reinforcing policies, and creates differential, discriminatory access to financial services, in part because of the digital divide in broadband internet access, with racialized communities disproportionately affected (Friedline & Chen, 2020).

Digital redlining with regard to sexual speech is the part of a systematic, multifaceted campaign undertaken by multiple actors within the online payment ecosystem that disproportionately affects small businesses and individual operators in the sex and adult entertainment industries (see Blue, 2015a). Regulatory efforts often appear arbitrary, as shown by PayPal's bans of sales relating to escorts' memoirs or corsets discussed in the introduction. However, a common element of many of the media-reported cases of sexual censorship is the targeting of small businesses and individuals, but facilitating business for larger companies. Visa and Mastercard have permitted transactions for pornography businesses, as anti-pornography campaigns against the credit card companies demonstrate (Mohan, 2020). Even sex-adverse American Express is trialing a program allowing users in the United States to pay for a pornography-streaming service (Mohan, 2020).

Companies within the sex and adult entertainment industries, typically larger players that have a certain scale of transactions, may be able to negotiate a contract with an ISO specializing in high-risk transactions or guarantee a low rate of chargebacks. Simply put, bigger actors can better manage risk and afford fees, while smaller operators have few options (see Swartz, 2020). This power dynamic is evident in the case of Pornhub. In December 2020, Visa and Mastercard pulled their services from Pornhub in response to allegations that the platform hosted child sexual abuse videos and non-consensual sexual acts, following PayPal's similar move in 2019. Pornhub is a big business, as the largest free, adult-video streaming service, and it makes its money from paid subscriptions and advertising. With the loss of its major payment providers, Pornhub has shifted to cryptocurrencies. While the platform will likely continue to generate significant advertising and subscription revenue, the move has hit hard its Model

Program, which allows adult performers to upload and earn revenue from videos they create themselves (see Cole, 2019).[2] These performers seemingly operate like contractors and earn revenue independently through Pornhub and, as individuals, have no collective way to negotiate with an ISO specializing in high-risk transactions.

Payment actors' systemic financial discrimination against sex workers is evident even where sex work is decriminalized, such as in the Australian states of Victoria and New South Wales. Australia's Small Business and Family Enterprise Ombudsman, Kate Carnell, accuses financial institutions in Australia of wrongly discriminating against legitimate small businesses in the adult sector, which is estimated at 25,000 people with an annual turnover of AUS$2.6 billion (Simpkins, 2019). Despite the overall size of the sector, its small businesses and individual operators struggle to receive merchant accounts with platforms like PayPal, Square, and Stripe banning or restricting their services related to the sale of sexual services (Sex Work Law Reform Victoria, 2020). As advocacy groups point out, payment actors perceive risk in differing legal approaches to sex work among Australian states: a nationwide decriminalization of sex work could ensure equitable access financial services (see Sex Work Law Reform Victoria, 2020).

FOSTA/SESTA have intensified payment platforms' anti-sex measures as the laws have increased platforms' risk of allowing any content or transactions that may be perceived to promote or facilitate prostitution. The chilling effect predicted by the law's critics is occurring, not only in the United States but also internationally. A 2019 survey undertaken by a US collective of sex workers and allies of 98 internet-based sex workers, mostly in the United States, but also some in Canada, Europe, Australia, and South America found sex workers reported less ability to advertise their services, vet clients, and control their working environments (Blunt & Wolf, 2020). In Vancouver, the PACE Society, which provides sex worker-led programs, finds many United States-based websites have shut down, and without the ability to advertise and digitally screen clients, Canadian sex workers are forced to operate in unsafe conditions (Tierney, 2018).[3]

CONCLUSION

The internet war on sex rages globally with many countries seeking to suppress various forms of sexual speech, but it has distinct roots in the United States given the commercial dominance of US platforms and the role of the United States in shaping rules and standards on the internet that are instituted worldwide (see Tusikov, 2017). Simply put, US norms and rules become universalized into global rules through large US companies and their global networks. As new digital payment methods like Apple Pay, PayPal, or Square are often built upon the existing banking and credit card infrastructure, over which the United States exerts structural power, banks, and credit card companies act as gatekeepers. Visa and Mastercard wield particular structural power given their dominant

market share. When Visa and Mastercard forbid certain types of transactions, the prohibitions reverberate throughout the online payment industry with their networks of banks and payment providers required to institute their rules or face stiff fines. While Visa and Mastercard allow some sexual content, including legal pornography, this permission is contingent upon the credit card networks' interpretation that the transactions are not bringing the companies' brands "into disrepute" (Visa, 2020, p. 79).

This is the essence of payment actors' structural power. People in smaller countries like Canada, Australia, or Mexico which use US platforms and businesses find they must also accept US rules, even when they may conflict with national or local laws.

In the traditional payment model, credit card companies surveilled consumers to manage risk in order to facilitate transactions (see Lauer, 2018). The platform payment model focuses not only on risk management but also consumer surveillance as part of its business model (Swartz, 2020, see also Westermeier, 2020). Platforms' data-intensive business model enables them not only to discern patterns in users' behavior for the development of future products and services, but also to sort users into risk categories based on their sales of sexual content and services. In some cases, platforms have shifted their focus from targeting adult content as a "high-risk *transaction*," to designating those in the sex industry as "high-risk *people*" (Swartz, 2020, p. 97, emphasis in original). PayPal did this in 2010 when it banned dominatrix January Seraph and any business run or owned by her "for life" (Blue, 2015a), a grossly disproportionate and discriminatory penalty that bluntly targets a person not a specific prohibited activity. Payment platforms' mass accumulation of data from different sources, including social media, means that they have a capacity to make inferences about people's economic circumstances and their social activities and a largely opaque and unaccountable power to designate people as high risk.

This chapter highlights the role of dominant US actors in setting rules regarding permitted sexual content, but significant gaps remain in our understanding of payment platforms' regulation of online sexual speech. More empirical research is needed regarding the scope and nature of payment actors' restrictions on sexual expression, particularly among smaller and non-US payment platforms. What political, regulatory, and socio-cultural factors figure into non-US payment platforms' regulation of sexual content? To what extent might there be differing legal or socio-cultural interpretations of permitted and prohibited sexual speech among payment actors, and with what consequences? Cindy Gallop, founder of the female-focused pornography platform Make Love Not Porn, argues that laws and platforms' terms-of-conditions agreements "were written for porn," meaning that there is a "need for a new legal definition of adult content" with regards to "sextech and sex education" (Deighton, 2020). Analysis determining the legal, regulatory, or material incentives that might encourage payment actors to adopt a more nuanced definition of adult content could be a useful first step to address the problems of digital redlining and the financial exclusion of those in the sex and adult entertainment industries.

NOTES

1. Problematically, FOSTA and SESTA conflate consensual sex work with sex trafficking, a common practice strategically employed by opponents of the sex industry, some of whom are prominent civil society organizations with the goal of outlawing all sex work (see Musto, 2016).
2. Pornhub is a controversial site with critics contending that it does not effectively remove revenge porn, which is the non-consensual uploading of porn, or content relating to rape and sexual abuse (Cole & Maiberg, 2020).
3. For a list of websites and online services shut down because of SESTA/FOSTA, see: https://survivorsagainstsesta.org/documentation/

REFERENCES

American Express. (2020, July). Merchant operating guide. United States, including Puerto Rico and the US Virgin Islands. Retrieved from https://icm.aexp-static.com/content/dam/gms/en_us/optblue/us-mog.pdf. Accessed on July 18, 2020.

Andrews, I. (2019, October 30). Cryptocurrency for sex workers and their clients. *Medium*. Retrieved from https://medium.com/@lafemmeisobel/cryptocurrency-for-sex-workers-and-their-clients-bee59466eb46. Accessed on July 11, 2020.

Birch, K., Chiappetta, M., & Artyushina, A. (2020). The problem of innovation in technoscientific capitalism: Data rentiership and the policy implications of turning personal digital data into a private asset. *Policy Studies*, *41*(5), 468–487.

Blue, V. (2015a, December 2). PayPal, Square and big banking's war on the sex industry. *Engadget*. Retrieved from https://www.engadget.com/2015-12-02-paypal-square-and-big-bankings-war-on-the-sex-industry.html. Accessed on July 23, 2020.

Blue, V. (2015b, February 24). Timeline: Google's role in global sex censorship. *ZDNet*. Retrieved from https://www.zdnet.com/article/timeline-googles-role-in-global-sex-censorship/. Accessed on July 22, 2020.

Blue, V. (2018, December 7). The internet war on sex is here. *Engadget*. Retrieved from https://www.engadget.com/2018-12-07-the-internet-war-on-sex-is-here.html. Accessed on July 20, 2020.

Blunt, D., & Wolf, A. (2020). Erased: The impact of FOSTA–SESTA and the removal of Backpage on sex workers. *Anti-Trafficking Review*, *14*, 117–121.

Bridy, A. (2015). Internet payment blockades. *Florida Law Review*, *67*(5), 1523–1568.

Chamberlain, L. (2019). FOSTA: A hostile law with a human cost. *Fordham Law Review*, *87*, 2171–2211.

Chemaly, S. (2015). Slut-shaming and the sex police: Social media, sex, and free speech. In S. Tarrant (Ed.), *Gender, sex and politics: In the streets and between the sheets in the 21st century* (pp. 125–140). New York, NY: Routledge.

Cole, S. (2019, November 14). PayPal pulls out of Pornhub, hurting 'hundreds of thousands' of performers. *Vice*. Retrieved from https://www.vice.com/en_ca/article/d3abgv/paypal-pulls-out-of-pornhub-payments. Accessed on July 3, 2020.

Cole, S., & Maiberg, E. (2020, February 6). Pornhub doesn't care. *Vice*. Retrieved from https://www.vice.com/en_us/article/9393zp/how-pornhub-moderation-works-girls-do-porn. Accessed on August 3, 2020.

Deighton, K. (2020, May 10). Why sex startups face uphill payments battle. *Raconteur*. Retrieved from https://www.raconteur.net/finance/sextech-payments. Accessed on July 11, 2020.

DeLacey, H. (2020, May 10). The precaritisation of sex worker during the COVID-19 pandemic. *Leiden Law Blog*. Retrieved from https://leidenlawblog.nl/articles/the-precarisation-of-sex-workers-during-the-covid-19-pandemic. Accessed on July 16, 2020.

Do Thi Duc, H. (2017). Public by default. Retrieved from https://22-8miles.com/public-by-default/. Accessed on July 25, 2020.

Faust, G. (2017). Hair, blood and the nipple. In U. Undine Frömming, S. Köhn, S. Fox, & M. Terry (Eds.), *Digital environments: Ethnographic perspectives across global online and offline spaces* (pp. 159–170). Bielefeld: Transcript-Verlag.

Friedline, T., & Chen, Z. (2020). Digital redlining and the Fintech marketplace: Evidence from US zip codes. *The Journal of Consumer Affairs*, https://doi-org.ezproxy.library.yorku.ca/10.1111/joca.12297.

Gilman, S. (2019). Proliferating predation: Reverse redlining, the digital proliferation of inferior social welfare products, and how to stop it. Retrieved from https://ssrn.com/abstract=3511892. Accessed on July 10, 2020.

Heartpie. (2017, May 8). nsfw's payment processing problem. *Medium*. Retrieved from https://medium.heartpie.com/nsfws-payment-processing-problem-aa0d53c0ea2f. Accessed on July 8, 2020.

Karlis, N. (2019, August 21). "Digital redlining:" Faceook's housing ads seem designed to discriminate. *Salon*. Retrieved from https://www.salon.com/2019/08/21/digital-redlining-facebooks-housing-ads-seem-designed-to-discriminate/. Accessed on July 17, 2020.

Krauskopf, L. (2020, January 31). Swiping their way higher: Visa, Mastercard could be the next $1 trillion companies. *Reuters*. Retrieved from https://www.reuters.com/article/us-visa-Mastercard-stocks/swiping-their-way-higher-visa-Mastercard-could-be-the-next-1-trillion-companies-idUSKBN1ZU0JA. Accessed on July 10, 2020.

Laidlaw, E. (2015). *Regulating speech in cyberspace: Gatekeepers, human rights and corporate responsibility*. Cambridge: Cambridge University Press.

Lauer, J. (2018). *Creditworthy: A history of consumer surveillance and financial identity in America*. New York, NY: Columbia University Press.

Liang, F., Das, V., Kostyuk, N., & Hussain, M. M. (2018). Constructing a data-driven society: China's social credit system as a state surveillance infrastructure. *Policy & Internet*, *10*(4), 415–453.

Lyon, D. (2016, August 12). Why some businesses aren't allowed. *Strip Blog*. Retrieved from https://stripe.com/blog/why-some-businesses-arent-allowed. Accessed on July 12, 2020.

Malcolm, J. (2017, March 15). Payment processors are still policing your sex life, and the latest victim is FetLife. *Electronic Frontier Foundation*. Retrieved from https://www.eff.org/deeplinks/2017/03/payment-processors-are-still-policing-your-sex-life. Accessed on July 2, 2020.

Martin, A. (2019). Mobile money platform surveillance. *Surveillance & Society*, *17*(1/2), 213–222.

Mastercard. (2019). Mastercard rules. Last updated December 19. Retrieved from https://www.Mastercard.us/content/dam/mccom/global/documents/Mastercard-rules.pdf. Accessed on July 18, 2020.

Mastercard. (2020, February 27). Security rules and procedures. *Merchant Edition*. Retrieved from https://www.merchantmaverick.com/learning-terminated-merchant-file-tmf-aka-match-list/. Accessed on July 27, 2020.

McNamara, H. (2020, February 24). Credit card companies should stop partnering with porn websites. *Washington Examiner*. Retrieved from https://www.washingtonexaminer.com/opinion/op-eds/credit-card-companies-should-stop-partnering-with-porn-websites. Accessed on July 27, 2020.

Mohan, M. (2020, May 8). Call for credit card freeze on porn sites. *BBC*. Retrieved from https://www.bbc.com/news/world-52543508. Accessed on August 5, 2020.

Musto, J. (2016). *Control and protect: Collaboration, carceral protection, and domestic sex trafficking in the United States*. Oakland, CA: University Press.

PayPal. (2020a). Acceptable use policy. Last updated March 19. Retrieved from https://www.paypal.com/us/webapps/mpp/ua/acceptableuse-full?locale.x=en_US. Accessed on July 19, 2020.

PayPal. (2020b). Privacy statement. Effective May 13. Retrieved from https://www.paypal.com/us/webapps/mpp/ua/privacy-full?locale.x=en_US. Accessed on July 11, 2020.

PayPal. (2020c, January 29). PayPal reports fourth quarter and full year 2019 results. Retrieved from https://www.paypal.com/stories/us/paypal-reports-fourth-quarter-and-full-year-2019-results. Accessed on July 27, 2020.

PayPal. (2020d). Platform Seller Account Agreement, Section 9.2(3 and 4). Last updated April 2. Retrieved from https://www.paypal.com/us/webapps/mpp/ua/platform-seller-agreement?locale.x=en_US. Accessed on July 19, 2020.

PayPal. (2020e). User agreement. Effective as of April 28. Retrieved from https://www.paypal.com/us/webapps/mpp/ua/useragreement-full?locale.x=en_US. Accessed on July 19, 2020.

PayPal. (2020f, April 8). 2020 Annual Meeting of Stockholders and Proxy Statement & 2019 Annual Report. Retrieved from https://investor.paypal-corp.com/static-files/6b4a31d7-9941-464d-846d-3859fd7058dc. Accessed on July 28, 2020.

Roberts, S. (2019). *Behind the screen: Content moderation in the shadows of social media*. New Haven, CT: Yale University Press.

Sex Work Law Reform Victoria. (2020). Financial institutions: Which ones discriminate? Retrieved from https://sexworklawreformvictoria.org.au/financial-institutions-which-ones-discriminate/. Accessed on August 1, 2020.

Simpkins, S. (2019, November 18). Sex worker blasts fintechs for discrimination. *Fintech Business*. Retrieved from https://www.fintechbusiness.com/payments/1586-sex-worker-blasts-fintechs-for-discrimination. Accessed July 19, 2020.

Srnicek, N. (2017). *Platform capitalism*. Cambridge: Polity Press.

Swartz, L. (2020). *New money: How payment became social media*. New Haven, CT: Yale University Press.

Tierney, A. (2018, April 12). How the US 'sex trafficking' crackdown is hurting sex workers in Canada. *Vice*. Retrieved from https://www.vice.com/en_ca/article/9kggwe/how-the-us-sex-trafficking-crackdown-is-hurting-sex-workers-in-canada. Accessed on August 5, 2020.

Tumblr. (2020). Community guidelines. Last modified July 16. Retrieved from https://www.tumblr.com/policy/en/community. Accessed on July 17, 2020.

Tusikov, N. (2017). *Chokepoints: Global private regulation of the internet*. Oakland, CA: University of California Press.

Tusikov, N. (2019). Defunding hate: PayPal's regulation of hate groups. *Surveillance & Society*, *17*(1/2), 46–53. https://doi.org/10.24908/ss.v17i1/2.12908

Van Dam, A. (2020, July 2). Black families pay significantly higher property taxes than white families, new analysis shows. *Washington Examiner*. Retrieved from https://www.washingtonpost.com/business/2020/07/02/black-property-tax/. Accessed on July 22, 2020.

van Dijck, J., Poell, T., & de Waal, M. (2018). *The platform society: Public values in a connective world*. New York, NY: Oxford University Press.

Venmo. (2020). Privacy policy. Last updated June 30. Retrieved from https://venmo.com/legal/us-privacy-policy/. Accessed on July 10, 2020.

Visa. (2020, April 18). Visa core rules and visa product and service rules. Version 1.1. Retrieved from https://usa.visa.com/support/consumer/visa-rules.html. Accessed on July 27, 2020.

Westermeier, C. (2020). Money is data: The platformization of financial transactions. *Information, Communication & Society*, *23*(4), 2047–2063.

Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the New Frontier of Power*. New York, NY: Public Affairs.

LEGISLATION

Allow States and Victims to Fight Online Sex Trafficking Act, 1H.R. 1865, 115th Congress. Retrieved from https://www.govtrack.us/congress/bills/115/hr1865/text

Communications Decency Act. S. 314. 47 USC. § 230, 104th Congress (1995–1996). Retrieved from https://www.congress.gov/bill/104th-congress/senate-bill/314/

Stop Enabling Sex Traffickers Act of 2017, S.1693, 115th Congress (2017–2018). Retrieved from https://www.congress.gov/bill/115th-congress/senate-bill/1693

CHAPTER 5

GAFAM AND HATE CONTENT MODERATION: DEPLATFORMING AND DELETING THE ALT-RIGHT

Tanner Mirrlees

ABSTRACT

Purpose – *This chapter demonstrates the power that Google, Apple, Facebook, Amazon and Microsoft (or the "GAFAM") exercise over platforms within society, highlights the alt-right's use of GAFAM sites and services as a platform for hate, and examines GAFAM's establishment and use of hate content moderation apparatuses to de-platform alt-right users and delete hate content.*

Approach – *Drawing upon a political economy of communications approach, this chapter demonstrates GAFAM's power in society. It also undertakes a reading of GAFAM "terms of service agreements" and "community guidelines" documents to identify GAFAM hate content moderation apparatuses.*

Findings – *GAFAM are among the most powerful platforms in the world, and their content moderation apparatuses are empowered by the US government's cyber-libertarian approach to Internet law and regulation. GAFAM are defining hate speech, deciding what's to be done about it, and censoring it.*

Value – *This chapter probes GAFAM's hate content moderation apparatuses for Internet platforms, and shows how GAFAM enable and constrain the alt-right's hate speech on their platforms. It also reflexively assesses the politics of empowering GAFAM to de-platform the alt-right.*

Keywords: GAFAM; platforms; content moderation; hate speech; censorship; Alt-right

Media and Law: Between Free Speech and Censorship
Sociology of Crime, Law and Deviance, Volume 26, 81–97

ISSN: 1521-6136/doi:10.1108/S1521-613620210000026006

INTRODUCTION

From the election of Donald Trump in 2016 to the present day, something called the "alt-right" has been using Internet platforms to produce, spread, and consume hateful anti-Semitic, Islamophobic and anti-Black messages and imagery, far and wide (Neiwert, 2018; Wendling, 2018). Headlines such as "YouTube's neo-Nazi music problem" (BBC, 2018) and "Democratic Senators Demand Facebook Answer For Its White Supremacist Problem" (Durkee, 2020) recognize the alt-right's platform presence as a social problem. In the United States and elsewhere, public officials, advertisers and citizens have pressured the owners of Internet platforms to counter the alt-right's cyber-hate. The biggest of these – Google, Apple, Facebook, Amazon and Microsoft (or "GAFAM") – have established and operationalized hate content moderation apparatuses to investigate alt-right interactions and expressions, ban their accounts and delete millions of pieces of hateful content.

As GAFAM shutter some alt-right influencers and shut some noxious expressions out of their platforms, celebrations and denunciations of GAFAM's de-platforming of hate speech mount. Yet, the social conditions that empower GAFAM to ban alt-right users and remove posts from their platforms are often obscure. The institutional policies and mechanisms these corporations wield to define, identify and respond to hate speech are likewise opaque, and often hidden from view. This chapter sheds light on GAFAM's hate content moderation apparatuses with an overview of the economics and politics of these corporations' hate speech policies, human and technological identification, reporting and enforcement mechanisms, and sanctions. The chapter's first section gives a snapshot of GAFAM's ubiquitous power in society, and reviews the alt-right's use of GAFAM platforms to create, impart, and receive hateful content. The second section explains how the US Federal government's "cyber-libertarian" approach to media and Internet law and regulation has empowered GAFAM's development of hate content moderation apparatuses for their Internet platforms. The third section identifies GAFAM's hate content moderation apparatuses through a reading of their "terms of service agreements" and "community guidelines" regarding hate speech, and illustrates GAFAM's operationalization of these apparatuses with some recent examples of the alt-right being de-platformed and deleted. The conclusion assesses the efficacy of GAFAM's content moderation "solution" to the problem of alt-right hate speech on platforms in light of political arguments about "what's to be done" to resist the alt-right.

THE RISE OF GAFAM, PLATFORMS AND THE ALT-RIGHT

The United States is arguably the world center of "cyber-libertarianism" or the idea that the Internet, the World Wide Web and related digital communication technologies can and should be a space where individuals freely interact and express themselves in whatever way they choose, absent interference by Federal

law and regulatory action (Dahlberg, 2019). Although the US State has shaped and supported the Internet from its earliest days and subsidized all kinds of digital developments, each new Silicon Valley venture seems to recharge the ideological battery of cyber-libertarianism (Dahlberg, 2019). Yet, digital capital, not the State, have long exercised the most power over the modern Internet (McChesney, 2013; Schiller, 2014). For two decades, GAFAM, not governments, concentrated their ownership over innumerable platforms, merging with and incorporating them into their business operations, and the GAFAM now preside over an Internet oligopoly (Smyrnaiois, 2018). The Big Five operate the biggest Internet platforms in the world, and the billions who use their sites and services must abide by their executive decisions and follow their platform rules.

GAFAM are at the center of a globalizing "platform capitalism" (Srniceck, 2017) and "platform imperialism" (Jin, 2019), and these corporations "shape the core technological infrastructure, economic models, and ideological orientation of the [global platform] ecosystem as a whole" and "steer how sectoral platforms, societal institutions, companies and billions of users interact" for profit-oriented ends (van Dijck, Poell, & de Waal 2018, p. 32). In 2019, the "Big Five" ranked among the largest and most valuable global publicly traded corporations in the world. In 2020, they reached a combined total market capitalization valuation of $6.4 trillion, a 53% increase from the previous year (Tucker, 2020). GAFAM sell products and services thorough their platforms, which intermediate between users, customers, service providers and advertisers (Srniceck, 2017). Their platforms are also designed to meet capitalism's need for an efficient advertising system with "a programmable digital architecture" that organizes "interactions between users" and monitors, collects, processes, circulates and monetizes the data of billions of users worldwide (van Dijck et al., 2018, p. 4). As platform capitalists, GAFAM are at the helm of a "new business model capable of extracting and controlling immense amounts of data" about people to turn a profit (Srniceck, 2017, p. 6). Google, Facebook and Amazon take in more than half of the world's total Internet advertising revenue when using the data they extract from users to algorithmically sell advertisers customized batches of user attention. The net worth of GAFAM's ruling class is immense: Google's founders Larry Page and Sergey Brin ($117.6 billion); Apple CEO Tim Cook ($1.02 billion); Facebook's Mark Zuckerberg ($68.2 billion); Amazon CEO Jeff Bezos ($138.2 billion); and Microsoft founder Bill Gates ($104.3 billion) (Forbes, 2020).

In addition to being economic entities, GAFAM are political entities, communicative spaces and instruments. Like all corporations, they lobby US government agencies and politicians to support laws and regulations that's bolster their bottom line, and from 2005 to 2018, they spent over half a billion dollars pressuring Congress (Polat, 2019). Frequently, the US State buttresses GAFAM's oligopoly powers abroad and generously dolls out tax breaks and subsidies to them at home. At the same time, GAFAM's platforms are mediated battlespaces upon which politocos of the Left, Centre and Right fight for recognition and attention, try to organize consent to their ideologies, and seek to influence how people think and behave. While all kinds of political actors log on to and wield

GAFAM's platforms for strategic communications, far from being neutral tools or passive intermediaries, these platforms are made to shape the uses people make of them, and to act upon their users. CEOs, Board Members and managers press these platforms into profit-making on behalf of shareholder values, and their legal teams, engineers and programmers make design choices about what users may do with platforms, and what platforms do to them.

GAFAM are also "cultural" entities, embedded in the ways billions of people interact, commune and create each day. When they launched a little over 15 years ago, GAFAM platforms seemed like a fleeting fad, but by 2020, they were everywhere and among the top 30 most visited websites in the world (Clement, 2020). In 2020, about 4.5 billion people worldwide were using the Internet, and 3.8 billion of these were using at least one social media platform, with Facebook and YouTube being the most popular (Clement, 2020). GAFAM platforms also provide a new means for the production, circulation and consumption of all kinds of digital cultural works. In 2005, when the cyber-entrepreneur Tim O'Reilly touted a radical shift from Web 1.0 to Web 2.0 as the dawn of a new Internet that was more user-friendly, interactive and participatory, platforms were celebrated as "Web 2.0" writ large and cheered for empowering the once passive media consumer to become an interactive producer and mass communicator of media (the "prosumer"). Platforms such as Facebook, YouTube and Twitter are indeed "built around the convergence of content sharing, public communication and interpersonal connection" (Burgess, Marwick, & Poell, 2018, p. 1), and they support personalized one-to-one and many-to-many communication practices that have increased "our ability to share, to co-operate with one another" (Shirky, 2008, p. 20). Nowadays, billions of people, or platform "users," are creating, imparting and receiving abundant comments, essays, stories, photos, memes and videos about everything from why Donald Trump is a neo-fascist to the Proud Boys' idolatry of his cult of personality.

Evidently, GAFAM's platforms are not separate from society, but integrated with every sphere – economic, political and cultural. When GAFAM's platforms launched in the early 2000s, cyber-libertarians hoped they would bring about a different and much better kind of society then before (Morozov, 2014). Yet, instead of doing that, GAFAM's platforms have mostly propped up the capitalist order, served State power and big parties, and at worse, supported the rise of the "alt-right," an umbrella term for white supremacist hate groups that reject "tolerant" neoconservatism and progressive neoliberalism, wish to dismantle the Republican and Democratic "establishment," and strive to build an ethno-State that compels all of society's institutions to protect and promote the values of an idealized white race against the supposedly culturally diluting "threat" of non-white Others (Hawley, 2017; Neiwert, 2017; Wendling, 2018). During Barack Obama's two-term presidency, the alt-right was busily building a cross-border digital media network of its own and organizing consent to its hateful ideas across GAFAM's platforms (Wendling, 2018).

Adherents to the alt-right assembled on message boards such as 4chan and 8chan's/pol/and discussion forum sites such as Reddit, and from massively populated platforms, such as Facebook, Twitter and YouTube, they launched

trans-national campaigns for propagandizing, fundraising, recruiting and mobilizing (Berger, 2018; Conway, Scrivens, & Macnair, 2019; Hawley, 2018; Lewis, 2018; Wendling, 2018; Wong, 2019). Alt-right propagandists have used the Internet and GAFAM's platforms to spread their ideology, become more audible and visible, amplify movement leaders, engage, interpellate, and influence new recruits, and coordinate cross-border communities of white supremacy. Concurrently, the platform business model and search algorithm arguably reproduce white nationalist ideology by blowing alt-right "filter bubbles" (Bryant, 2020). The more one searches for, views and clicks through alt-right sites, posts, works, channels and videos, the more the platform search algorithm recommends to users these hateful media sources (Conway et al., 2019; Lewis, 2018; Wendling, 2018; Wong, 2019). Having "connected" and "shared" with one another, some alt-right acolytes go on "to commit hate crimes and terrorism" (Conway et al., 2019, p. 1). In 2018, people who subscribed to white supremacist ideology killed at least 40 people in North America and in that same year, these ideologues perpetrated 39 of 50 extremism-related killings in the United States (Friedman, 2019).

While the alt-right uses platforms to produce and consume hateful content, attract new foot soldiers to its ranks, commune across geographical distances, and cheer on hate crimes and even terroristic violence, the alt-right's platform presence is being resisted by a new "digital united front" (Mirrlees, 2019). Platforms are neither inherently far-Right, far-Left or Centrist in their designs, uses and impacts, but instead, constitute a dynamic, multifaceted and evolving communications battlespace. So, while GAFAM's platforms are frequently associated with the alt-right's rise, they may also be serviceable to bringing it down. As of late, GAFAM and other Internet owners have established hate content moderation apparatuses and operationalized these to de-platform alt-right ideologues and delete the hateful content they "prosume." Given how powerful GAFAM are, it is important to consider how they got into the business and politics of policing and punishing the hateful communicative actions of platform users. The next section looks at how the US State's cyber-libertarian legal and regulatory framework empowered GAFAM to do that.

HATE SPEECH, CYBER-LIBERTARIANISM AND CONTENT MODERATION

Self-Policing Platforms

In many democratic countries – Australia, Canada, Denmark, France, Germany, New Zealand and the United Kingdom – the lion's share of the communications and media systems is not owned or controlled by States. Rather, they are capitalist, but subject to governmental laws, policies and regulations oriented to the public interest. In these countries, States prohibit the free flow of hate speech in print and broadcasting media, and some have extended this form of censorship to the Internet and social media platforms (Howard, 2019). Around the world, many States define, criminalize and censor mediatized hate speech, but the United States is mostly an exception to cross-border hate speech rules.

In the United States, State-enshrined hate speech laws and efforts to censor hate speech would likely be unconstitutional. While the Constitution does not protect speech that incites imminent lawless action (e.g., an anti-Semite's posting of a manifesto that directly motivates a neo-Nazi to assault or murder a Jewish person) or true threats (e.g., a white supremacist's call to perpetrate real violence against Black Americans), the First Amendment largely protects the rights of American citizens to express hatred toward one another and hear it. The United States does not have an official Federal hate speech Law, and the US Supreme Court has long ruled that hate speech is protected free speech (Howard, 2019). In the United States, negative liberty (individual freedom from the State's interference in society and placing of constraints on agency) has long been privileged over positive liberty (individual freedom to self-actualize with help from a State's intervention in society and placing of limits on agency). Arguably, hate speech infringes upon the 14th Amendment's guarantee of equal protection under the law and affronts democratic values (Tsesis, 2009), but in the Federal and judicial main, the American citizen's constitutional right to be free from State censorship of speech defined as hateful trumps the citizen's right to be free of speech that targets or afflicts them with hatful invective. In the United States, the work of governing mediatized hate speech is exempted from State and public media law and regulatory entities and left to the decision-making powers of market-oriented and private media corporations. The US Federal government's laissez-faire media and Internet legal and regulatory apparatus empowers broadcast and Internet corporations to govern and choose for themselves what to do about hateful expressions.

The US Federal Communications Commission (FCC) is an independent Federal agency overseen by Congress that is responsible for implementing and enforcing America's communications law and regulations for radio, TV, wire, satellite, and cable within a jurisdiction encompassing all American interstate, national and international communications (FCC, 2020a). The FCC's own website clarifies it will not censor messages or images by TV or radio networks, stations or channels, their workers or their guests that may be perceived as "extreme, incorrect or somehow improper political, economic or social statements," nor will the FCC prohibit the flow of statements that may "criticize, ridicule, 'stereotype' or demean individuals or groups because of the religion, race, nationality, gender, gender identification, or sexual orientation, or other characteristics of the group or individual" (FCC, 2020b). Barred by law from censoring any point of view, the FCC takes a hand's off approach to the problem of hate speech. It seems to abide by the idea that "the public interest is best served by permitting free expression of views" and assumes or hopes that a wide range of "opposing opinions will be expressed, even though some may be highly offensive." The FCC enables convergent media conglomerates, such as the Walt Disney Company, Comcast and ViacomCBS, to decide for themselves and for society what's fit for print and broadcast. And this Federally enshrined laissez-faire approach also means that they and other Big media entities may prohibit or censor voices, views and statements for whatever reason they decide. Overall, this Federal "corporate libertarian" approach to media law and regulation has empowered corporations

to govern themselves and the content they impart and receive, to the detriment of even the weakest public interest safeguards and redistributive mechanisms (Pickard, 2013).

The Federal government's corporate libertarian media framework extends to the corporations that provide access to the Internet (e.g., an Internet Service Provider) and other digital services (e.g., a social networking platform). In *Reno* v. *ACLU* (1997), the US Supreme Court ruled that speech on the Internet is entitled to the highest level of First Amendment protection, and so the State cannot prohibit any Internet corporation from letting people use its service or site to create, impart and receive hateful content. Stormfront.org is given as much protection as some print copy of Richard Spencer's *The Uprooting of European Identity* or an American Identity Movement pamphlet. That said, Internet corporations are not legally obliged to uphold the First Amendment rights of their users, but the law does empower them to support or suppress their platform users' free expression, and without the liabilities of a broadcaster or publisher. The legal centerpiece of this power is a section in the short-lived Communications Decency Act of 1996, which had tried to make the display or circulation of obscene or indecent material online to anyone under 18 years of age a criminal act (Gillespie, 2018, p. 30). Known as Section 230, this section was adopted by Congress to support maximal free speech on the World Wide Web and encourage Internet corporations and their service users to "self-police." Section 230 treats Internet corporations as intermediaries between two or more communicators that are not liable for the *content* communicated on their sites, but it also grants these corporations the right to act upon users and intervene in the content they communicate with one another, hateful or not. Simply put, Section 230 lets Internet corporations "claim 'the right but not the responsibility' to remove users and delete content," and largely protects these corporations "from as much legal liability as possible" while ensuring they have "the most discretionary power" (Gillespie, 2018, p. 31). Basically, Internet corporations can choose to take a hand's off or hand's on approach to users and the content they generate, impart and receive through their sites and services.

The US State's cyber-libertarian Internet law and regulatory framework empowers Internet corporations, not the government or public officials, to do as they like with the user interactions and expressions contained by their proprietary platforms. In that regard, the Internet is in no way *absolutely* empowering of individual freedom and freewheeling expression. There is no corporate-run Internet site or service in the modern world "that does not impose rules" upon the interactive agency and expression of the people who use them (Gillespie, 2018, p. 5). Most platforms will not even allow people to use them without first eliciting their consent to have their communicative freedom circumscribed. By requiring users to click "I agree" or "I accept" to "terms of service" and defining appropriate and inappropriate communicative behavior and expression via "community guidelines," corporations decide what is and is not allowed or permitted on their platforms (Gillespie, 2018, p. 46). While Internet corporations might prefer it "if either the community could police itself or, even better, users never posted objectionable content in the first place" (Gillespie, 2018, p. 5),

they are powerful "setters of norms, interpreters of laws, arbiters of taste, adjudicators of disputes, and enforces of whatever rules they choose to establish" (Gillespie, 2018, p. 5).

In sum, Internet corporations have developed content moderation apparatuses that combine a normative communicative interaction and content framework with an enforcement and sanction mechanism. Using content moderation apparatuses, Internet corporations try to govern user interactions and expressions related to the "sexual" (nudity, sex and pornography), the "graphic" (violence and obscenity), the harassing (abuse, trolling and true threats) and the "hateful" (Gillespie, 2018, pp. 45–73). The following section describes GAFAM's hate content moderation apparatuses and shows these corporations' attempt to de-platform the hateful conduct and content of alt-right users.

GAFAM'S HATE CONTENT MODERATION APPARATUSES

Policy, Enforcement and Sanction

Over the past decade, non-profit organizations, non-US governments, advertisers and activists worldwide have invited, prodded and demanded GAFAM to do something about the spread of hate speech on their platforms. The Anti-Defamation League (ADL, 2014) released a "Best Practices for Responding to Cyberhate" doctrine that urged "the Internet Community, including providers, civil society, the legal community and academia, to express their support for this effort" (ADL, 2014). That year, some Silicon Valley titans responded with public pledges to combat cyber-hate. Google announced that to "maintain a safe and vibrant community across our platforms, we offer tools to report hateful content, and act quickly to remove content that violates our policies" (Google, 2014). Facebook stated its support for "efforts to address and counter cyber hate" and committed to "creating a safe and respectful platform for everyone" (ADL, 2014). Microsoft affirmed it aimed to provide a "safe and enjoyable online experience for our customers" and would enforce "policies against abuse and harassment on our online services" (ADL, 2014).

In the wake of the 2016 Trump election, the 2017 Unite the Right rally, and the 2020 BLM uprising, GAFAM faced immense pressure to more aggressively target and combat digital hate. On May 31, 2016, the European Union worked with Google, Facebook and Microsoft to create a (non-legally binding) code of conduct mandating they review the majority of valid notifications for removal of hate speech posted on their platforms within 24 hours (Hern, 2016). Some States subsequently developed their own sovereign legal and regulatory mechanisms for compelling GAFAM platforms to do this. In early 2018, Germany began enforcing the Network Enforcement Act (NetzDG), which requires Internet companies to remove illegal hateful posts from their platforms within 24 hours, or face up to 50 million euros in fines. From 2020 forward, Germany has required GAFAM to report to the Office of the Federal Criminal Police (BKA) any user post that promulgates far-right hate propaganda, threatens violence or prepares for a

terrorist attacks (Guerrini, 2020). In June 2020, numerous civil rights groups, non-profit organizations and companies launched "#StopHateForProfit," a global campaign to boycott Facebook (Wong, 2020). Over 1,000 corporations paused their advertising expenditure on Facebook in July 2020 to cajole it to do more to counter digital hate. Accompanying these government and corporate moves but different from them are anti-fascist activist campaigns to engage and combat alt-right-related hate sites, pages, personas and propaganda across Internet platforms (Mirrlees, 2019).

Over the past few years, GAFAM have crafted and added *explicit* and *tacit* "hate speech" policies to their content moderation apparatuses, and they've used these to act upon hateful user behaviors and expressions. Whereas the former is openly defined by the corporation as a "hate speech policy," the latter tends to be embedded in the sections and clauses of "terms of service" documents and "community guidelines" pertaining to objectionable, offensive or discriminatory content. GAFAM are enforcing these policies by acting upon the users responsible for hate speech, sometimes on their own terms, and sometimes in consultation with hate speech experts and other interested publics.

In a 2019 HBO interview, Google CEO Sundar Pichai said Google must do more to tackle hate speech. Google's most popular platform is YouTube, the world's largest video sharing site. On the YouTube Policies page for "Hate Speech Policy," Google declares "Hate speech is not allowed on YouTube" and says it will "remove content promoting violence or hatred against individuals or groups" based on attributes such as "age, caste, disability, ethnicity, gender identity, nationality, race, immigration status, religion, sex/gender, and sexual orientation" (Google, 2020a). For Google, hate speech encompasses "content that dehumanizes individuals or groups with these attributes, claims they are physically or mentally inferior, or praises or glorifies violence against them." Google's hate speech policy also stipulates that YouTube does not "allow use of stereotypes that incite or promote hatred based on these attributes, or racial, ethnic, religious, or other slurs where the primary purpose is to promote hatred." Furthermore, Google "prohibits content that alleges the superiority of a group over those with any of the attributes noted above to justify violence, discrimination, segregation, or exclusion." If a Google user's channel, page or content violates this policy, Google will "remove the content" and send a warning to the user. For every subsequent violation, Google gives the user a "strike," and if a user amasses three, they get pushed out of YouTube. Additional sanctions include Google's limiting of a user's "YouTube features" for content that "comes close to hate speech" and termination of a user's channel or account "due to a single case of severe abuse" (Google, 2020b).

In 2018, Apple CEO Tim Cook said "we [Apple] only have one message for those who seek to push hate, division and violence: You have no place on our platforms" (cited in Feiner, 2018). The "Apple Media Services: Terms and Conditions" document outlines Apple's tacit hate speech policy for iTunes, the App Store, Apple Books, Apple Music, Apple Podcasts and all other Apple platforms (Apple, 2020a). Apple users may "submit or post materials such as comments, ratings and reviews, pictures, videos, and podcasts" to some of its platforms, but

those who do must "comply with the Submissions Guidelines" which "may be updated from time to time" (Apple, 2020a). The Guidelines stipulate that users "may not use the Services to" post or spread "objectionable, offensive, unlawful, deceptive, inaccurate, or harmful content" and declares that Apple "may monitor and decide to remove or edit any submitted material" (Apple, 2020a). The closest Apple comes to articulating an explicit hate speech policy is in its App Store's "Review Guidelines" (a section on "Safety" pertaining "Objectionable Content") (Apple, 2020b). Apple states that "Apps should not include content that is offensive, insensitive, upsetting, intended to disgust, in exceptionally poor taste, or just plain creepy" and notes it will reject content that is

> Defamatory, discriminatory, or mean-spirited content, including references or commentary about religion, race, sexual orientation, gender, national/ethnic origin, or other targeted groups, particularly if the app is likely to humiliate, intimidate, or harm a targeted individual or group. (Apple, 2020b)

Apple "does not tolerate hate speech," and the users, creators and developers who don't abide by its Terms and Guidelines may have their expressions deleted and accounts terminated.

In 2020, Facebook head Mark Zuckerberg declared "I stand against hate or anything that incites violence" (cited in Vega, 2020). Facebook "Community Standards" are detailed, and the "Hate Speech Policy Rationale" section explains that Facebook does not allow hate speech "because it creates an environment of intimidation and exclusion and in some cases may promote real-world violence." Facebook defines "hate speech as a direct attack on people" based on "protected characteristics" that encompass "race, ethnicity, national origin, religious affiliation, sexual orientation, caste, sex, gender, gender identity, and serious disease or disability." Facebook defines "attack" as "violent or dehumanizing speech, harmful stereotypes, statements of inferiority, or calls for exclusion or segregation." Facebook delineates specific "attacks" by "three tiers of severity," each which is a detailed registry of all kinds of hateful expressions. For example, Tier 1 attacks include "dehumanizing comparisons" (e.g., comparing people to insects, filth, disease and feces) and "designated dehumanizing comparisons" (e.g., "Black people and apes," "Jewish people and rats" and "Muslim people and pigs"). Tier 2 attacks include "Generalizations that state inferiority" as related to physical appearance, physical, mental and moral deficiency, and "Expressions of contempt and disgust." Tier 3 attacks encompass overt "Calls for segregation" and "Explicit Exclusion," as well as targeted slurs. Facebook's compendium of hate attacks highlights hundreds of additional types of hateful interactions and expressions prohibited by the platform.

Amazon CEO Jeff Bezos once said "It's ugly speech that needs protection" (cited in Calhoun, 2017), but Amazon doesn't protect hate speech on its platforms. The corporation's tacit hate speech policy is encapsulated by its "Offensive and Controversial Materials" guidelines for vendors who sell and post content about "all products except books, music, video and DVD" (Amazon, 2018) as well as its "Amazon Community" guidelines for consumer users who share "authentic feedback about products and services – positive or negative" with

other of the online store. Amazon prohibits vendors from selling "products that promote, incite or glorify hatred, violence, racial, sexual or religious intolerance or promote organizations with such views," and it also bars vendors for posting *listings* for products that "promote, incite, or glorify hate or violence towards any person or group." If a vendor violates Amazon's hate speech policy, the corporation may remove their product listing, suspend or terminate their account, withhold payments, and destroy their inventory without reimbursement. Amazon "Community Guidelines" regulate user interaction and "any content (including text, images, video, and links)" posted to the platform (e.g., in the Customer Review section). Amazon users must not conduct themselves or post content that expresses "hatred or intolerance for people on the basis of race, ethnicity, nationality, gender or gender identity, religion, sexual orientation, age or disability, including by promoting organizations with such views." Amazon may sanction users who violate this policy by restricting their use of community features, removing content, delisting products, or suspending or terminating their account.

In 2020, Microsoft CEO Satya Nadella tweeted "There is no place for hate and racism in our society" (cited in Alspach, 2020). Embedded in Microsoft's Services Agreement for all of its platforms (e.g., Microsoft Account, Windows, Office 365 products, Xbox, Azure, Outlook.com and Skype) is a tacit hate speech policy (Microsoft, 2020a). In the Agreement's "code of conduct" section, Microsoft tells users "Don't engage in activity that is harmful to you, the Services or others" and lists "communicating hate speech, or advocating violence against others" as a harmful activity. Microsoft does not define "hate speech" in the Agreement, but on a Microsoft web page entitled "Report Hate Speech Content Posted to a Microsoft Hosted Consumer Service," hate speech is defined as "content that advocates violence or promotes hatred based on: age, disability, gender, national or ethnic origin, race, religion, sexual orientation, gender identity." Microsoft also has a hate speech policy for its advertising clients. In the "Disallowed Content Policies" (Microsoft, 2020b) section, Microsoft defines hate speech as

> any content intended to degrade, intimidate or incite violence or prejudicial action against a group of people based on their race, gender, ethnicity, national origin, religion, sexual orientation, disability or other differentiating factors. (Microsoft, 2020b)

Microsoft says "Advertising that facilitates or promotes hate speech is not allowed, whether directed at an individual or a group." In the "enforcement" section of the Agreement, Microsoft clarifies that if any user violates the Agreement, it may "block delivery of a communication (like email, file sharing or instant message)," "remove or refuse to publish" the user's content, "stop providing Services," and "close" their account.

GAFAM's hate speech policies are enforced by users (who are invited to collaboratively self-police the platforms) and paid content moderators (who are hired to review and make judgments about controversial content). GAFAM downloads the labor of monitoring their platforms for hateful user interactions and content to the users of their services, and expects these users to identify and report this to the content review system. After a potentially hateful user profile, page, channel,

post or piece of content is identified and reported by a user, a paid content moderator will then investigate and judge the content's quality. If they find the content to be in violation of the hate speech policy, they may delete it from the platform. If they determine a user's conduct transgresses the policy, they may issue a warning, suspend their account, or banish them from the platform.

GAFAM's reliance upon their users to enforce hate speech policies may not be the best way to purge hate from Internet platforms because most people use these sites and services for entertainment consumption and socialization, not because they wish to co-police what far-right ideologues do and say. Also, after a hard day of work, many people may prefer to use platforms to watch funny videos, look at pictures or chat with friends than to partake in the unpaid labor of alt-right content surveillance. For those users that do have the time, energy and interest to search for hate content, the subcultural argot of hate groups might cloud their recognition of some hateful expressions. GAFAM's expectation that the billions of users whose data they already monitor, mine and exploit and whose interactions they massage to attract more advertisers will freely partake in the labor of monitoring the platforms for potential violators and violations of their hate speech policy is quite high. And GAFAM's attempt to use their users to service their hate speech policies is *reactive*, and assumes hate content will always already be on the platforms. Perhaps this helps GAFAM save on costs associated with the paid labor of *proactive* content monitoring.

To be fair, GAFAM do employ content moderators, who frequently work on contract and for low wages within a cross-border network of contractor firms that service clients around the world (Chen, 2014; Roberts, 2019). These content moderators review and vet content, often before it appears on user pages. They also make judgments about what content is appropriate or inappropriate, and decide what content will appear on platform pages and what will be hidden from view. Each day, these digital gatekeepers review thousands of pieces of hateful content and their labor of content moderation or curation has major mental health drawbacks, including post-traumatic stress disorder (PTSD) (Newton, 2019). Given GAFAM accumulate billions in profit each year, they should create secure and high paying jobs for content moderators and pay even more to proactively monitor their platforms.

In any case, GAFAM have operationalized their hate content moderation apparatuses to de-platform many alt-right users and delete many pieces of hate content. In 2019, Google removed over 100,000 videos, 17,000 channels and 500 million comments from YouTube that violated its hate speech policy (Yurieff, 2019). In 2020, YouTube terminated over 25,000 channels for doing the same, ridding the site of David Duke and other alt-right influencers (Hern, 2020). Apple Pay retracted service from alt-right websites that relied upon it for processing merchandise payments and donations. Apple iTunes removed white power bands, playlists and tracks, and Apple Podcasts took down alt-right podcasts, including Alex Jones' (Vincent, 2018). In 2018, Facebook removed over 2.5 million pieces of hate content, and took down two pages associated with Richard Spencer: the National Policy Institute and AltRight.com. In early 2020, it deleted nearly 10 million hateful posts and banned 250 white supremacist organizations (Abril, 2020).

Amazon removed some products and listings by white supremacist vendors from its online store, including neo-Nazi t-shirts and books such as Greg Johnson's *The White Nationalist Manifesto*. It also pulled some hate-filled books from Kindle Direct Publishing (KDP). In 2018, Microsoft threatened to stop hosting Gab unless it deleted posts advocating for another Jewish holocaust, and deleted some related hate content.

CONCLUSION

GAFAM are some of the most powerful corporations on the planet, and their platform sites and services interweave and interact with twenty-first century capitalism, State and civic politics and cultures around the world. The US Federal Government's cyber-libertarian media law and regulatory framework prohibits the US State from censoring hate speech in print, via broadcasting and on the Internet, and this has empowered GAFAM to build their own platform content moderation apparatuses. GAFAM have enforced hate speech policies through their content moderation apparatuses and exploited the free labor of users and the waged labor of content moderation workers, who have identified, investigated and de-platformed many alt-right users and deleted a lot of hateful content.

In many ways, GAFAM have become the legislators, judges, juries and executors of their own hate speech policies on the Internet. Long gone are the days when white supremacists could freely use GAFAM platforms in any way they like for whatever hateful ends they please without the risk or threat of some kind of penalty or consequence. GAFAM's efforts have disrupted some alt-right propagandist–entrepreneurs and denied them a platform to speak, be heard and make money. For this reason, GAFAM's de-platforming and deleting of the alt-right is frequently applauded, and yet, GAFAM content moderation apparatuses are no panacea for the social problem of the alt-right, online or offline. The alt-right's presence on platforms persists, and its use of platforms to spread hate is ongoing. As recent as August 2020, millions of QAnon conspiracy theorists were active on Facebook, as were numerous pages promoting the violent "boogaloo bois" (Sonnemaker, 2020). In Spring of 2020, white supremacist books were still being sold on Amazon's KDP: about 200 books recommended by the Colchester Collection, a white nationalist reading space, had been self-published through this platform (Kofman, 2020).

GAFAM's content moderation apparatuses have not stopped the alt-right from continuing to build its movement, online and offline. When "no-platformed" by GAFAM's platforms, the alt-right user might just create a new anonymous profile or account, and then start producing, spreading and consuming hate content again. The labor process of hate content moderation then repeats, with unwaged platform users identifying and reporting potentially hateful content to waged content moderator workers, who then review, judge and act on that content and the user behind it. The dialectic of platforming and de-platforming hate content in GAFAM sites and services is like a never-ending game of "whack-a-mole": smash

one alt-right Facebook profile down, five more pop up. Suspend one neo-Nazi YouTube account, and new creators emerge. Take one Amazon KDB hate book down, and 10 more with slightly different titles and descriptions get published soon after. Even if it was possible to completely purge GAFAM's platforms of all alt-right users, some upstart Internet provider looking to make a quick buck would likely be glad to service their hate.

Apropos the hashtag #SiliconValleySoWhite, GAFAM is intertwined with the structural inequities of racial capitalism (McIlwain, 2019), and so they seem ill-suited to combatting the white supremacist Right. GAFAM are mostly run by white and male CEOs and managers and in 2016, a paltry 3% of the total employees of Silicon Valley's top 75 technology firms were Black (Guynn, 2020). GAFAM's de-platforming and deletion of the alt-right may helpfully reduce the quantity of hate content available on mainstream Internet platforms, but this does not constrain the power of Big Tech's shareholders and CEOs to cash in on the reproduction of racial capitalism. GAFAM are combating digital hate, not because they are committed to some democratic socialist project of emancipatory anti-racist politics, but because they want to appease shareholders and advertisers, augment their brand identities, and drive maximal network effects. In that regard, it is unlikely that GAFAM will disrupt the social order that incubated the alt-right. The global economic slump, neoliberalism's legitimacy crisis, and a reloaded far-Right continue.

Around the world, GAFAM are universalizing a particularly American mode of dealing with digital hate speech. The United States is the world's biggest economic, military and media-technological power, an Empire like no other. GAFAM are the world's leading platform imperialists, and they advance broader US economic, geopolitical and cultural–ideological goals in other countries (Jin, 2019). When GAFAM entered other societies, they networked together billions of people around the world but also opened these people's homes, screens, hearts and minds to the American alt-right. Backed by the US State's cyber-libertarian foreign policy, GAFAM are now trying to curb the cross-border networking of hate groups and constrain the free flow of hate speech. As a consequence, the work of governing GAFAM's platforms, hate groups and hate content on the Internet is being exempted from democratic, public and national legal and regulatory deliberation and decision-making, and efficiently uploaded to oligopolistic, private and trans-national corporations empowered to decide for themselves, and for the rest of us, what the terms and conditions of free speech and hate speech are, and what guidelines we all must follow, for them.

REFERENCES

Abril, D. (2020, May 12). Facebook removes 10 million posts for hate speech as tensions rise on the social network. *Fortune*. Retrieved from https://fortune.com/2020/05/12/facebook-removes-10-million-hate-speech-posts/

Alspach, K. (2020, June 1). Microsoft CEO Satya Nadella: 'There is no place for hate and racism in our society'. *CNN*. Retrieved from https://www.crn.com/news/cloud/microsoft-ceo-satya-nadella-there-is-no-place-for-hate-and-racism-in-our-society-

Amazon. (2018, August 11). Offensive and controversial materials. *Amazon Seller Central*. Retrieved from https://sellercentral.amazon.com/gp/help/external/200164670#:~:text=Amazon%20does%20not%20allow%20products,violence%20or%20victims%20of%20violence

Amazon. (2020). Community guidelines. *Amazon Help and Customer Service*. Retrieved from https://www.amazon.com/gp/help/customer/display.html?nodeId=201929730&language=en_US&ref=efph_home_cont_200164670

Anti-Defamation League (ADL). (2014). Best practices for responding to Cyberhate. *ADL*. Retrieved from https://www.adl.org/best-practices-for-responding-to-cyberhate

Apple. (2020a, September 16). Apple media services: Terms and conditions. *Apple*. Retrieved from https://www.apple.com/legal/internet-services/itunes/us/terms.html

Apple. (2020b, September 11). App Store review guidelines. *App Store*. Retrieved from https://developer.apple.com/app-store/review/guidelines/

BBC (2018, March 20). YouTube's neo-Nazi music problem. *BBC News*. Retrieved from https://www.bbc.com/news/blogs-trending-43416117

Berger, J. M. (2018). The Alt-Right Twitter Census: Defining and describing the audience for Alt-Right content on Twitter. *Vox-Pol*. Retrieved from https://www.voxpol.eu/download/vox-pol_publication/AltRightTwitterCensus.pdf

Bryant, L. V. (2020). The YouTube algorithm and the Alt-Right filter bubble. *Open Information Science*, *4*, 85–90.

Burgess, J., Marwick, A., & Poell, T. (Eds.). (2018). *The Sage handbook of social media*. London: Sage.

Calhoun, L. (2017, July 17). Here's what Jeff Bezos says to do when people criticize you. *Inc*. Retrieved from https://www.inc.com/lisa-calhoun/when-folks-hate-on-you-heres-how-jeff-bezos-says-h.html

Chen, A. (2014, November 23). The laborers who keep Dick pics and beheadings out of your Facebook feed. *Wired*. Retrieved from https://www.wired.com/2014/10/content-moderation/

Clement, J. (2020, August 20). Global social networks ranked by number of users 2020. *Statista*. Retrieved from https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/

Conway, M., Scrivens, R., & Macnair, L. (2019). Right-Wing extremists' persistent online presence: History and contemporary trends. *ICCT Policy Brief*, *1*, 1–24. Retrieved from https://icct.nl/wp-content/uploads/2019/11/Right-Wing-Extremists-Persistent-Online-Presence.pdf

Dahlberg, L. (2019). Cyberlibertarianism. In D. Cloud (Ed.), *The Oxford encyclopedia of communication and critical cultural studies*. Oxford: Oxford University Press. https://www.oxfordreference.com/view/10.1093/acref/9780190459611.001.0001/acref-9780190459611-e-70

Durkee, A. (2020, June 20). Democratic Senators demand Facebook answer for its White Supremacist problem. *Forbes*. Retrieved from https://www.forbes.com/sites/alisondurkee/2020/06/30/senators-demand-facebook-answer-for-its-white-supremacist-problem/#19c3e43228ac

Facebook. (2020). Hate speech. *Facebook Community Standards*. Retrieved from https://www.facebook.com/communitystandards/hate_speech

Federal Communications Commission (FCC). (2020a). About. *FCC*. Retrieved https://www.fcc.gov/about/overview

Federal Communications Commission (FCC). (2020b). FCC and freedom of speech. Retrieved from https://www.fcc.gov/consumers/guides/fcc-and-freedom-speech

Feiner, L. (2018, December 4). Tim Cook's message to hatemongers: 'You have no place on our platforms'. *CNBC*. Retrieved from https://www.cnbc.com/2018/12/04/apple-ceo-tim-cook-says-hate-has-no-place-on-tech-platforms-at-adl.html

Forbes. (2020, September 1). World's Billionaires list. *Forbes*. Retrieved from https://www.forbes.com/billionaires/

Friedman, U. (2019, August 4). How many attacks will it take until the White-Supremacist threat is taken seriously? *The Atlantic*. Retrieved from https://www.theatlantic.com/politics/archive/2019/08/how-can-fbi-fight-far-right-extremism-ideology/595435/

Gillespie, T. (2018). *Custodians of the Internet: Platforms, content moderation and the hidden decisions that shape social media*. New Haven, CT: Yale University Press.

Google. (2014, September 23). Fighting online hate speech. *Google Public Policy Blog*. Retrieved from https://publicpolicy.googleblog.com/2014/09/fighting-online-hate-speech.html

Google. (2020a). YouTube hate speech policy. *Google*. Retrieved from https://support.google.com/youtube/answer/2801939?hl=en

Google. (2020b). Google transparency report. *Google*. Retrieved from https://transparencyreport.google.com/youtube-policy/featured-policies/hate-speech?hl=en

Guerrini, F. (2020, March 3). The problems with Germany's New Social Media Hate Speech Bill. *Forbes*. Retrieved from https://www.forbes.com/sites/federicoguerrini/2020/03/03/the-problems-with-germanys-new-social-media-hate-speech-bill/#210946c9592a

Guynn, J. (2020, February 10). #SiliconValleySoWhite: Black Facebook and Google employees speak out on big tech racism. *USA Today*. Retrieved from https://www.usatoday.com/story/tech/2020/02/10/racial-discrimination-persists-facebook-google-employees-say/4307591002/

Hawley, G. (2017). *Making sense of the Alt-Right*. New York, NY: Columbia University Press.

Hern, A. (2016, May 31). Facebook, YouTube, Twitter and Microsoft sign EU hate speech code. *The Guardian*. Retrieved from https://www.theguardian.com/technology/2016/may/31/facebook-youtube-twitter-microsoft-eu-hate-speech-code

Hern, A. (2020, June 30). YouTube bans David Duke and other US far-right users. *The Guardian*. Retrieved from https://www.theguardian.com/technology/2020/jun/30/youtube-bans-david-duke-and-other-us-far-right-users

Howard, J. W. (2019). Free speech and hate speech. *Annual Review of Political Science*, *22*(1), 93–109.

Jin, D. Y. (2019). Facebook's platform imperialism: The economics and geopolitics of social media. In O. Boyd-Barrett & T. Mirrlees (Eds.), *Media imperialism: Continuity and change* (pp. 187–198). New York, NY: Rowman & Littlefield.

Kofman, A. (2020, April 7). The Hate Store: Amazon's self-publishing arm is a haven for White Supremacists. *ProPublica*. Retrieved from https://www.propublica.org/article/the-hate-store-amazons-self-publishing-arm-is-a-haven-for-white-supremacists

Lewis, R. (2018). Alternative influence: Broadcasting the Reactionary Right on YouTube. *Data & Society Research Institute*. Retrieved from https://datasociety.net/wpcontent/uploads/2018/09/DS_Alternative_Influence.pdf

McChesney, R. (2013). *Digital disconnect: How capitalism is turning the internet against democracy*. New York, NY: The New Press.

McIlwain, C. (2019). *Black software: The internet & racial justice, from the AfroNet to Black Lives matter*. New York, NY: Oxford University Press.

Microsoft. (2020a, August 1). Microsoft Services Agreement. Retrieved from https://www.microsoft.com/en-ca/servicesagreement/

Microsoft. (2020b, September 1). Disallowed content policies. *Microsoft Advertising*. Retrieved from https://about.ads.microsoft.com/en-ca/resources/policies/disallowed-content-policies

Mirrlees, T. (2019). The Alt-Right's platformization of fascism and a New Left's Digital United Front. *Democratic Communiqué*, *28*(2), 28–46.

Morozov, E. (2014). *To save everything, click here: The Folly of technological solutionism*. New York, NY: Public Affairs.

Neiwert, D. (2018). *Alt-America: The rise of the radical right in the Age of Trump*. New York, NY: Verso.

Newton, C. (2019, February 15). The Trauma floor: The secret lives of Facebook moderators in America. *The Verge*. Retrieved from https://www.theverge.com/2019/2/25/18229714/cognizant-facebook-content-moderator-interviews-trauma-working-conditions-arizona

Pickard, V. (2013). Social democracy or corporate libertarianism? Conflicting media policy narratives in the wake of market failure. *Communication Theory*, *23*(4), 336–355.

Polat, G. (2019, April 30). The issues that matter to the Big Tech Lobby. *vpmMentor*. Retrieved from https://www.vpnmentor.com/research/us-lobby-report/vpmMentor

Roberts, S. T. (2019). *Behind the screen: Content moderation in the shadows of social media*. New Haven, CT: Yale University Press.

Schiller D. (2014). *Digital depression: Information technology and economic crisis*. Champaign, IL: University of Illinois Press.

Shirky, C. (2008). *Here comes everybody*. New York, NY: Penguin.

Smyrnaios, N. (2018). Internet Oligopoly: The Corporate Takeover of Our Digital World. Bingley: Emerald Group Publishing.

Sonnemaker, T. (2020, August 15). Facebook had a very unsuccessful week in its fight against misinformation and hate speech. *Business Insider*. Retrieved from https://www.businessinsider.com/facebooks-unsuccessful-week-fighting-misinformation-and-hate-speech-2020-8

Srniceck, N. (2017). *Platform capitalism*. New York, NY: Polity.

Tsesis, A. (2009). Dignity and speech: The regulation of hate speech in a democracy. *Wake Forest Law Review*, *44*, 497–502. Retrieved from https://lawecommons.luc.edu/cgi/viewcontent.cgi?referer=https://www.google.com/&httpsredir=1&article=1040&context=facpubs

Tucker, H. (2020, May 13). World's largest technology companies 2020. *Forbes*. Retrieved from https://www.forbes.com/sites/hanktucker/2020/05/13/worlds-largest-technology-companies-2020-apple-stays-on-top-zoom-and-uber-debut/#b7254a63c030

van Dijck, J., Poell, T., & de Waal, M. (2018). *The platform society: Public values in a connective world*. New York, NY: Oxford University Press.

Vega, N. (2020, June 26). Zuckerberg says Facebook will crack down on hate speech as ad boycott widens. *New York Post*. Retrieved from https://nypost.com/2020/06/26/facebook-twitter-stocks-plummet-after-unilever-joins-ad-boycott/

Vincent, J. (2018, August 6). Apple removes Alex Jones podcasts from iTunes for hate speech. *The Verge*. Retrieved from https://www.theverge.com/2018/8/6/17655168/alex-jones-infowars-apple-itunes-podcasts-removed

Wendling, M. (2018). *Alt Right: From 4chan to the White House*. Halifax: Fernwood.

Wong, Q. (2020, July 30). Facebook ad boycott: Why big brands 'hit pause on hate'. *CNET*. Retrieved from https://www.cnet.com/news/facebook-ad-boycott-how-big-businesses-hit-pause-on-hate/

Yurieff, K. (2019, September 3). YouTube says it's removing more hate speech than before but controversial channels remain up. *CNN*. Retrieved from https://www.cnn.com/2019/09/03/tech/youtube-hate-speech/index.html

CHAPTER 6

PUBLIC ACCUSATION ON THE INTERNET

Sarah Lageson and Kateryna Kaplun

ABSTRACT

Purpose – *In a digital environment, a simple accusation has the potential to permanently attach to a person's identity. Our purpose here is to identify several types of accusations that persist in the internet environment: person to person accusations, media documented accusations, and accusations by the state.*

Approach – *Using a typology of cases and legal analyses, the authors trace how accusations proliferate and persist across the internet and offer a set of social and legal explanations for the salience of public accusation online.*

Findings – *The authors ultimately find that in contemporary society, the act of accusing increasingly replaces the desire or need for a fair and just outcome. The authors close by discussing implications for the accused and potential avenues for remedy.*

Originality – *Our contribution bridges sociological and legal perspectives on the intersection of free speech, defamation, and digital media.*

Keywords: Internet; accusation; due process; defamation; privacy; punishment

INTRODUCTION

In October 2017, widespread buzz began to rise surrounding a seemingly innocuous Google Spreadsheet. Designed initially for private circulation among women

Media and Law: Between Free Speech and Censorship
Sociology of Crime, Law and Deviance, Volume 26, 99–114

ISSN: 1521-6136/doi:10.1108/S1521-613620210000026007

in the media industry, this crowdsourced database aimed to warn other women about sexual misconduct perpetrated by the men in their field. Within hours, however, news outlets were breaking stories about the document, titled "Shitty Media Men." The creator of the list wanted the spreadsheet to work as a "whisper network" where women in the industry could help each other and warn each other about potentially dangerous men (Martin, 2020).

Any person contributing to the list would remain anonymous. A disclaimer read: "This document is only a collection of misconduct allegations and rumors. Take everything with a grain of salt. If you see a man you're friends with, don't freak out" (Martin, 2020). The list allowed contributors to anonymously list men's names, media affiliations, alleged misconduct, and highlighted those names which appeared more than once. The database was built through a network of trust that no one would delete content or share the document with a man. Once the list went public, the original creator attempted to delete the spreadsheet; however, copies remained forever ensconced in the public sphere (Martin, 2020).

On its face, this list seemed to operate in the public good by allowing women who are unconnected, new to the industry, or particularly vulnerable to protect one another in a professional context seemingly full of predatory men. This was particularly beneficial for women of color who are left out of these conversations and who typically have less professional connections that alert them to the harmful side of the media industry (Donegan, 2018). There were also pragmatic considerations at play, which helped explain why public accusation was useful: the American criminal justice system has a long history of poor management of sexual misconduct cases (Morabito, Williams, & Pattavina, 2019). Research has shown that, too often, police are reluctant to investigate, and prosecutors are reluctant to charge someone associated with these types of cases (Morabito et al., 2019). This becomes increasingly true when the accused are powerful, rich white men, and the victims are women, and, in particular, women of color (Ontiveros, 1993).

Shitty Media Men allowed victims to seek justice by crowdsourcing information and releasing an organized set of public accusations. Lacking faith in the justice system and institutionalized media gatekeepers, the public accusations sprang hope that legal actors or employers would take action or begin investigations. There was some initial response: a small handful of terminations and resignations resulted, and some investigations continued in the months after the release of the list (Grady, 2018). However, most people were more concerned with learning the identity of the creator of the list (Grady, 2018).

In its popularity, the list has also sparked debates about the rights of the accused. Men accused of sexual misconduct on this list, as well as some members of the public, pointed to the lack of due process and trial by jury as a race to assume guilt (Grady, 2018). Men on the list felt they were "tried" by the public sphere and deemed guilty without actual mechanisms of justice. Critics noted that the presumption of innocence quickly dissipated as victims struggled to be heard.

The list continues to remain online despite the controversy and emerging litigation. One "media man" on the list filed a defamation suit against the spreadsheet's named creator and 30 other unnamed women who edited the document (Martin, 2020). But the lawsuit has raised even more speech and defamation issues. The

creator has argued that the plaintiff is a public figure through his professional role in the media (which would require a showing of actual malice in knowingly spreading false accusations), but the plaintiff countered that his journalistic work (though centered on issues of sex and sexual assault) is unrelated to the personal privacy issues at hand. The court agreed, and the case is proceeding with this "media man" painted as a private figure who was publicly accused of an act he claims is false.

But the digital context has thrown another wrench into the legal debate over the spreadsheet. The creator of the list has claimed a defense under the guise of Section 230 of the Communications Decency Act. This provision serves as broad immunity to liability for content provided by other users on an internet platform created by the defendant. Likening the "Shitty Media Men" spreadsheet to an online message board, the court has held that the creator of the spreadsheet does qualify as an "interactive computer service provider." The case now turns to whether or not the creator of the spreadsheet also created the specific *entries* that accused the plaintiff of sexual misconduct. This leaves the creator in a tricky bind: for the Section 230 immunities to attach, the creator must disclose the identity of the person who actually leveraged the specific accusation by entering it onto the spreadsheet, which would publicize the identity of the victim onto the internet as well (Martin, 2020). Google has been thrust into the fold to provide this evidence by subpoena, but has refused to do so (Martin, 2020).

While the lawsuit continues to unfold, the emergence of and reactions to the list set up an important set of questions. The list existed outside of any formal grievance structure and did not provide any legal justice to victims. Many of the men on the list faced little, if any consequences. And yet, the popularity of the list and the very nature of its existence seemed to deliver some sense of collective, public justice. Given the significant barriers to real justice, is a public accusation itself now the fastest and most effective way to obtain justice? On the other hand, extralegal accusations come wholly without due process protections and the presumption of innocence. What about the rights of the accused?

SOURCES OF PUBLIC ACCUSATION

In the past, people who wanted to leverage an accusation faced significant obstacles: they could turn to a more powerful gatekeeping institution, such as a newspaper or the criminal legal system. If these institutional arrangements failed to deliver justice, accusers could then rely on informal channels of community-level gossip. Leveraging an accusation and finding justice in all three spaces, however, has long been directly related to the accuser's relative power, social value, and believability of the person who was harmed (Goodey, 2005; Ontiveros, 1993; Walklate, 1989).

In the digital age, a public accusation can be leveraged anywhere and by anyone and rapidly spread through the public sphere. The power to accuse no longer only resides in traditional institutions but can be exercised by less powerful actors as a type of "online self justice" (Loveluck, 2019). We open by typologizing three

forms of public accusation and discuss how each type has evolved in the digital era: accusation by the media, accusation by the state, and accusation by another individual. Although these typologies describe accusations that originate in distinct institutional and social sources, the "context collapse" (Davis & Jurgenson, 2014) of the digital environment means that the implications for the accused are increasingly similar as information rapidly spreads across the internet.

Accusation by the Media

The news media has long been the central institution for publicizing accusation. Yet, in recent decades, entertainment media has increasingly marketed punishment and shaming as part of a broader turn toward using crime as "infotainment," particularly on television (Kohm, 2009) and operating as a form of "gonzo justice" (Altheide, 1992). Blurring the line between fact and fiction, television shows like *Cops, Americas Most Wanted, To Catch a Predator,* and *LivePD* not only reinforce law and order ideologies, exaggerate the extent of crime, and stereotype "criminals" (Cavender, 1998; Jewkes, 2004; Kohm, 2009), but they also contribute to a broader, market-based commodification of crime and punishment (Kohm, 2009).

But while the media was traditionally the key gatekeeper to the public's awareness of a transgressor, the digital landscape has rapidly shifted power balances. While control over content has long been central to the professional logic of media production, media producers today must negotiate the "contested boundary space between producers and users in the digital environment" (Lewis, 2012, p. 836). This "strikes at the heart of a professional model that was built on an implicit bargain between journalists and the public – an assumption about how society should handle the collection, filtering, and distribution of news information" (Lewis, 2012, p. 838). Professional journalism was built upon the assumption that society relies only on them to fulfill the functions of "watch-dog publishing, truth-telling, independence, timeliness, and ethical adherence in the context of news and public affairs" (Lewis, 2012, p. 845)

But today, the "convergence culture" (Jenkins, 2006) of contemporary media means that users feel increasingly more in control of content, making ever more fluid the distinction between media creator and media consumer – and representing a cultural shift in what participatory democracy looks like in a digital world. Similar to the "crowdsourcing" described in the "Shitty Media Men" spreadsheet, digital media in all its forms reflects a shift from the professional media's expert control over content to that of de-professionalized engagement facilitated through open systems of communication. This means that media reports of accusations today come from a broader variety of sources, is increasingly user-generated, and must be responsive to a dynamically engaged audience. The media, in many ways, now reports *on* the internet, rather than simply using the internet as a distribution tool.

Accusation by the State

In the digital age, accusation by the state via arrest or criminal charges is also increasingly publicized on the internet by governmental entities (Lageson, 2020).

As police and courts increasingly expand their digital and data-driven operations, criminal justice agencies disclose online rosters and dockets of people who have been arrested, detained, or charged with a crime. Often, these rosters are accompanied by personal information, including people's full names, birthdates, home addresses, and photographs. In short, technological development has widely expanded the availability of a broad variety of criminal justice system records that have begun to exact punishment-related harms to millions of individuals with legal system contact, well before a conviction (Lageson, 2020).

This public disclosure of pre-conviction records is protected by public records and transparency laws, even though the accused person has not yet been convicted of a crime – and again, the digital environment quickly creates a context collapse similar to that in media. Once information about the accused is disclosed, private data brokers and websites duplicate and re-disseminate pre-conviction records for a myriad of uses, including online extortion, "people search" websites, commercial background check services, and community-level shaming and surveillance (Corda & Lageson, 2020). Even though these records often contain accusations without dispositions, the unprecedented rise of the information age has fundamentally changed the scope and permanence of criminal punishment that extends the impact of formal legal sanctions.

Accusation by Another Individual

Public shaming has always been centered within the community and has remained locally embedded in personal and professional networks developed through face-to-face interactions. But today, the most visible and salient forms of public accusation often originate between two individuals, with the media and justice system trailing behind the internet response. For instance, recent years have seen the meteoric rise of "cancel culture," where traditionally private information shared between individuals, such as direct messages or social media posts, are elevated into the public sphere for shaming purposes. Typically centered on "canceling" financial support for celebrities or other public figures accused of misdeeds, this public disclosure is then used as the basis for large-scale "cultural boycotts" of the person in a range of public domains (Bromwich, 2018).

Online shaming can also be leveraged against previously unknown people, bringing them into immediate infamy (Ronson, 2016). While a public accusation is leveraged at first from one individual to another, the internet allows the accusation to blossom into a charge leveraged by millions of digital supporters who can reshare and retweet a transgression. For instance, in 2013, Justine Sacco became infamous over the course of hours after posting a racist Tweet while boarding a flight to South Africa. Once she landed, her reputation and career had been upended (Gallardo, 2017).

There is a broad spectrum of uses and effects of public accusations leveraged between people on the internet. Public accusations have brought light and justice to longstanding abuses of power, as evidenced in the #metoo movement and in efforts to hold police accountable. Public accusations have also been leveraged to create further victimization, such as through "revenge porn" and other types of

non-consensual distribution of intimate content. But perhaps most importantly, the rapid diffusion of public accusation signals a shift of power and agency in who can now leverage an accusation with traction against a much more powerful figure or quickly grow an internet mob against a heretofore unknown transgressor. Today, "'ordinary people' armed with digital tools and publishing platforms exercise a new form of surveillance in digital culture" (Hess & Waller, 2013, p. 102). While there are benefits to this grassroots and crowdsourced method of public accusation in bringing to light injustices that would in previous times remained secret, the rapid rise and reach of such accusations raises privacy and due process concerns.

THE INTERNET AND ETERNAL ACCUSATION

Once leveraged, a public accusation has nearly endless digital spaces to inhabit. While this communicative power has the ability to upend longstanding hierarchies and inequalities, there is also enormous potential for harms. As Lawrence Lessig warned in 1999, "cyberspace does not guarantee its own freedom but instead carries an extraordinary potential for control" (Lessig, 1999, p. 58). While public accusations leveraged on the internet have potential to bring injustices to light and hold transgressors accountable to their actions, the eternal shelf life of the accusation itself raises questions about due process, forgiveness, and rehabilitation.

The architecture of internet search engines ensures that a public accusation will remain firmly lodged to a persons' name for perpetuity. Search engines, like Google, continuously "crawl" public webpages and incorporate text into search results. When a link in a search result reveals information with "shock value," people are more likely to click – and links with the most clicks maintain a dominant place in internet search results. Thus, the most salacious and most viewed information about a person will dominate their Google search identity, such as a mugshot or social media commentary about a person accused of wrongdoing (Pierce, 2013).

There has also emerged a strong market of websites designed to aggregate and collate a mix of public records, social media posts, and website content into "people search" sites that essentially function as unregulated and informal background check websites. For instance, the website "MyLife" gathers personal information on the internet to create a "MyLife Public Page" for any individual. MyLife claims to operate as a website focused on the safety of individuals by alerting them to search results and records that are associated with their names (Wash & Benitez, 2019). However, court filings have alleged that MyLife actually operates as an extortion scam where people are forced to pay money to access the records about them maintained by the company (Cohen, 2020). Because of the raw aggregation of multiple data sources, the lawsuit also alleges that oftentimes these records are also inaccurate. As a result, MyLife is facing two class action lawsuits alleging extortion after the company sent mass email solicitations notifying website subjects that "someone" is searching for them, then charging a

subscription fee to reveal a list of allegedly bogus names (Cohen, 2020; Wash & Benitez, 2019).

Extortion victims also allege that MyLife's use of a "reputation score," which is supposed to help people determine whether people and their businesses are reputable is actually *damaging* their reputations by spreading false information (Cohen, 2020; Wash & Benitez, 2019). This "reputation score" furthers the scam by instantaneously rising in value when an individual pays to see more information about it (Cohen, 2020). MyLife claims a proprietary algorithm calculates the score and refuses to be transparent about how the score is calculated and does not allow website subjects to submit new or accurate information to the site in an effort to increase their reputation score (Cohen, 2020).

Search engine optimization and the commodification of public accusations into profit-generating content signal how even the smallest of online accusations can be permanently cemented into the internet canon. However, the digital permanence of these accusations is sustained and perpetuated by maintaining an audience willing to click, share, and re-publicize an accusation. This begs the questions of the social and psychological functions of public accusation, which we explore next.

SOCIOLEGAL AND PSYCHOLOGICAL PERSPECTIVES

Sociology of Punishment

As the cases described above imply, accusation itself can pragmatically constitute punishment as it is solidified in the public sphere. Here, we explore sociological and psychological theories that explain the salience, functions, and attractions of public punishment in society. First, public accusation might be framed as a form of public punishment that serves social solidarity functions. Second, we explore the psychological roots of "clickbait" to analyze the persistence of public accusation long after the triggering incident.

The social–functional components of punishment are central to classic social theory. Émile Durkheim (1893) argued that punishment held a moral function to solidify and express shared values through collective responses to crime. Following this tradition, David Garland (1991) observed that

> the rituals of punishment are directed less at the individual offender than at the audience of impassioned onlookers whose cherished values and security had been momentarily undermined by the offender's actions. (p. 123)

In this sense, it is *the accusation itself* that is central to solidifying the social function of punishment, the so-called "declaration" of punishment rather than the actual "delivery" of punishment (Garland, 1991, p. 125).

The visual and emotional salience of participatory internet also deepens the salience of contemporary accusations, wherein "the power of the image moves that passion which Durkheim called the very soul of punishment" (Valier, 2004, p. 252). And while digital accusation is the most recent, and perhaps most public, example of this trend, punishment scholars have broadly noted the rise of

"expressive" punishments since the late 1970s, including the renewed use of the death penalty in the United States, increasingly punitive sentencing regimes, and "shame penalties," such as the infamous "sandwich board" cases, where judges require convicted thieves to wear public proclamations of their crimes. These explicit emotional appeals are central to Garland's conception of the "culture of control," where contemporary law and order ideologies have been characterized by a focus on victims, crime as a political issue, and the recasting and othering of people convicted of juvenile or sex offense crimes (Garland, 2001). These cathartic policies constitute a form of politically motivated "acting out," rather than empirically driven efforts to reduce crime, wherein policy rationales sound more like "sound-bite" statements (Garland, 2001, pp. 132–133). These sociological perspectives, in many ways, help explain the social acceptance, popularity, and salience of public accusation.

Publicity, in some ways, has also begun to constitute the punishment itself – replacing traditional forms of legal sanctions. The rise in sex offender registries and mugshot galleries, for instance, might be conceptualized as "gonzo" justice: the "use of extraordinary means to demonstrate social control and moral compliance through rule enforcement and punishment designed to stigmatize publicly" to "demonstrate the moral resolve of those mandating the punishment" (Altheide, 1992, pp. 69 & 71). While Altheide, writing in the early 1990s, focuses on state sanctioned punishments (such as a judge ordering a woman convicted of child abuse to remain on birth control pills for her lifetime), the elements of such punishments he describes can be logically applied to public accusation on the internet. For instance, Altheide (1992) spells out elements of gonzo justice as including expressive and evocative responses that are "intended to be interpreted and presented as an exemplar for others to follow," but seldom includes contrary or challenging statements: the corollary of "taking the law into one's own hands" (p. 76).

Clickbait: Uses and Gratifications Theory

The social value of accusations on the internet rests within the desire of users to consume, share, and discuss the transgression. The psychological theory of "uses and gratifications" helps explain why people utilize mass communication for social interaction (Katz, Blumler, & Gurevitch, 1974). Audiences today are active, rather than passive, and people utilize mass communication for socialization purposes (Katz et al., 1974). While a broad set of social-cognitive factors explain how people use the internet to create social interactions that give immediate positive reinforcements, psychological theories point toward the role of the internet in heightening the effects of uses and gratifications into the digital context that centers on instant gratification (LaRose, Mastro, & Eastin, 2001).

Against the backdrop of increasing disillusionment with the criminal legal system, grassroots public accusations may offer a sense of justice in a manner much faster – and less corrupted – than traditional legal remedies. For instance, the publicization of "Shitty Media Men" provided instant gratification for accusers far before legal or media systems could mobilize. Further gratification comes from control over the content itself as information is now free from the evidentiary

requirements of the court system and editorial constraints of traditional media. Accountability of the perpetrators and relief on the behalf of victims deepens the salience and justification for making the accusation in the first place. There becomes justice when no other form of justice might exist.

On the user end, public accusations on the internet also constitute a type of "clickbait." Clickbait, the all-too-familiar sensational headlines on the internet that are designed to be too irresistible for an internet user to ignore, invokes an instant psychological response, ranging from mild curiosity to the "fear of missing out." Psychologists argue that the success of clickbait lies in deeper and darker realities of the modern consumer. Prior to the internet, people were generally more comfortable acknowledging that there was just plenty of information out there that they didn't have to access to. The internet has changed the situation dramatically by allowing us direct access to a volume of information so large that we will never be able to consume all of it. This knowledge of the volume of unknowable information provokes an emotional response (Lowenstein, 1994), and creates an anxiety that bite-sized clickbait helps to remedy so the world becomes more understandable, digestible, and ultimately controllable (Yarrow, 2014).

Procedural Justice

A final explanation for the rising salience of public accusation on the internet is through a procedural justice lens (Tyler, 2003). As broad public faith in the court system wanes, people are turning to taking the law into their own hands. There are many examples of how this legal self-help process emerges from broader social constraints or concerns; take for instance the Minutemen who patrol the American border as a means to cope with anxieties toward people of other racial and ethnic groups (Cabrera & Glavac, 2010), or men who carry guns in response to a perceived decline in American values (Carlson, 2015). And even though we live in an information-rich environment, many parts of the criminal legal system remain largely obfuscated, such as through the "vanishing" of jury trials (Galanter, 2004) and the "black box" around prosecutorial discretion (Vera Institute of Justice, n.d.).

Digital technologies have also been mobilized to assist long-overlooked issues with victims rights. For instance, VINELink, a nation-wide website that receives metadata from thousands of correctional facilities for most states of the United States, stands for "Victim Information and Notification Everyday" and alerts victims if the perpetrator of the crime committed against them has been released from custody (Appriss Insights, 2020). The "Citizen" crime watch app, Amazon Ring Doorbell systems, and community social media sites like NextDoor all contribute to a sense of taking control over a legal system that has failed to effectively deliver justice.

LEGAL LIMITATIONS FOR THE DIGITALLY ACCUSED

Our typology of public, digital accusations by media, state, and individual offers separate potential legal remedies and legal limitations. For media and

individual claims, cyber-defamation law provides some answers. However, cyber-defamation claims are muddied by internet law (particularly Section 230 of the Communications Decency Act), and further complicated by jurisdictional problems for internet-based claims. Accusation by the state is governed by a different set of public records laws, leaving the accused with even less legal recourse.

Cyber-defamation Law and Its Limits

Defamation is framed through the lens of reputational damage. Traditional defamation consists of the torts of libel (written or recorded words) and slander (oral communications and speech) that involve *false* allegations that harm a person's reputation (Restatement of Torts 2d § 559). But amidst an incredibly uneven legal and theoretical landscape, defamation has been limited by free speech concerns (Solove, 2007), exemplified in the 1964 US Supreme Court case *New York Times Co*. v. *Sullivan*. Recalling the legal determinations in the "Shitty Media Men" case, this case hinged on proving "actual malice" when the allegedly defamed person is a public figure. Here, the plaintiff was the police commissioner who was attacked in a *New York Times* advertisement for his mistreatment of protesters during civil rights protests in Montgomery, Alabama. The Supreme Court determined that although there were factual errors or defamatory statements published, these were insufficient to overcome protection of the First Amendment because no "actual malice" was proven (*New York Times Co*. v. *Sullivan*, 1964). The ruling encouraged spirited public debate around public and governmental figures without fear of extensive litigation. For the private actor, there is no burden to prove that the person who made defamatory statements acted with malice in spreading false information.

This defamation standard is made more complex by the digital context. In the information age, a person can become a public figure in an instant, blurring the line between legal categories of public officials, public figures, and "limited-purpose" public figures (those whose notoriety is linked directly to the issue at hand). Importantly, any defamation cause of action must also involve the distribution of demonstrably false information that causes a person harm. Accusations are often leveraged well before any sort of systemized or independent fact-finding process.

A second issue relates to personal jurisdiction for content posted to an online publication. This idea presents new challenges because the plaintiff must show that people in the state where the publication was written had actually viewed it and that the content caused reputational damage (see, for instance, *Scottsdale Capital Advisors Corp*. v. *The Deal, LLC*, 2018). A final issue relates to the "single publication rule" as applied to online speech, which provides that a publisher can only be within the statute of limitations for the *original* content and not subsequent disclosures. This is obviously complicated by the Internet as revisions, modifications, and retractions can occur in an instant (see *Petro-Lubricant Testing Labs & Inc*. v. *Adelman*, 2017).

Thus, while defamation law has always had limits, the digital context has raised many new questions about the line between public and private figures, the jurisdiction of cyberspace, and the problems of online accusations having an eternal

shelf life as they are re-discovered and re-shared across the internet. In response, cyber-defamation has developed as a distinct common law tort that focuses on the communication of false statements in cyberspace (Medenica, 2008), significantly widening where and when defamation can occur. However, unlike traditional defamation cases that focus on First Amendment elements, cyber-defamation is largely governed and limited by the Communications Decency Act (CDA) of 1996.

Specifically, Section 230 of the CDA states that "no provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider" (47 U.S.C. § 230). This provision provides immunity for internet service providers and requires courts to determine whether or not a website is a publisher capable of speech and, by extension, responsible for harmful speech on its platforms. Section 230 also makes defamation lawsuits that have an online element almost unwinnable cases for the individuals who are defamed, constantly putting cyber-defamation cases in a battle with a third-party "intermediary" (Medenica, 2008). Traditional defamation occurs between two parties: the party being defamed and the party accused of writing the defamation (Medenica, 2008). Cyber-defamation results in a tricky situation where a third party owns and operates the platform where the public defamation occurred, creating headaches for determining who is ultimately responsible for content (Medenica, 2008). This also raises an interesting defense for victims seeking justice through public accusations: in the Shitty Media Men case, the spreadsheet creator has claimed a CDA230 defense. Unfortunately, by painting herself as an internet platform, the case will now focus on revealing the identity of the anonymous women who updated the spreadsheet. Bringing in yet another player, Google will have to decide whether or not to release that information, adding layers of discretion and complexity to an already complicated case.

Public Records Law

Accusations officially leveraged by the state are much more difficult for the accused person to challenge. Accusations by the state are released through public records laws, which then legitimize their re-dissemination across private websites, platforms, and search engine results. In the paper age, records were shielded by "practical obscurity," requiring a bit of patience and skills to physically obtain copies of public records. Digitization rapidly changed the shape and reach of records: no longer a physical item of a person is required to actively seek out, today's public records can be inadvertently discovered through a Google search for a person's name.

After a government agency makes a record public, subsequent disclosure of criminal information on the internet is legal. The First Amendment allows for the reproduction and further publication of these public records (see, for instance, *Cox Broadcasting Corp* v. *Cohn*, 1975). Tort law cannot impose damages for truthfully publishing this information (see, for instance, *Florida Star* v. *B. J.F.* 1989), or for failing to remove outdated information, even when an arrest record has been expunged (*Martin* v. *Hearst Corporation,* 2d. Cir. 2015).

The internet has made information transmitted via media, the state, or other internet users essentially synonymous. The internet archive means that, regardless of source, an accusation can live on forever. This includes false accusations by other people, the publicization of a person through the media, or even illegal or wrongful arrests by police. A final legal consideration exists within the shadowy contours of American privacy law.

Privacy Law and the Right to Be Forgotten

Privacy torts may be leveraged if a person publicizes private information about another individual that is highly offensive to a reasonable person and is *not* of legitimate concern to the public (Restatement of Torts 2d § 652D). In contrast to defamation law that creates liability for spreading falsehoods, this privacy tort addresses the dissemination of truths (Solove, 2007, p. 119). Again, however, courts are reluctant to recognize a privacy violation, particularly under a First Amendment analysis.

A related limitation to privacy suits is that the accused person must re-enter the public record to file a claim. Commonly referred to colloquially as the "Streisand effect," people experience real barriers to making privacy claims through the risk of re-publicizing the initial disclosure. The label is derived from an incident where the singer attempted to limit online photos of her Malibu home from being distributed by media companies, thereby drawing more public attention to it. When a person makes a fuss in attempting to remove online content, they inadvertently drive traffic to it (Feinberg, 2015). The result may be a chilling effect for those already on the fence about invoking their already tenuous claim to a privacy right.

Finally, the European Union's "Right to Be Forgotten," offers an interesting legal counterpoint to the problems of public accusation, reputation, and free speech posed in the United States. In a 2014 ruling, the European Court of Justice ruled that EU citizens have the right to make requests that search engines de-link particular search results by request of the search subject. Requesters must show that the information is inaccurate, inadequate, irrelevant, or excessive for the purposes of data processing. Each removal request is assessed on a case-by-case basis by a team of Google employees to ensure that the loss of the public's access to the information is balanced fairly against the requester's right to privacy.

But American companies have publicly disagreed with the European court's decision. Google challenged the ruling in a Guardian op-ed and argued that forcing the search engine to remove links

> means that the Guardian could have an article on its website about an individual that's perfectly legal, but we might not legally be able to show links to it in our results when you search for that person's name. It is a bit like saying the book can stay in the library but cannot be included in the library's card catalogue. (Drummond, 2014)

The New York Times editorialized that

> the European position is deeply troubling because it could lead to censorship by public officials who want to whitewash the past. It also sets a terrible example for officials in other countries who might also want to demand that Internet companies remove links they don't like. (The Editorial Board, 2015)

Still, these public pronouncements don't reflect American sentiments: a 2015 poll found nearly nine in 10 Americans support Right to Be Forgotten legislation in the United States (Trujillo, 2015).

However, the prevailing view of big tech and American law posits that once information is online, it can legally remain online, even though such preservation and permanence represent an unrealistic view of how human memory works. Viktor Mayer-Schönberger (2011) argues that digital documentation negates time, and argues that "through perfect memory, we may lose a fundamental human capacity – to live and act firmly in the present" (p. 12). Here, unshakable digital memory "prevents society from moving beyond the past because it cannot forget the past" (Jones, 2018, p. 20).

IMPLICATIONS FOR ACCUSERS AND THE ACCUSED

A single public accusation made today can supersede all other layers of a person's public identity. Particularly in an American legal context, the internet never forgets. Daniel Solove (2007) calls this the "dark side" of the agency and accountability goals of public accusations, where "information that was once scattered, forgettable, and localized is becoming permanent and searchable" (p. 4). In Solove's view, "the free flow of information threatens to undermine our future" (p. 4).

While there may be social solidarity impacts, positive norm enforcement is also potentially upended when public accusations result in permanent punishment or ostracization. If a public accusation remains firmly lodged in place as a person's master public identity, there is enormous potential for increased normlessness. Undercutting Durkheim's notions of social solidarity through shared and public punishments, today's public accusations may operate as a *de facto* sorting mechanism to exclude people out of society. The circular consequence may be less public engagement, less free speech, and less creative endeavors as people begin to withdraw from society even before a public accusation; the result of the "chilling effects" of a society increasingly comprised of community surveillance efforts.

Evidence shows that this personal and social stigmatization leads people to purposefully "opt out" of formal institutional arrangements and relationships that might trigger a Google search, also referred to as "institutional avoidance." This avoidance has professional, economic, personal, and familial consequences (Brayne, 2014; Goffman, 2014), and has been linked to decreases in civic and political engagement (Lerman & Vesla, 2014) and volunteering (Uggen & Janikula, 1999). Decades of research on labeling theory have shown individual and social harms through the wholesale exclusion of people after a widely acknowledged accusation. In the eyes of society, the person with a stigma is not quite human (Goffman, 1963). While the desire for justice may underlie the rationale for taking the law into one's own hands, we know too little about the long-term consequences.

CONCLUSION

There are broad implications beyond the individual who is accused when accusations eclipse due process, forgiveness, and forgetting. As the traditional pillars of formal accusations become secondary to crowdsourced accusations, society increasingly focuses on the front end of the justice-seeking process, losing valuable sight of constitutional protections and diminishing the accountability needed to ensure truth and justice prevail at the end.

The case of the "Shitty Media Men" provides just one vivid example of a phenomenon that unfolds on a minute by minute basis across the internet. As the case illustrates, an earnest effort to seek justice outside the legal system or mainstream media has been met with even more litigation, threats to anonymity, and the shifting of the spotlight to the list's creator (rather than the abusive men the list was meant to reveal). As for now, the original spreadsheet contributors remain anonymous, but the pendulum may swing in the other direction if people begin to publicly shame the accusers.

The legal system was designed to hold people accountable, to seek justice, and to restore victims. The media has historically functioned as the fourth estate, watchdogging those in power and bringing justice to the masses. Within small communities, reputations and accusations made local ripples. Today, however, accusations proliferate and persist across the internet, heightening the punishment of the accused and potentially re-victimizing the person who was originally harmed and turned to the internet to find justice. In cases of wrongful or false accusations, permanent public stigma is unlikely to outweigh any chance of redemption. Here, we offer a set of social and legal explanations for the salience of public accusation on the internet, highlighting how existing legal regimes limit the remedies available to those who are caught in the web of internet accusations. We ultimately find that in contemporary society, the act of accusing increasingly replaces the desire or need for a fair and just outcome. However, the long-term consequences of public accusation are perhaps still yet to come.

REFERENCES

Altheide, D. (1992). 'Gonzo justice.' *Symbolic Interaction*, *15*(1), 69–86.

Appriss Insights. (2020). What is Vine? Retrieved from https://apprissinsights.com/solutions/vine/. Accessed on September 18, 2020.

Brayne, S. (2014). Surveillance and system avoidance: Criminal justice contact and institutional attachment. *American Sociological Review*, *79*(3), 367–391.

Bromwich, J. E. (2018, June 28). Everyone is canceled. *New York Times*. Retrieved from https://www.nytimes.com/2018/06/28/style/is-it-canceled.html. Accessed on September 18, 2020.

Cabrera, L., & Glavac, S. (2010). Minutemen and desert Samaritans: Mapping the attitudes of activists on the United States' immigration front lines. *Journal of Ethnic and Migration Studies*, *36*(4), 673–695.

Carlson, J. (2015). Mourning Mayberry: Guns, masculinity, and socioeconomic decline. *Gender & Society*, *29*(3), 386–409.

Cavender, G. (1998). In "The shadow of shadows": Television reality crime programming. In G. Cavender & M. Fishman (Eds.), *Entertaining crime: Television reality programs* (pp. 79–94). New York, NY: Transaction Publishers.

Cohen, S. (2020, May 14). MyLife class action alleges site displays false background reports. *Top Class Actions*. Retrieved from https://topclassactions.com/lawsuit-settlements/lawsuit-news/privacy/mylife-class-action-alleges-site-displays-false-background-reports/. Accessed on September 9, 2020.

Corda, A., & Lageson, S. E. (2020). Disordered punishment: Workaround technologies of criminal records disclosure and the rise of a new penal entrepreneurialism. *British Journal of Criminology*, *60*(2), 245–264.

Cox Broadcasting Corp v. *Cohn*, 420 U.S. 469 (1975).

Davis, J. L., & Jurgenson, N. (2014). Context collapse: Theorizing context collusions and collisions. *Information, Communication & Society*, *17*(4), 476–485.

Donegan, M. (2018, January 10). "I started the media men list: My name is Moira Donegan." *The Cut*. Retrieved from https://www.thecut.com/2018/01/moira-donegan-i-started-the-media-men-list.html. Accessed on September 9, 2020.

Drummond, D. (2014, July 10). We need to talk about the right to be forgotten. *The Guardian*. Retrieved from http://www.theguardian.com/commentisfree/2014/jul/10/right-to-be-forgotten-european-ruling-google-debate. Accessed on September 18, 2020.

Durkheim, E. (1893). *The division of labor in society*. New York, NY: Free Press.

Feinberg, A. (2015, January 7). The Streisand effect: Celebrating 10 years of Internet pile-ons. *Gizmodo*. Retrieved from https://gizmodo.com/the-streisand-effect-celebrating-10-years-of-internet-1677579192. Accessed June 9, 2019.

Florida Star v. *B. J.F.*, 491 U.S. 524 (1989).

Galanter, M. (2004). The vanishing trial: An examination of trials and related matters in federal and state courts. *Journal of Empirical Legal Studies*, *1*(3), 459–570.

Gallardo, K. (2017). Taming the internet pitchfork mob: Online public shaming, the viral media age, and the Communications Decency Act. *Vanderbilt Journal of Entertainment & Technology Law*, *19*(3), 721–746.

Garland, D. (1991). Sociological perspectives on punishment. *Crime and Justice*, *14*, 115–165.

Garland, D. (2001). *The culture of control*. Chicago, IL: The University of Chicago Press.

Goffman, A. (2014). *On the run: Fugitive life in an American city*. Chicago, IL: The University of Chicago Press.

Goffman, E. (1963). *Stigma: Notes on the management of spoiled identity*. New York, NY: Touchstone.

Goodey, J. (2005). *Victims and victimology: Research, policy and practice*. Harlow: Pearson Education.

Grady, C. (2018, January 11). The "shitty media men" list, explained. *Vox*. Retrieved from https://www.vox.com/culture/2018/1/11/16877966/shitty-media-men-list-explained. Accessed on September 9, 2020.

Hess, K., & Waller, L. (2013). The digital pillory: Media shaming of 'ordinary' people for minor crimes. *Continuum: Journal of Media & Cultural Studies*, *28*(1), 101–111.

Jewkes, Y. (2004). *Media and crime*. Thousand Oaks, CA: Sage Publications.

Jenkins, H. (2006). *Convergence Culture: Where Old and New Media Collide*. New York, NY: New York University Press.

Jones, M. L. (2018). *Ctrl + Z: The right to be forgotten*. New York, NY: New York University Press.

Katz, E., Blumler, J. G., & Gurevitch, M. (1974). Uses and gratifications research. *The Public Opinion Quarterly*, *37*(4), 509–523.

Kohm, S. A. (2009). Naming, shaming, and criminal justice: Mass-mediated humiliation as entertainment and punishment. *Crime, Media, Culture*, *5*(2), 188–205.

Lageson, S. E. (2020). *Digital punishment: Privacy, stigma, and the harms of data-driven criminal justice*. Oxford: Oxford University Press.

LaRose, R., Mastro, D., & Eastin, M. S. (2001). Understanding Internet usage: A social-cognitive approach to uses and gratifications. *Social Science Computer Review*, *19*(4), 395–413.

Lerman, A. E. & Vesla, M. W. (2014). *Arresting citizenship: The democratic consequences of American crime control*. Chicago, IL: University of Chicago Press.

Lessig, L. (1999). *Code and other laws of cyberspace*. New York, NY: Basic Books.

Lewis, S. C. (2012). The tension between professional control and open participation: Journalism and its boundaries. *Information, Communication & Society*, *15*(6), 836–866.

Loewenstein, G. (1994). The psychology of curiosity: A review and reinterpretation. *Psychological Bulletin*, *116*(1), 75.

Loveluck, B. (2019). The many shades of digital vigilantism. A typology of online self-justice. *Global Crime*, *21*(3–4), 213–241.

Martin v. *Hearst Corporation*, 777 F.3d 546 (2d. Cir. 2015).

Martin, B. (2020, July 1). Defamation suit against the 'Shitty Media Men' list creator can move forward, judge rules. *LA Magazine*. Retrieved from https://www.lamag.com/citythinkblog/shitty-media-men-list/. Accessed on September 9, 2020.

Mayer-Schönberger, V. (2011). *Delete: The virtue of forgetting in the digital age*. Princeton, NJ: Princeton University Press.

Medenica, O. (2008). The immutable tort of cyber-defamation. *Journal of Internet Law*, *11*(7), 3–10.

Morabito, M. S., Williams, L. M., & Pattavina, A. (2019). *Decision making in sexual assault cases: Replication research on sexual violence case attrition in the U.S.: Final Technical Report #252689*. Washington, DC: U.S. Department of Justice.

New York Times Co. v. *Sullivan*, 376 U.S. 254 (1964).

Ontiveros, M. L. (1993). Three perspectives on workplace harassment of women of color. *Golden Gate University Law Review*, *23*(3), 817–828.

Petro-Lubricant Testing Labs, Inc. v. *Adelman* 160 A.3d 712 (2017).

Pierce, D. (2013, October 7). The SEO behind mugshot websites. *Cogney*. Retrieved from https://www.cogney.com.hk/blog/mugshot-seo/. Accessed on September 18, 2020.

Ronson, J. (2016). *So you've been publicly shamed*. New York, NY: Riverhead Books.

Scottsdale Capital Advisors Corp. v. *The Deal, LLC 887* F.3d 17 (1st Cir. 2018).

The Editorial Board. (2015, February 4). Europe's expanding 'Right to be Forgotten.' *The New York Times*. Retrieved from https://www.nytimes.com/2015/02/04/opinion/europes-expanding-right-to-be-forgotten.html. Accessed on September 18, 2020.

Trujillo, M. (2015, March 19). Public wants 'Right to be Forgotten' online. *The Hill*. Retrieved from https://thehill.com/policy/technology/236246-poll-public-wants-right-to-be-forgotten-online. Accessed on September 18, 2020.

Tyler, T. R. (2003). Procedural justice, legitimacy, and the effective rule of law. *Crime and justice*, *30*, 283–357.

Uggen, C., & Janikula, J. (1999). Volunteerism and arrest in the transition to adulthood. *Social Forces*, *78*(1), 331–362.

Valier, C. (2004). Introduction: The power to punish and the power of the image. *Punishment & Society*, *6*(3), 251–254.

Vera Institute of Justice. (n.d.). *Unlocking the Black Box of prosecution*. Retrieved from https://www.vera.org/unlocking-the-black-box-of-prosecution. Accessed on September 18, 2020.

Walklate, S. (1989). *Victimology: The victim and the criminal justice system*. London: Unwin Hyman.

Wash, S., & Benitez, G. (2019, November 1). A website gathering info on 'almost everyone in America' says a poor reputation score could hurt you. *Good Morning America*. Retrieved from https://www.goodmorningamerica.com/living/story/website-gathering-info-america-poor-reputation-score-hurt-66666178. Accessed on September 9, 2020.

Yarrow, K. (2014). *Decoding the New Consumer Mind: How and Why We Shop and Buy*. San Francisco: Jossey-Bass.

PART III

REGULATING SPEECH ACROSS NATIONS

CHAPTER 7

FREEDOM OF EXPRESSION AND HUMOR IN CANADA: THE CASE OF *JÉRÉMY GABRIEL* v *MIKE WARD*

Anne-Marie Gingras

ABSTRACT

Purpose: *This chapter examines how two basic rights, freedom of expression, and the right to equality based on one's dignity, reputation, and honor, were balanced in a case involving a stand-up comedian and an adolescent suffering from Treacher Collins syndrome.*

Methodology/Approach: *The case is contrasted with Jürgen Habermas' concept of the public sphere and with the intrinsic and utilitarian values that Canadian courts have attributed to free speech.*

Findings: *Because the case was dealt with first in a human rights tribunal and then by a court of appeal, a number of considerations were overlooked in court proceedings: how laughter occurs; the broadening of Ward's audience and its consequences; and Ward's publicity strategy. These aspects are explored here to give a more complete picture of the case beyond the court decisions.*

Originality/Value: *In Canada, freedom of expression is usually dealt with ordinary courts. A whole new avenue for dealing with this right is human rights bodies and tribunals. Contesting free speech in the name of defamation is being replaced by rights entrenched in human rights charters, such as the right to equality based on the preservation of one's dignity, reputation, and honor.*

Media and Law: Between Free Speech and Censorship
Sociology of Crime, Law and Deviance, Volume 26, 117–131

ISSN: 1521-6136/doi:10.1108/S1521-613620210000026008

Keywords: Freedom of expression; right to preserve dignity; reputation and honor; Canada; humor; disabled adolescent.

INTRODUCTION

Is discrimination a valid impediment to free speech? This question summarizes the litigation between Jérémy Gabriel, at the time a teenager suffering from Treacher Collins syndrome, which causes some head deformities, and Mike Ward, a popular stand-up comedian who performs in French and English in Canada (*CDPDJ* v. *Ward*, 2016, para. 8). As a child, Gabriel had begun a singing career and had had some success, having performed for the pope and having visited Céline Dion, among others. In a skit called "Les Intouchables," played in bars and show venues from 2009 to 2013, Ward joked about Gabriel's physical appearance in relation to his disability and the means for remedying it. He also joked about Gabriel's mother, making her out to be a "showbiz mom" eager to profit from her son's career. Ward's jokes sparked mockery at Gabriel, especially at high school and on social media, where he became an object of bullying. He became socially withdrawn, lost confidence, and suffered from depression. His grades dropped, he had to leave school, and he had suicidal thoughts.

The dispute opposed two rights protected by the Quebec Charter of Human Rights and Freedoms: the right of Gabriel and his parents to safeguard their dignity, honor, and reputation, and Ward's right to free speech. The case was first handled by the Commission of Human Rights and Youth Rights of Quebec (Commission des droits de la personne et des droits de la Jeunesse or CDPDJ), with which a complaint was filed in 2012, and was then referred to the Human Rights Tribunal. Evidence of discrimination against Gabriel was found in 2016; Ward appealed to the Court of Appeal of Quebec, losing in 2019.

Freedom of expression is one of the most fundamental pillars of democracy: the right to express oneself creates a lively public sphere – an environment in which a wide variety of ideas can be discussed, including unconventional or loathsome ones. However, free speech is not absolute; exceptions exist in matters of national security, protection of privacy, hate propaganda, and so on. Violent expression is not included in protected expression (*Irwin Toy* v. *Quebec (Attorney General)*, 1989; Zhou, 2019). In Canada, other rights included in the Charter also limit free speech, and the courts are called upon to strike a balance between this right and a range of other rights, such as the right to equality. Humor is included in speech protected by freedom of expression, as it sometimes paves the way for new ideas and for debates about political or economic power.

In the first part of this chapter, I explain my approach for dealing with free speech and humor. The second part is devoted to the case at hand, with details about how free speech is dealt with by human rights bodies in Canada. The third part focuses on three aspects ignored in the courts that I believe essential to grasping the social repercussions of the issue: how laughter occurs, the broadening of Ward's audience in the digital era, and Ward's publicity strategy.

FREEDOM OF EXPRESSION: PRINCIPLE AND LIMITS

Strengths and Weaknesses of the Public Sphere

Freedom of expression is intimately linked to the exercise of democracy, and it is interpreted mainly from a liberal perspective that privileges public discussion as a basis for making sense of the world. In this perspective, the concept of the public sphere (Habermas, 1962, 1989) is useful for understanding how citizens make informed opinions thanks to a variety of ideas – in other words, through free speech. According to Jürgen Habermas' thesis, first published in German in 1962 and later translated into English, public opinion was born in England, Germany, and France during the seventeenth and eighteenth centuries. Public opinion was able to contest the state's might during a period when public authorities began to make rules for economic exchanges. Public discussion then extended to the cultural domain. Cafes, lounges, and the press helped to build public opinion, a prelude to democracy as it is known in the West.

This thesis is based on a pair of conditions. First, hierarchies related to status, wealth, or expertise would be suspended, allowing for more egalitarian exchanges and opening the way to a public conversation in which all citizens could participate. Second, the public sphere would be based on high-quality arguments – that is, on the public use of reasoning. Habermas argued that the public sphere deteriorated from the mid-nineteenth century on, as it was instrumentalized by political and economic powers (Gingras, 2009, pp. 13–14).

Habermas' habilitation thesis has been hotly contested from the historical, sociological, and political points of view. Although it was a scholarly work, Habermas' objectives were political, and his methods were unconventional. Indeed, his thesis was a defence of liberalism in the context of Germany's reconstruction in the 1950s. The German conservative elites wanted to preserve the economic and cultural substratum of liberalism while reinstating authoritarian structures. Habermas was opposed to this idea and tried to reconcile bourgeois civility and "publicity" – "representative publicity" associated with status – by adding normative values (access, equality, and transparency) to the concept. In a word, he wanted to reconcile liberalism – bourgeois civility – with publicity dissociated from its veneer of nobility (Simard, 2018). Interpreted in the light of his subsequent works, the public sphere became a paean to rational and egalitarian public conversation.

However, a public conversation in which expertise, status, and wealth have no sway, in which all citizens have access and equal treatment, and in which political, economic, and social questions are dealt with complete transparency is a utopia. The concept of public sphere, in its formation stage, rests on a prerequisite that is difficult to reconcile with the social conflicts and unequal relationships that exist in any society. Habermas admitted some weaknesses in his theory 30 years after writing his book in German (Habermas, 1992), but his vision of a society capable of self-government through informed public opinion remained largely intact. The intellectual success of his thesis can be explained by two factors: first, it embodies the liberal ideology based on belief in the ability of peoples to govern themselves; second, it embodies the critical perspective that is based on the idea of a

generalized instrumentalization of the public sphere by economic and political powers. Habermas therefore convinced the proponents of competing ideologies of the value of his thesis.

In liberal societies, elections are the privileged mechanism for self-government, but everything that represents the public sphere, devoid of its over-rationalism, embodies democracy's ideals: public conversation, the expression of conflict, and, of course, free speech (Gingras, 2008).

Freedom of Expression: Foundations and Limitations

Canadian courts have established three foundations for protecting freedom of expression (Hamilton, 2009, p. 2). First, free speech has an instrumental value, thanks to the exchange of ideas that allows for the emergence of new ideas and truth (*Bou Malhab* v. *Diffusion Métromédia CMR inc.*, 2011) in a "market of ideas" (Garton Ash, 2016, p. 75). False statements need not be censured in a marketplace of ideas: their existence in public debate is justified by citizens' ability to separate the wheat from the chaff. Indeed, Canadian courts rely on the "reasonable person" to assess the veracity of statements and their effects, even if they are unpleasant or hurtful. The reasonable person lives in a society in which free speech is of the utmost value and the courts do not judge morality or good taste.

Second, free speech, and political discourse in particular, is fundamental to democracy (*Irwin Toy* v. *Quebec (Attorney General)*, 1989). It strengthens institutions and makes it possible to criticize actions taken by the state. The right to vote would be sheer non-sense without citizens' ability to develop an informed opinion.

Third, free speech has an intrinsic value; it allows individuals to create and to inform themselves, and it ensures that everyone can express their thoughts, opinions, and beliefs, however unpopular, distasteful, or contrary to the mainstream (*Irwin Toy* v. *Quebec (Attorney General)*, 1989). In short, it is essential for individual autonomy. Both legal scholars and philosophers, from John Stuart Mill to C. Edwin Baker, have emphasized the importance of fostering citizens' autonomy (Maclure, 2017). The sovereignty of the individual applies to both the speaker and the recipient of information (Garton Ash, 2016, p. 75). Art is a form of legally protected expression. It enables us to understand the experience of others in many ways, with what Martha Naussbaum has called the "inner eye" of imaginative sympathy and Ricoeur has dubbed "Oneself as Other" (Garton Ash, 2016, p. 244).

The more closely linked a contentious expression is to one of intrinsic or instrumental value, the more essential free speech is considered to the balancing of rights. However, freedom of expression is not without its limits, and Western societies do not operate solely on the basis of the individualistic and market foundations of liberalism. In fact, the liberal approach fails to reconcile the right to free speech that protects racist, sexist, xenophobic discourse with other values essential to personal development, such as dignity, self-esteem, and reputation for members of the groups targeted by this type of discourse.

> The free speech debate reveals a limit of liberal political morality and therefore leaves liberal normative theorists with an uncomfortable predicament, as they have to rely more on the

> complementary role of pro-social personal provisions and civic virtues than they generally wish to. (Maclure, 2017, p. 135)

In Canada, free speech is limited by provisions in the Criminal Code, such as those covering hate propaganda and child pornography. It is also limited by rights enshrined in charters, such as the right to equality, the right to life and security of the person, and the right to preserve one's dignity, honor, and reputation (*CDPDJ* v. *Ward*, 2016, at para. 75; Mahoney, 1992, p. 79; Zhou, 2019). Furthermore, it is limited by Section 1 of the Canadian Charter of Rights and Freedoms (equivalent to Section 9[1] of the Quebec Charter of Rights and Freedoms), which "guarantees the rights and freedoms set out in it subject only to such reasonable limits prescribed by law as can be demonstrably justified in a free and democratic society" (Canada, 1982). The courts apply a multistage test that examines the balance between the positive and negative effects of the provision limiting the Charter right and its objective. To limit a right, the matter must be of sufficient social importance (Hamilton, 2009).

> The section 1 process is one of the key elements that distinguishes the Canadian and American approaches of free expression. Because there is no similar limit built into its constitution, United States courts have had to read limits into the right itself, and have been reluctant to do so except in the most extreme instances. In Canada, there are virtually no limits on the rights as stated in section 2(b) [covering freedom of thought, belief, opinion, and expression, including freedom of the press and other communication media], but the courts have frequently found that legislative limits on free expression meet the section 1 test. (p. 4)

Over a three-decade period, Canadian courts have adopted a more inclusive, egalitarian approach to free speech. When the courts assess the claims of adversaries in a litigation, they engage in a balancing exercise that takes into account the greater impact of the speech on disadvantaged members of society than on non-disadvantaged members. Thus, although the value of autonomy is always sought in questions of free speech, it is possible not to treat autonomy as always equivalent for the two parties (Little, 2012, p. 97; Mahoney, 1992, p. 78; Zhou, 2019).

A more egalitarian approach also presupposes that some dispute mechanisms, such as human rights bodies, should exist to take into account situations in which there is inequality between the parties. Canadian jurisprudence, like jurisprudence in the UK and New Zealand and unlike that in the United States, is inclined to favor the redress of wrongs on the grounds of equality rights (Little, 2012, p. 149).

It is this quest for balance that the courts engage in cases involving freedom of expression. In the case under consideration, the question is whether, and under what conditions, comments made in a comedy show can be limited by virtue of the right to equality based on dignity, honor, and reputation.

The issue of free speech has been addressed in famous cases concerning humor, such as that of Muhammad's cartoons in Denmark and France (Klausen, 2009), that involving Dieudonné in France (Bem, 2015; Leplongeon, 2015; Leloup & Leplongeon, 2015), and that, less well known, of the radio station CHOI-FM in Quebec City in 2004–05 (Gingras, 2017). It should be noted from these cases that concerns other than the issues before the courts – that is, in addition to free speech and the rights claimed by the complainants – must be considered. For

example, to fully understand the case of Muhammad's cartoons, one must take into account diplomatic negotiations, the instrumentalization of the cartoons by the Muslim Brotherhood, attacks on Danish and French businesses, and demonstrations in Africa and Asia that resulted in several hundred deaths. Much more than free speech was involved in this case.

GABRIEL VERSUS WARD

Litigation: Discrimination Versus Freedom of Expression

The dispute under discussion here concerns Ward's jokes about Gabriel in the skit "Les Intouchables," part of the show *Mike Ward s'eXpose*. The skit was first performed in bars in 2009, then presented 230 times between 2010 and 2013 in theaters. Recorded on video in 2012, *Mark Ward s'eXpose* was also offered on DVD, which had sold 7,500 copies as of 2016. Three video capsules on Gabriel were also posted on Ward's website, and they had been viewed 60,000 times as of 2016 (Patriquin, 2016). It should be noted that the prosecution was based on an allegation not of defamation but of discrimination. The aim was to assess whether free speech can be limited by the right to equality based on the preservation of reputation, dignity, and honor. If it can, what elements and what circumstances justify this violation of free speech?

An allegation of defamation would have required hiring a lawyer, and because Gabriel's parents were unable to do so, they turned to the CDPDJ. In Canada, the Supreme Court has recognized that human rights commissions can deal with cases in which free speech is used as a defence (*Saskatchewan* v. *Whatcott*, 2013).

The Facts About Jérémy Gabriel

At the time of the first alleged events, Gabriel was a teenager with Treacher Collins syndrome, which causes certain head malformations, especially in the ears and palate. Gabriel was also born with severe deafness, a problem partially resolved in 2003, when he was six years old, by the implantation of a bone-anchored hearing aid. Because of an immune deficiency, he had to receive blood transfusions every three weeks. By 2016, he had undergone 23 surgeries requiring general anesthesia (*CDPDJ* v. *Ward*, 2016).

As soon as he could hear and speak, Gabriel longed for a singing career, and his parents decided to help him in order to prove that his handicap was not an obstacle to his personal development. Between 2005 and 2010, Gabriel appeared on several television programs, sang the national anthem before a hockey game in Montreal, met Céline Dion and the pope, released an album, published an autobiography, and participated in a documentary on Treacher Collins syndrome. In 2012, he became a patient ambassador for Shriners Hospitals, for which he traveled in Canada and the United States. His activities were generally performed on a voluntary basis, and the income from his book sales helped to pay some expenses (*CDPDJ* v. *Ward*, 2016).

The Facts About Mike Ward

Ward has been a professional stand-up comedian since 1996. According to his lawyer, he uses "black humour." Ward claims that he wants to break taboos and assess his audience's ability to laugh at everything. "Les Intouchables" focused on Quebec public figures of whom it would be impossible to make fun "without getting into shit." Ward attacked Céline Dion for changes to her physical appearance, Dion's manager René Angelil for his sensitivity, singer, and author Ariane Moffat for being overweight, and radio host Jacques Languirand for his eyebrows, among others. The only controversy in the skit that raised concerns were the jokes about "little Jérémy." Referring to his bone-anchored hearing aid, Ward described Gabriel as "the young man with a sub-woofer on his head." He claimed to believe that Gabriel had participated in the Make a Wish charity for children with terminal diseases but that he was mistaken, as Gabriel was still not dead five years later. He went on to say that he met him in a pool store and tried to drown him, but ultimately found out that he was "unkillable." Gabriel's disease, Ward said, was simply that he was utterly ugly. "Les Intouchables" ends with an admission of the reckless nature of his words: "I didn't know how far I could go with this gag. I said to myself at one point, I will go too far, they will stop laughing. But no you didn't … ya bunch of …" (*CDPDJ* v. *Ward*, 2016).

Ward has also produced three video capsules about Gabriel for his website: in one, he has him sing a bawdy song, and another one has a connection with pedophilia, in reference to his meeting with the Pope and to the accusations and suspicions that Catholic priests are pedophiles (*CDPDJ* v. *Ward*, 2016).

The Legal Foundations of the Case

In 2012, when asked about "Les Intouchables" in a television interview, Ward explained that he was like a cocaine addict: he needed to cross boundaries. The interview convinced Gabriel's family to file a complaint with the CDPDJ. The first step in the commission's procedure is mediation between the parties involved; if there is no satisfactory settlement, the CDPDJ may send the case to the Human Rights Tribunal. This is what happened with the Gabriel–Ward litigation.

In the analysis of the case, the Tribunal referred to several articles of the Quebec Charter of Human Rights and Freedoms. Article 3 deals with several fundamental freedoms, including freedom of expression. Article 4 deals with the right to the safeguard of one's dignity, honor, and reputation. Article 10 deals with the right to recognition and the exercise of the right to equality without distinction, exclusion, or preference based on a series of grounds, including age, disability, and the use of a means to overcome the handicap.

The combination of Articles 4 and 10 of the Charter protects the right to equality in the recognition and exercise of the rights to dignity, reputation, and honor (*CDPDJ* v. *Ward*, 2016). Before the court, evidence of the effects of Ward's jokes about Gabriel was demonstrated: bullying concerning his mouth, which does not close completely, and his bone-anchored hearing aid, directly inspired by Ward's jokes, and his academic difficulties, depression, and suicidal thoughts. Ward's jokes were also picked up by people in the workplaces of both

of Gabriel's parents, which explains why they were first included in the human rights submission to the court.

Ward claimed that being in court for a joke is ridiculous. He explained that he chose Gabriel as a fall guy because he was a public figure and a child, because he is "weak," and because jokes about him would cause discomfort. He said that he was aware that his comments about Gabriel were harsh, and he relied on the laughter of the audience as a criterion for assessing their acceptability. Ward contends that laughing at a disabled person is evidence that we do not pity him or her, but that we integrate him or her into society. He was never concerned with the impact of his jokes on Gabriel (*CDPDJ* v. *Ward*, 2016, at para. 132). When his manager told him about the complaint to the CDPDJ, he thought that it was a "just a question of money" and did not stop using the jokes about Gabriel (*CDPDJ* v. *Ward*, 2016, at para. 166). If Ward's free speech was to be restricted, his lawyer argued, all Quebec comedians would work with a sword of Damocles over their heads from then on; there would be a definite chilling effect (*CDPDJ* v. *Ward*, 2016).

The human rights tribunal ruled that some of Ward's jokes were discriminatory (those about Gabriel's physical appearance and bone-anchored hearing aid) and infringed on Gabriel's equality based on the preservation of his dignity, reputation, and honor. The tribunal considered the pedophilia jokes and the bawdy song not discriminatory. The court then looked at the alleged defence, free speech: could it be limited by the right to equality based on the preservation of dignity, honor, or reputation? The Court of Appeal decided that freedom of expression was applicable not as a ground of defence but as the second step in the test leading to determination of the presence, or absence, of discrimination. In defamation matters, the courts have established two criteria for determining whether the damage to a person's reputation is justified by free speech: truthfulness and public interest (*G. Néron* v. *Chambre des notaires du Québec*, 2004). The Human Rights Tribunal finds these criteria appropriate in cases of discriminatory offence as well. The analysis of disputes must also take into account the context of utterance, the tone used, the identity of the author of the remarks, and the identity of the victim (*CDPDJ* v. *Ward*, 2016, para. 130). When the victim is a group, the Supreme Court of Canada has often explained that hurtful words, taunts, or name-calling have to be protected by freedom of expression. In the case under discussion here, however, the victim was not a group but an individual.

Humor and Freedom of Expression: The Results of the Prosecutions

"If humor can have educational virtues," wrote the human rights tribunal, it must be done with respect for the dignity of the person concerned. "Humor cannot be used as a pretext, a screen or a justification for discriminatory conduct" (*CDPDJ* v. *Villemaire*, 2010). In fact, it can even be an aggravating factor:

> Depending on the circumstances, it may even constitute a particularly insidious form of discrimination. Attempting to make people laugh about personal characteristics which are prohibited grounds of discrimination carries an obvious risk of trivializing the prohibition of discrimination. Humor can have the effect of further isolating the person who is the subject of

> discrimination, by drawing laughers to the side of the mocker and by discrediting the protests of the victim of discrimination. (*CDPDJ* v. *Villemaire*, 2010, at para. 46, our translation)

The court emphasized that having a digital platform imposes some responsibilities. A stand-up comedian cannot act solely on the basis of his audience's laughter; he must also take into account the fundamental rights of the victims of his jokes (*CDPDJ* v. *Ward*, at para. 134).

The tribunal's decision was based on several elements: Ward's jokes did not raise an issue of public interest or contribute to a debate about questions of general interest; they were repeated from 2009 to 2013 and contributed to the comedian's financial success; Ward refused to stop joking about Gabriel in his performances even though he stated that he was ready to do so if the targeted person felt injured; he arrogantly claimed that the complaint was only a question of money (*CDPDJ* v. *Ward*, at para. 137 and 139). The Court of Appeal of Quebec also specified that in 2019, the capsules about Gabriel were still likely to be available. Furthermore, although public figures are exposed to increased media attention, accepting the risks of the profession do not amount to a renunciation of their fundamental rights, especially when the person concerned is a teenager (*Ward* v. *CDPDJ*, 2019 at para. 207). By attacking a person with a disability and associating him with a notion of weakness (para. 208), Ward revived or perpetuated a stereotype that disabled persons are worthless than non-disabled ones. Gone were the days when handicaps could be used for entertainment (*Ward* v. *CDPDJ*, at para. 208).

Both the Human Rights Tribunal and the Court of Appeal have stated that humor impinging on the grounds of discrimination set out in Section 10 of the Charter of Human Rights is not prohibited per se. Such cases can, in no way, be interpreted as a green light to censor comedians. However, the repercussions of comedians' actions must be considered when assessing the claims in a litigation, since free speech has no prerogative over other rights in the charter (*CDPDJ* v. *Ward*, 2016, at para. 135).

THE OTHER ISSUES IN THE LITIGATION

How Laughter Occurs

Labeled as "black humor" by his lawyer in court, Ward's humor is rather hybrid: a little bit of black humor, some trash talk – as Ward himself says in his show *Mike Ward au bordel* (at the brothel) – and vulgar humor, according to his own description (Kelly, 2013). Ward's shows and video capsules are similar to mild jock shock: a mix of stories of a sexual or scatological nature featuring himself and his friends, jokes, rumors, swear words, and attacks on public figures. Ward is described as a "habitual line-stepper" (Patriquin, 2016). He says that he wants to break taboos. For example, racism is dealt with unlikely situations that bring out the ridiculousness of racists, but these can be first- or second-degree jokes. He uses Quebec slang and a direct, raw, inflammatory style with which he aims to surprise his audience, a phenomenon that leads to laughter.

In a classic work, *Le rire* (translated as *Laughter*), written in 1900, French philosopher Henri Bergson explained that laughter is an automatic reaction, a spontaneous response caused by surprise, by the association of the body with a mechanism, or by the exploitation of character defects, among others. It is incompatible with emotion – that is, it is insensitive, indifferent: there is, in laughter, an "anaesthesia of the heart." Often caused by malice or wickedness, it discourages sympathy and strikes without distinction. Laughter does not provoke reflection; it develops intrinsically, without thought, in a moment of relaxation. People laughing are further stimulated by hearing the laughter of others, in a context removed from reality (Bergson, 1900, 1972).

No wonder, then, that Ward's audience did not back away from his jokes. There are, in fact, no limits to jokes when collective laughter prevails. Ward's audience knows and appreciates his style, as evidenced by his immense success in Canada (Gagnon, 2019; Kelly, 2013; Patriquin, 2016). He has won several awards over the years, including that of Comedic Artist of the Year in 2016. Audiences for his shows come to listen to him for fun, and they expect to laugh a good deal, especially when he makes unexpected jokes in which he alternates the congruous and the incongruous.

The jokes about Gabriel in *Mike Ward s'eXposed* are incongruous, because they follow ones about several famous figures. Among them is Céline Dion: Ward talks about a well-known pedophile who would not have deigned to touch her when she was young, because she was too ugly. He then adds, "If I don't have a lawsuit against me now, it's not for the lack of trying" (Patriquin, 2016). Making fun of artists and public figures is common in humor. So, when he makes jokes about Gabriel, Ward is changing the type of victim, a fact that increases incongruity. After the Human Rights Tribunal decision, Ward conceded, "The difference between the Céline and Jérémy bit was that with Céline, I was punching up. But with the kid, I was punching down. In the court of public opinion I already lost" (Patriquin, 2016). These words echo a concept of humor that Ward knows but does not apply, that of American satirist H. L. Mencken, who described his role as "To comfort the afflicted and afflict the comfortable" (Garton Ash, 2016, p. 244).

Gabriel was not the first child to have been targeted in Ward's shows, although most of the others were not affected by his jokes as they had died. In 2008, Ward joked about a child from Trois-Rivières who had been abducted in front of witnesses while she was playing in a public park, a fact that moved all of Quebec – she was found 10 years later, murdered. It was similar with the joke about "Guy Turcotte who has no children" in the show *Mike Ward au bordel*. Turcotte is a Quebec doctor who murdered his three- and five-year-old children with 46 stab wounds in 2009. Ward said those words in the middle of the show, while the audience's contagious laughter was spreading without restraint; the discomfort amplified the reactions. Black humor does indeed elicit a huge amount of laughter in an audience that has been conditioned.

The alternation between congruence and incongruity is part of the method aimed at fostering an atmosphere of collective laughter (Little, 2012, pp. 104–105):

> Another quality known to enhance comedic effect derives from the social quality of humor. A listener often knows from a joke teller's cue that the joke teller seeks to make them laugh. Through symbiotic mental cooperation with the joke teller, the listener anticipates that a punch line is coming, doesn't know what it will be, and experiences both fulfillment and surprise when it arrives. (Little, 2012, p. 107)

Ward's Growing Audience and Its Repercussions

Ward's words and the video capsules about Gabriel reached an audience beyond those who bought tickets for his shows. In 2016, it was estimated that the capsules had been viewed 600,000 times (Patriquin, 2016). Ward's profile went viral on social media, so the jokes about Gabriel in "Les Intouchables" were taken up by Internet users – and by Gabriel's classmates – just when he was starting high school. The human rights complaint was filed after Ward appeared in television interviews and the bullying at school and on social media was triggered by his comments. Intention is not dealt with matters of discrimination, which means that the effects of Ward's jokes, intended or not, were at issue (*CDPDJ* v. *Ward*, 2016, para. 80).

After the complaint was filed with the CDPDJ, conciliation would have been possible if Ward had agreed. He was asked to stop making jokes about Gabriel and refused to do so, claiming that the complaint was only a question of money.

But for Gabriel, it was not: he suffered from being mocked by his peers. Bullying at school is a widespread phenomenon. In Western countries, it is estimated that between 10 and 20 percent of children are involved, as victim or aggressor, in bullying incidents (Carter & Spencer, 2006; Salmivalli & Peets, 2018, p. 304). There are three elements in this kind of incident: asymmetry in power relations, repetition, and the intention to do harm (Salmivalli & Peets, 2018, p. 303). Children with disabilities are frequent prey (Houchins, Oakes, & Johnson, 2016; Mulvey, McMillian, & Irvin, 2018; Yell et al. 2016), more than students without disabilities (Carter & Spencer, 2006). Groups tend to intimidate students who receive additional help at school (Carter & Spencer, 2006), a measure taken in programs aimed at integrating students with disabilities. Being weak and having few friends are other predictors of victimization (Salmivalli & Peets, 2018, p. 310).

Bullying in school is one of the most traumatic experiences a child or teenager can have. Through the internet, messages and videos can be viewed repeatedly, expanding the audience for such incidents and making them even more hurtful for the victims. There may also be more than one bully, a fact that multiplies attacks (Salmivalli & Peets, 2018, p. 303). Bullying at school can have major consequences: severe damage to self-esteem, feelings of estrangement, diminished academic performance, reduced concentration, truancy, running away from home, avoidance behavior, depression, and suicide (Carter & Spencer, 2006; Salmivalli & Peets, 2018, p. 305).

In a study on disparagement humor, Ford and Ferguson (2004) contend that this type of speech opens the door to greater tolerance for discriminatory events. The rules of the social game are modified, distorted, extended, so that tolerance of discrimination is increased. The many hate messages received by Gabriel

the day after his victory in the Court of Appeal illustrate this phenomenon (Poirier, 2019).

In his defence before the Human Rights Tribunal, Ward explained that laughing at a disabled person is a message that we do not pity him or her but include him or her in society. So, what is inclusion?

> Inclusiveness … is the property of society which makes it possible for everyone to feel that they belong to them as equal members, that their place and status is not put into question, regardless of their identity-conferring attributes or conceptions of the good. (Maclure, 2017, p. 141)

In Gabriel's case, it is clear that, far from inclusion, Ward's comments had the effect of isolating the targeted person, making his life difficult. Even after the complaint to the CDPDJ, Ward continued to joke about Gabriel.

Ward's Publicity Strategy

Although free speech is a legitimate ground of defence, it should also be emphasized that laughter is a source of income for Ward. His interest in Gabriel was based on his belief that the discomfort of a joke about a disabled child would amplify laughter among an already conditioned audience. Ward's career thrives on show business, and the more shows he does, the more his income increases. Any judicial proceedings are bound to give him free publicity.

Indeed, faced with a conciliation procedure at the Human and Youth Rights Commission, Ward decided not to cooperate; rather, he chose to be sued. This allowed him to publicize the lawsuit and to generate support for his cause in the cultural community. He had previously expected to be sued by Céline Dion, and that would have had a resounding media effect, but it had not occurred. The media impact of Gabriel's complaint to the CDPDJ was smaller, but it was nonetheless free publicity.

Once the complaint was filed, in 2012, the various stages of the prosecution and the outcome of the legal proceedings generated free publicity for Ward, bringing him to the attention of people who had not heard of him (Hamelin, 2019). He never missed an opportunity to talk about his trial, whether when receiving an award or on stage, as in his show *Mike Ward au bordel*. The media covered the case extensively, particularly in 2016, when the CDPDJ rendered its decision, and in fall 2019, before and after the ruling by the Court of Appeal of Quebec, which upheld the commission's decision two to one. Ward received almost unanimous support from the Quebec humor community (Gagnon, 2019), though not from the journalism community (Cassivi, 2019; Tremblay, 2019). He appealed the decision to the Supreme Court of Canada, which had indicated it would hear the case.

CONCLUSION

Free speech disputes are rarely limited to the confrontation of two rights. They fall within a social environment involving multiple issues, including commercial profitability, notoriety, rivalry, human rights issues, ethical considerations, and

profound cultural differences. The courts are limited to the framework imposed by law in their exercise of balancing the rights of parties to a dispute, but a broader context can shed more light on free speech conflicts.

What is free speech for? Do "Les Intouchables" and the jokes about Gabriel contribute to an enlightened public sphere in an egalitarian conversation? What can we say about the intrinsic value of free speech expressed in Ward's show – that is, the freedom for individuals to emancipate themselves, to create, and to inform themselves – in short, to promote personal autonomy? What can we say about the instrumental value of free speech, which should encompass exchanges of opinions, the emergence of new ideas, criticism of state action, and the emergence of truth?

The above-mentioned virtues usually attributed to free speech do not appear to have been present in the case at hand. Rather, I contend, the intrinsic values of personal development and autonomy for Ward are precisely those that are denied to Gabriel. As for the instrumental value, Ward's goal of showing that there are untouchables in Quebec society that we cannot laugh at without risk of being sued, this idea amounted to knocking on an open door. Public figures have always been fair game in humor in Quebec, Canada, and elsewhere in the West. Some have sued; some have not.

Humor is evolving, as the internet broadens audiences, invades privacy, and heightens culture shock. Jokes can be received by the public in many ways, a fact that takes on a particular twist when they are *double entendres*. We have to take into account that there are unexpected audiences for Western adult humor, whether it is people from other countries – as in the case of Muhammad's cartoons – or children.

Humor also evolves because political cultures change. Cartoons featuring the caricature of greedy, hooked-nosed Jews, which abounded in European newspapers until the end of the last world war, are no longer being drawn. Jokes about *bamboulas* and *bougnoules*, which infuriated audiences in former colonies and regions with a past involving enslavement, are no longer tolerated (Tremblay, 2019).

Freedom of expression in the age of cyberspace and disinformation poses immeasurable challenges. Without the intrinsic or instrumental values, is free speech worth defending?

REFERENCES

Bem, A. (2015). Spectacle de Dieudonné contraire à la liberté d'expression et aux valeurs de justice et de paix. *Cabinet Bem*. Retrieved from https://www.legavox.fr/blog/maitre-anthony-bem/spectacle-dieudonne-contraire-liberte-expression-19095.htm. Accessed on June 3, 2020.

Bergson, H. (1900, 1972). *Le rire. Essai sur la signification du comique* (303rd ed.). Paris: Presses Universitaires de France.

Bou Malhab v. *Diffusion Métromédia CMR inc*. (2011). Retrieved from https://scc-csc.lexum.com/scc-csc/scc-csc/en/item/7921/index.do. Accessed on May 25, 2020.

Canada. (1982). Constitution Act. Retrieved from https://laws-lois.justice.gc.ca/eng/Const/page-15.html. Accessed on June 10, 2020.

Carter, B. B. & Spencer, V. G. (2006). The Fear Factor: Bullying and Students with Disabilities. International Journal of Special Education, *21*(1), 11–23.

Cassivi, M. (2019, December 9). Il y a (toujours bien) des limites. *La Presse*. Retrieved from https://www.lapresse.ca/arts/humour/201912/09/01-5252965-il-y-a-toujours-bien-des-limites.php. Accessed on May 12, 2020.

Commission des droits de la personne et des droits de la jeunesse v. *Villemaire*. (2010). QCTDP 8. Retrieved from http://t.soquij.ca/k2F9D. Accessed on May 25, 2020.

Commission des droits de la personne et des droits de la jeunesse (Gabriel et autres) v. *Ward*. (2016). QCTDP 18 (CanLII). Retrieved from http://canlii.ca/t/hq9d9. Accessed on June 3, 2020.

Ford, T. E., & Ferguson, M. A. (2004). Social consequences of disparagement humor: A Prejudice Norm Theory. *Personality and Social Psychology Review*, *8*(1), 79–94.

Gagnon, H. (2019, December 12). La puissante industrie de l'humour. *Le Devoir*. Retrieved from https://www.ledevoir.com/opinion/idees/568880/la-puissante-industrie-de-l-humour. Accessed on May 5, 2020.

Garton Ash, T. (2016). *Free speech. Ten principles for a connected world*. New Haven, CT: Yale University Press.

Gilles E. Néron Communication Marketing Inc. v. *Chambre des notaires du Québec* (2004). 3 S.C.R. 95. Retrieved from https://scc-csc.lexum.com/scc-csc/scc-csc/en/item/2168/index.do. Accessed on May 16, 2020.

Gingras, A.-M. (2008). Les représentations sociales de la démocratie: Réflexivité, effervescence et conflit. *Politique et Sociétés*, *27*(2), 11–40.

Gingras, A.-M. (2009). *Médias et démocratie. Le grand malentendu*. Quebec: Presses de l'Université du Québec.

Gingras, A.-M. (2017). Freedom of expression, entertainment, hate speech, and defamation: Where do we draw the line? In L. Taylor & C.-M. O'Hagan (Eds.), *The unfulfilled promise of press freedom in Canada* (pp. 84–97). Toronto: Toronto University Press.

Habermas, J. (1962, 1989). *The structural transformation of the public sphere: An inquiry into a category of Bourgeois Society*. Cambridge: Polity.

Habermas, J. (1992). L'espace public trente ans après. *Quaderni*, *18*, 161–191.

Hamelin, M. (2019, December 2). Quand le mieux et l'ennemi du bien. *L'Actualité*. Retrieved from https://lactualite.com/culture/mike-ward-quand-le-mieux-est-lennemi-du-bien/. Accessed on May 15.

Hamilton. S. (2009). *Law's expression. Communication, laws and media in Canada*. Markham, ON: LexisNexis.

Houchins, D. E., Oakes, W. P., & Johnson, Z. G. (2016). Bullying and students with disabilities: A systematic review of intervention studies. *Remedial and Special Education*, *37*(5), 259–273.

Irwin Toy v. *Quebec (Attorney General)*. (1989). Retrieved from https://scc-csc.lexum.com/scc-csc/scc-csc/en/item/443/index.do. Accessed on April 29, 2020.

Kelly, B. (2013, July 23). Mike Ward speaks our languages. *The Gazette*. Retrieved from http://www.montrealgazette.com/entertainment/Mike+Ward+speaks+languages/8693272/story.html. Accessed on May 15, 2020.

Klausen, J. (2009). *The cartoons that shook the world*. New Haven, CT: Yale University Press.

Leloup, D., & Laurent, S. (2015, January 14). "Charlie", Dieudonné … Quelles limites à la liberté d'expression? *Le Monde*. Retrieved from https://www.lemonde.fr/les-decodeurs/article/2015/01/14/de-charlie-a-dieudonne-jusqu-ou-va-la-liberte-d-expression_4555180_4355770.html. Accessed on June 3, 2020.

Leplongeon, M. (2015, November 10). Liberté d'expression: Dieudonné subit un camouflet devant la CEDH. *Le Point*. Retrieved from https://www.lepoint.fr/justice/liberte-d-expression-dieudonne-subit-un-camouflet-devant-la-cedh-10-11-2015-1980465_2386.php, 2020. Accessed on June 3, 2020.

Little, L. E. (2012). Just a Joke: Defamatory humor and incongruity's promise. *Southern California Law Journal*, *95*, 93–164.

Maclure, J. (2017). The regulation of Hateful and Hurtful Speech: Liberalism's uncomfortable predicament. *McGill Law Journal*, *63*(1), 133–154.

Mahoney, K. (1992). The Canadian constitutional approach to freedom of expression in hate propaganda and pornography. *Law and Contemporary Problems*, *55*(1), 77–105.

Mulvey, K. L., McMillian, L., & Irvin, M. J. (2018). Youth cognition surrounding bullying of peers with disabilities: Inclusion, intervention and the role of the group. *Journal of Emotional and Behavioral Disorders*, *28*(1), 17–28.

Patriquin, M. (2016, March 23). When comedy crosses the line in Quebec. *Maclean's*. Retrieved from https://www.macleans.ca/news/canada/should-mike-wards-bad-joke-be-censored/. Accessed on May 15, 2020.

Poirier, C. (2019, December 1). Jérémy Gabriel ciblé par de nombreux messages de haine après sa victoire en Cour d'appel. *Le Journal de Montréal*. Retrieved from https://www.journaldemontreal.com/2019/12/01/jeremy-gabriel-cible-par-de-nombreux-messages-haineux-apres-sa-victoire-en-cour-dappel. Accessed on May 15, 2020.

Salmivalli, C., & Peets, K. (2018). Bullying and victimization. In W. M. Bukowski, B. Laursen, & K. H. Rubin (Eds.), *Handbook of peer interactions, relationship and groups* (2nd ed., pp. 302–321). New York, NY: The Gilford Press.

Saskatchewan (Human Rights Commission) v. *Whatcott*. (2013). S.C.C. 11. Retrieved from https://scc-csc.lexum.com/scc-csc/scc-csc/en/item/12876/index.do. Accessed on June 2, 2020.

Simard, A. (2018). Le charme discret de la bourgeoisie. Jürgen Habermas et les origines de l' "espace public". In A.-M. Gingras (Ed.), *Histoires de communication politique. Pratiques et état des savoirs* (pp. 237–258). Quebec: Presses de l'Université du Québec.

Tremblay, O. (2019, December 5). Par-delà l'affaire Jérémy Gabriel. *Le Devoir*. Retrieved from https://www.ledevoir.com/opinion/chroniques/568462/par-dela-l-affaire-jeremy-gabriel. Accessed on May 16, 2020.

Ward v. *Commission des droits de la personne et des droits de la jeunesse (Gabriel et autres)*. (2019). QCCA 2042 (CanLII). Retrieved from http://canlii.ca/t/j3p58. Accessed on June 3, 2020.

Yell, M. L., Katsiyannis, A., & Rose, C. A. (2016). Bullying and harassment of students with disabilities in schools: Legal considerations and policy formation. *Remedial and Special Education*, *37*(5), 274–284.

Zhou, H.-R. (2019). Ford et Irwin Toy 30 ans plus tard: une conversation avec le juge de Montigny (interview). *Constitutional Forum*, *28*(3), 39–50.

CHAPTER 8

HATE SPEECH, MEDIA, AND CANADIAN FEDERAL LAW

Allyson M. Lunny

ABSTRACT

Purpose – *This chapter has three general purposes: to trace Canada's hate speech laws from their policy inception to their current state; to identify the importance that media and mass communication have played in the creation and development of Canada's hate speech laws; and to demonstrate the critical relationship that media has had to significant legal cases on hate speech.*

Methodology/Approach – *This chapter historically maps the policy development of and legal challenges to Canada's hate speech laws. It takes directed notice of the relationship of media and mass communication to the development and implementation of those laws. It engages with libertarian and egalitarian arguments on free speech throughout the chapter testing these ideas through an examination of the legal cases cited.*

Findings – *Canadian legislators and courts have long grappled with the balancing of rights with respect to the issue of "hate speech." Advances in mass communication technology have added intricate challenges to that legal balancing. Awareness of media's allure to hatemongers and racial extremists and of media's protean characteristics make regulation of its hateful content a continuous legal challenge. Canada's greatest challenge yet to the regulation of hate speech will be its adaptive response to the growing phenomenon of online hate.*

Originality/Value – *This chapter highlights the little recognized prescient statements made by the Cohen Committee about the allure of media and the dangers of its technological advancements in Canadian free speech debates.*

Media and Law: Between Free Speech and Censorship
Sociology of Crime, Law and Deviance, Volume 26, 133–149

ISSN: 1521-6136/doi:10.1108/S1521-613620210000026009

Providing a comprehensive survey of Canada's "hate speech" laws, it recognizes the importance that advancements in mass communication have played in the creation and development of Canada's "hate speech" laws.

Keywords: Hate speech; freedom of expression; freedom of speech; libertarian rights; egalitarian rights; Cohen Report

INTRODUCTION

Canadian legal and policy experts have long recognized the inherent tension between the limiting of public expression in terms of regulation and criminalization and the right of free expression, without government interference, in a liberal democracy. Under Canadian law, criminalizing or regulating speech must be balanced against liberal rights and democratic principles. With respect to hate speech specifically, twentieth century historical atrocities, international human and civil rights commitments, and cultural shifts in the understanding of the impact of mass communication have led to the creation of Canada's hate speech laws (Lunny, 2017). Federal regulation of hate speech, prior to the repeal of section 13 of the *Canadian Human Rights Act* (*CHRA*) in 2013, fell primarily under criminal law and human rights jurisdictions. This chapter examines the history and rationale for the creation of Canadian federal hate speech laws, significant hate speech cases that involved the use of media for the dissemination of extreme hatred, and current policy debate with respect to the internet, hate speech, and governmental intervention.

This chapter historically maps the policy development of and legal challenges to Canada's hate speech laws. It takes directed notice of the relationship of media and mass communication to the development and implementation of those laws. It engages with libertarian and egalitarian arguments on free speech throughout the chapter testing these ideas through an examination of the legal cases cited.

Questions that this chapter will consider are: how have media and its technological advancements effected Canadian federal hate speech laws? Why is the communication and dissemination of ideas central to the debate between free speech and hate speech? With the internet as the foremost platform of hate speech internationally (Canada, 2019, p. 7; Mahoney, 2009, p. 322), how has the Canadian government responded to online hate?

THE ORIGINS OF CANADA'S HATE PROPAGANDA LAWS

In January 1965, the then minister of justice, the Honourable Guy Favreau, appointed a Special Committee on Hate Propaganda headed by Maxwell Cohen, dean of McGill University's Law School "to study and report upon the problems related to the dissemination of varieties of 'hate propaganda' in Canada" (Cohen, 1966, p. 1). Advisory in scope, the Cohen Committee, as it became

known, recommended that changes to Canadian law were necessary, particularly changes to the *Criminal Code*. Noting the inadequacy of Canadian law regarding group defamation, it recommended the creation of three new criminal offences: the advocation or promotion of genocide against an identifiable group, the public incitement of hatred against an identifiable group that would likely lead to a breach of the peace, and the willful promotion of hatred against an identifiable group. The framing of criminal necessity revolved around three significant developments at this time: an increasing rise in anti-Semitic incidents and the organized lobbying of the Canadian government by the Canadian Jewish Congress (CJC); human rights advances provincially, nationally, and internationally; and cultural shifts in the understanding of mass communication invigorated by the Marshall McLuhan media revolution (Lunny, 2017, p. 31).

Discriminatory and violent acts of anti-Semitism in Canada existed well before 1965. Since the early twentieth century, Canadian Jews experienced discrimination with respect to housing, employment, insurance, education, and access to public services (Tulchinsky, 2008). In some instances, these policies and practices of discrimination were accompanied by overt acts of violence, including those of intimidation, assault, and the destruction of property. By the 1960s, however, a number of socially significant shifts coalesced to produce the criminal law recommendations put forward by the Cohen Committee: a rise in anti-Semitic activity; developments in human rights legislation; and a prescient awareness of the advancements of media communications.

In the 1960s, there was an increased presence of neo-Nazi and white supremacist activity across Canada that included the wide-spread dissemination of printed hate propaganda materials (Cohen, 1966, pp. 11–25). Much of the printed white supremacist and anti-Semitic materials were produced by American organized hate groups and crossed the border either through the postal system or through customs (Cohen, 1966; Johnson, 2019, pp. 49–50). In its study of the distribution of hate materials, the Cohen Committee noted that

> the propaganda had been plastered and displayed in conspicuous places, distributed at public gatherings, placed in apartment mailboxes, sent through the regular mails, and scattered from the upper floors of downtown buildings to the streets below. (p. 12)

In the spring of 1965, the Canadian Nazi Party under the leadership of John Beattie organized an anti-Semitic rally in Toronto's Allan Gardens. Prominently wearing swastika armbands and using a megaphone to amplify their message, Beattie's group deliberately flaunted a city by-law that stated "no person shall publicly preach, lecture, declaim, recite, harangue or engage in any other form of public speaking in a City park except at a time and place approved by permit" (1963 amendment to Toronto Bylaw 21379). In response to this provocation, the neo-Nazi rally met with counter-protest and violent opposition by anti-Nazi groups whose objective was to raise public awareness of the threat and to fight for legislation against expressions of racial hatred (MacGuigan, 1966, pp. 233–235). The legislative response that followed were attempts at the municipal level to utilize city park bylaws to restrict unlawful assembly in public space. For example, one such bylaw prohibited the

> use [of] profane, indecent or abusive language […] or engage in any form of conduct that is likely to stir up hatred against any section of the public distinguished by colour, race, religion, ethnic or national origin in a city park. (MacGuigan, 1966, p. 237)

In the subsequent trial of Beattie for breach of the bylaws around assembly and speech, the court found in favor of Beattie. Opining that the relevant section of the bylaw as *ultra vires*, that is beyond the scope of Toronto's municipal jurisdictional powers, the court noted that "the area of freedom of speech belongs to the residual powers reserved to the federal parliament under section 91 of the *British North America Act*" (qtd. in MacGuigan, 1966, p. 241) and only the Government of Canada, not the city of Toronto, had the legislative authority to deal with the issue of freedom of speech and its legal restriction.

Following these acts of provocation and anti-Semitic propaganda, the CJC transitioned their political strategy of discreet lobbying and backroom diplomacy to raising the issue regularly with its annual *démarches* to the ministers of justice (Lunny, 2017, pp. 37–39). Further abandoning their don't-rock-the-boat strategy, the CJC engaged in a public campaign for legislation that sought support beyond the Jewish community and went beyond face-to-face consultations with Cabinet ministers. As a result of this campaign, unsolicited support in the form of demands for government redress and legal reform came from across the country from church leaders, the Manitoba Bar Association, the Canadian Federation of University Women, the National Convention of the Royal Canadian Legion, and several municipal newspapers, including the *Toronto Star* (Kaplan, 1993, p. 245; Kayfetz, 1970, p. 5). In a speech to a Montreal audience in April 1964, Opposition leader and "father" of the *Canadian Bill of Rights*, John Diefenbaker called for curbs on hate literature, stating the "distribution of anti-Jewish and anti-Negro literature is of an outrageous and offensive nature that cannot be justified as an exercise of freedom of speech" (qtd. in Kaplan, 1993, pp. 245–246), a position he would later abandon at the time of legislative debate (House of Common debates, 1970, April 9, 5679).

The second significant development converging at this moment in Canadian political and legal history was Canada's engagement with civil and human rights. Canadian human rights scholars have noted that it was not until the end of World War II that a real interest in anti-discrimination law developed in Canada (Clément, 2016; Kallen, 2010; Tarnopolsky, 1979). Despite early Canadian anti-discrimination law being largely symbolic because of a lack of enforcement mechanisms and clear penalties, Clément (2016) remarks that it "signalled a momentous change" (p. 57). In Ontario, for example, a number of bills were passed that attempted to address and redress discrimination: the *Racial Discrimination Act* (1944), the *Fair Employment Practices Act* (1951), a *Female Employees Fair Renumeration Act* (1951), and the *Fair Accommodations Practices Act* (1954). Clément notes that similar laws were enacted during this period in five other provinces (p. 63). Significantly, in 1962, Ontario enacted the country's first provincial *Human Rights Code*. Applying an approach that was enforceable and remedial focused, the Code prohibited discrimination on the basis of religion, race, or ethnicity in accommodation, employment, services, and the display of signs (Clément, 2016, p. 77).

At the federal level, the 1960 *Bill of Rights* was another "landmark piece of rights legislation" (Clément, 2016, p. 65). However, characterized as "weak" (Clément, 2016, p. 65) and "a disappointment" (Sharpe, Swinton, & Roach, 2002, p. 16), this Bill only applied to federal laws and was an ordinary Act of Parliament. As such, it did not form a part of the constitution and "the mandate it conferred upon the courts was suspect" (Sharpe et al., 2002, p. 17) running counter to the basic precepts of parliamentary supremacy. Despite this, legal scholars note that "it remains in the books […] and should not be ignored, for it contains certain guarantees not found in the *Charter of Rights and Freedoms*" (Sharpe et al., 2002, p. 17). On the international front, during this post-Holocaust period, Canada signed the *Convention on the Prevention and Punishment of the Crime of Genocide* in 1948 (ratified in 1952) and signed the *International Convention on the Elimination of All Forms of Racial Discrimination* in 1966 (ratified in 1970). Ratification of these conventions legally required member states to introduce federal legislation, including criminal law, to address these human rights issues. "By the 1940s, Canada's rights culture could be defined primarily in reference to free speech, association, religion, press, assembly, and due process," remarks Clément; "[b]y the 1960s, it was clear that a culture of rights was emerging that embraced the principle of non-discrimination" (Clément, 2016, p. 85).

The third factor in terms of legislative necessity was a prescient awareness by the Cohen Committee that the very nature of communication was rapidly changing and that the liberal arguments for free speech were not faring well when matched against twentieth century mass media. Significantly, according to the Cohen Committee, it was not the number of neo-fascist and white supremacist groups operating in Canada that was the measurement of a "clear and present danger" to Canadian democracy. Rather, it was the lessons of history and the stain of the twentieth century's past that eroded the confidence of the Committee in the belief that human beings had the ability to distinguish truth from falsehood and good from evil. "In a number of ways," they wrote, "we are less confident in the 20th century that the critical faculties of individuals will be brought to bear on the speech and writing which is directed at them" (Cohen, 1966, p. 8). Differentiating between the historical rise of democratic philosophy and liberalism of the eighteenth and nineteenth centuries and the modernist age of technological revolution, they questioned the rationality of man when confronted with the allure of modern communication and media. The arguments of early liberalism – namely that the confrontation and exchange of ideas elicited man's faculty of reason and produced more enlightened ways of thinking and being in the world – were inadequate and unrealistic expectations for the modern age. "The successes of modern advertising, the triumph of impudent propaganda such as Hitler's," they noted, "have qualified sharply our belief in the rationality of man" (Cohen, 1966, p. 8).

Particularly disturbing to the committee were the seductive elements of modern technological media – color, music, and spectacle. This concern echoed the historically contemporaneous popular media theory of Marshall McLuhan that claimed "the medium is the message" (McLuhan, 1964, p. 5). In *Understanding Media*, McLuhan argued that modern "electric" technologies of the early- to

mid-twentieth century – radio, film, and television – had effects upon audiences beyond their apparent content (the texts, narratives, and stories). Attributing "rationality" to the Western cultural mindset of textuality and literacy – that is, the medium of the "uniform and continuous and sequential" – McLuhan (1964) wrote that "we have confused reason with literacy, and rationalism with a single technology. Thus, in the electric age, man seems to the conventional West to become irrational" (p. 15). The Cohen Committee feared that when matched with speech that played to human emotion, particularly to anxiety, fear, and frustration, these new forms of alluring communication – radio, television, motion pictures, mimeographic print, and telephone – would reach mass audiences and have a mass effect. Under the effects of the technological spectacle, the committee warned, "the individual is swayed and even swept away by hysterical, emotional appeals" (Cohen, 1966, p. 8).

DEBATING CANADIAN HATE SPEECH LAWS

Upon recommending the necessity for the creation of criminal hate speech laws in Canada, Maxwell Cohen (1970) noted that there were two competing sentiments being expressed:

> On the one hand, there was a new emphasis on individual freedom now generally expressed through the international symbolism of "human rights" and domestic legislation dealing with non-discrimination which seemed to emphasize increasing "rights" for the individual, his political and social opportunities for liberty of expression as well as other freedoms. On the other side, there was a growing recognition that these liberties could be dangerously abused through their use in attacking the sense of well-being and group security, through such propaganda, and thereby threatening the goodwill and cohesion within democratic society itself. (p. 105)

For Cohen, "liberty of expression" was part of a burgeoning human rights regime focused on individual rights. Although Cohen framed individual freedom and rights of the individual within an international human rights paradigm, subsequent Canadian legal scholars and jurists have structured the debate on speech somewhat differently. For example, taking a more classic approach, Kallen (1991) framed the hate propaganda debate along two conflicting positions: that of the libertarian and that of the egalitarian. The libertarian free speech argument, noted Kallen (1991), established speech and its free expression as taking "precedence over all other rights and freedoms because all rights and freedoms depend on the existence of an effective right of dissent" (p. 47). The right of citizens to speak freely and to voice criticism without fear of government reprisal, censorship, and police action are understood to be a cornerstone of liberal democratic society. As noted by Hamilton and Robinson (2019), "parliament, legislatures, a free press, and open courts, all require that citizens be free from prior restraint in expressing their diverse views on matters of public interest" (p. 7).

John Stuart Mill (1975), in his seminal essay "On Liberty" (1858), identified liberty of the press as "one of the securities against corrupt or tyrannical government" (p. 17). Embedded in the liberal idea of free expression is the Miltonian

dictum "let truth and falsehood grapple." The idea here is grounded in the belief of rationality whereby the free exchange of ideas will ultimately lead to a rejection of falsehoods and a recognition of truth. Another libertarian defence of free speech claims that free expression of hateful speech provides a societal cathartic outlet or safety valve (Rosen, 2000, p. 5). The idea behind this notion of catharsis is that minor expressions of hate speech allow strain and anger, emotions behind that speech, to find release before building to extremes of conduct and violent action. Another perspective that plays to libertarian arguments is that, in the criminalization of speech, the trial provides a dedicated platform for the accused's ideas to be disseminated more broadly (Rosen, 2000, p. 5). To some extent then the libertarian approach may be understood as a guarantee and protection of individual liberties to free expression and speech, and that the protection of these liberties creates and strengthens a democratic liberal society.

The egalitarian perspective, on the other hand, reflects principles and values enshrined in the *Canadian Charter of Rights and Freedoms* (the *Charter*). In a balancing of rights, the *Charter* guarantees the fundamental right of freedom of thought, belief, opinion, and expression, including freedom of the press and other media of communication (section 2(b)), while simultaneously setting reasonable limits prescribed by law to said rights and freedoms as can be demonstrably justified in a free and democratic society (section 1). Other significant sections of the *Charter* that express egalitarian values and principles would be found in section 27 that recognizes Canada as a multicultural society and section 7 that guarantees the right of the individual to be equal before and under the law and to have equal protection and benefit of the law. Additionally, internationally, Canada is a signatory to numerous human rights conventions that "oblige Canada to combat racism and the advocacy of genocide and racial superiority" (Rosen, 2000, p. 5). Noting the ways in which the *Charter* "shares the values of other post-war rights protecting instruments, which promote stable democratic political functioning by rejecting excesses that can deprive certain members of the community of their self-dignity and the will to take part in public life" (Weinrib, 1991, p. 1432), Lorraine Weinrib distinguishes the Canadian policy and jurisprudential stance of freedom of expression from that of the United States. As noted by Rosen, "[t]his debate reveals as acutely as possible the conflict between generally accepted Canadian multicultural and egalitarian social values and the libertarian value of freedom of expression" (p. 2).

HATE PROPAGANDA OFFENCES

After much consultation with experts and stakeholders and years of parliamentary debate, in 1970, Canada enacted three separate hate propaganda offences under the *Criminal Code*: section 318 (advocating genocide), section 319(1) (publicly inciting hatred likely to lead to a breach of the peace), and section 319(2) (willfully promoting hatred). In addition to the three offences, there are provisions which authorize the courts to order the seizure of hate propaganda, either in physical formats (section 320) or, as was added in a later amendment, in electronic

formats (section 320.1). "Hate propaganda," as used in section 320 and section 320.1, is defined by section 320(8) to mean "any writing, sign or visible representation that advocates or promotes genocide or the communication of which by any person would constitute an offence under section 319" (*Criminal Code*). More broadly, section 319(7) defines "communicating" as including communicating by telephone, broadcasting, or other audible or visible means, and defines "statements" as including words spoken or written or recorded electronically or electro-magnetically or otherwise, and gestures, signs, or other visible representation. "Identifiable groups" protected under these sections have expanded from "the public distinguished by colour, race, religion or ethnic origin" to additionally include "sexual orientation, gender identity or expression, or mental or physical disability."

There are several safeguards protecting expression built into these provisions. The first is that the impugned speech only catches public expression; private conversation, no matter how repugnant, is not subject to criminal prosecution. Second, the consent of the Attorney General, either federally or provincially, must be given to lay charges under advocating genocide, the willful promotion of hatred, and warrant of seizure. Third, currently there are four defences available to those charged under the willful promotion of hatred:

- the communicated statements are true;
- an opinion or argument was expressed in good faith and either concerned a religious subject or was based on a belief in a religious text;
- the statements were relevant to a subject of public interest and were on reasonable grounds believed to be true; and
- the statements were meant to point out matters that produce feelings of hatred toward an identifiable group and were made in good faith for the purpose of their removal.

Last, these offences have no minimum punishment. Although imprisonment is a possible sentence, so too are punishments such as conditional sentence or fine.

CRIMINAL CODE CASES

One of the most widely publicized legal cases of hate speech in Canadian history was that of Ernst Zundel, a Holocaust-denier living in Canada who was one of the principal sources of neo-Nazi literature in West Germany at the time and whose materials were banned by Canada Post (Butovsky, 1985; Russell, 1986; Weimann & Winn, 1986). In 1983, Zundel was charged with two counts of willfully publishing false news that was likely to cause injury or mischief to a public interest (section 181, formerly section 177, of the *Criminal Code*). The subject of the charges was two pamphlets: *Did Six Million Really Die? Truth at Last Exposed* and *The West, the War and Islam* (Butovsky, 1985; Kallen, 1991). The first pamphlet proclaimed that the Holocaust was a hoax; the second advanced

the anti-Semitic conspiracy theory of Zionist world domination (Kallen, 1991; Weimann & Winn, 1986). Significantly, the CJC and the B'nai Brith League wanted Zundel charged under the hate propaganda section instead. However, the then Ontario Attorney-General would not give his support for a charge that required his official approval, citing that the legislation was so poorly written that conviction was unlikely (Weimann & Winn, 1986, pp. 18–19).

In March 1985, Zundel was convicted only on the first charge regarding his public claim of Holocaust hoax and was sentenced to 15 months in jail and three years' probation, during which he was prohibited from publishing hate propaganda (Kallen, 1991, p. 51). Zundel appealed against his conviction arguing legal errors (Kallen, 1991, p. 51). Additionally, he sought leave to appeal to the Supreme Court of Canada, but his request was denied. The Ontario Court of Appeal quashed the decision against Zundel and ordered a new trial. For a second time, he was convicted by a jury in 1988 and sentenced to nine months imprisonment (Cohen-Almagor, 2006). Like a dog with a bone, Zundel launched another appeal of his case, but in February 1990 his conviction was upheld by the Supreme Court of Ontario.

In the years between Zundel's first arrest for the publication of anti-Semitic and racist materials and his second conviction under section 181, the *Charter of Rights and Freedoms* was enshrined in the Canadian *Constitution* in 1982, coming into full effect in 1985 with the section 15 equality rights. Its significance with respect to hate speech laws was that it guaranteed the fundamental right of freedom of thought, belief, opinion, and expression, including freedom of the press and other media of communication (section 2(b)), while simultaneously setting reasonable limits prescribed by law to said rights and freedoms as can be demonstrably justified in a free and democratic society (section 1). In the simplest of terms, it both guaranteed rights of expression and the right of the government to censor that expression. In 1990, Zundel petitioned a second leave to appeal to the Supreme Court of Canada. This time, in the era of the *Charter*, the leave to appeal was granted on the issue of whether or not "false news" was a violation of the constitutional guarantee to freedom of expression.

The legal outcome of Zundel's prosecution ultimately produced a judgment by the Supreme Court of Canada finding that section 181 was in violation of the guarantee to freedom of expression under section 2(b) of the *Charter* and was not constitutionally protected. This decision rendered the provision as being of no force and effect and it was repealed in 2019. The majority opinion of the Court established that section 2(b) of the Charter "protects the right of a minority to express its view, however unpopular it may be" (*R.* v. *Zundel*, p. 2). The purpose of the guarantee, explained the Court, "is to permit free expression to the end of promoting truth, political or social participation, and self-fulfillment" (p. 2) despite whether the majority regards such belief as "wrong or false" (p. 2).

In their determination, the Court opined that a limit on that freedom could not be justified under section 1 of the *Charter*. However, in an examination of the original intention of Parliament in the drafting of the false news provision, the Court found that the provision "originally focused on the prevention of

deliberate slanderous statements against nobles of the realm to preserve political harmony in the state" (p. 3). Responding to this outdated intention, the Court stated that "to suggest now that its objective is to combat hate propaganda or racism is to go beyond its history" (p. 2). Citing that the Law Reform Commission recommended repeal of the law, labeling it "anachronistic," the Court found that it lacked ostensible purpose recognizing that the Criminal Code had sections that dealt with hate propaganda more fairly and effectively (pp. 2–3). Further, in applying the *Oakes* test of proportionality to the constitutional soundness of the law, the Court found it to be "too broad and more invasive than necessary" (p. 3) to achieve the aim of promoting racial and social tolerance. "There is a danger," opined the Court, "that s.181 may have a chilling effect on minority groups or individuals, restraining them from saying what they would like for fear that they might be prosecuted" (p. 3).

As noted in the libertarian argument calling for unrestrained speech, the effect of criminalizing those who spew vitriolic and hateful speech is that the public trial afforded to them in a liberal democracy provides them with a platform to reach potentially an even larger audience for their beliefs. According to scholars who have studied Zundel's trials (Butovsky, 1985; Cohen-Almagor, 2006; Russell, 1986; Weimann & Winn, 1986), the trials drew extensive national, as well as international, media attention. Characterized as a "spectacle" (Cohen-Almagor, 2006, p. 160) and a "media event" (Weimann & Winn, 1986, p. 83), Zundel's first trial has been criticized for giving the Holocaust-denier "one million dollars of free publicity" (Weimann & Winn, 1985, p. 30).

Reflecting on the media coverage given to Zundel's first trial, Cohen-Almagor lamented that "the media cover[ed] Zundel's views without qualification, presenting him as a legitimate thinker who is offering his truth in the free market of ideas" (p. 169). Criticism fell largely on the practices and conventions of Canadian media:

> the enormous uncritical publicity which Zundel received from the trial was a direct result of the rigid conventions of journalism, which hold that the press should report only what happens in the public [and] not transmit information which might help evaluate the truthfulness of what is said. (Weimann & Winn, 1985, p. 30)

Criticizing the Canadian press for reporting "both sides as if each were credible" (p. 31), Weimann and Winn claimed that "the media reports on the trial often bent over backwards to report Zundel's perspective in an apparently dispassionate way" (p. 103; Russell, 1986).

Using the example of a *Globe and Mail* headline that read, "No gas chambers in Nazi Germany, expert witness testifies" (p. 103), Weimann and Winn argued that the headline could have read, "No gas chambers in Nazi Germany, says Nazi apologist" (p. 103). Responding to this public criticism, "the *Globe and Mail*'s city editor admitted that some of his paper's headlines were 'indefensible' and promised that it would be 'a hell of a lot more sensitive in future' on questions related to the Holocaust" (Butovsky, 1985, p. 35). Whether foremost an issue of criminalizing hateful speech or of uncritical media coverage of its legal proceedings, critics of this particular speech law viewed the Zundel trials as "a convenient

platform to mislead and to rewrite history, and confer his ideas with undeserved legitimacy" (Weimann & Winn, 1985, p. 169).

One of the first cases to test the constitutionality of section 319(2) of the *Criminal Code*, "unlawfully communicating statements, other than in private conversation, which willfully promoted hatred against an identifiable group," was that of *R.* v. *Andrews* in 1990. Significantly, *Andrews* was a companion case to *R.* v. *Keegstra*, which was heard and decided by the Supreme Court at the same time. The accused were Don Andrews, the leader of the white supremacist Nationalist Party of Canada, and Robert Wayne Smith, the secretary of the party, who were responsible for publishing and distributing the bi-monthly *Nationalist Reporter*. The police seized written material, including 89 copies of the bi-monthly plus letters written by subscribers, subscription lists, and mimeographed sticker cards containing such racist messages as "Nigger go home," "Hoax on the Holocaust," and "Hitler was right. Communism is Jewish" (*Andrews*, p. 874). The materials seized expressed white supremacist, white nationalist, and anti-Semitic ideologies calling for a "white Canada," blaming violent crime on minority immigrant populations, and promulgating anti-Semitic sentiments and Holocaust denial. At trial, the district court judge concluded that the material under examination "[c]learly … indicates not only hatred, but hatred to an unbelievable degree" (*Andrews*, p. 878).

As in *Keegstra*, the question before the Supreme Court in *Andrews* was whether section 319(2) of the *Criminal Code* violated section 2(b) of the *Charter*, and if so, whether the violation was justifiable under section 1. The reasoning of the Court mirrored that of *Keegstra*. First, section 2(b) protects all content of expression, except for expression communicated directly through physical harm. For the majority, Dickson C. J. wrote, "Hate propaganda is not analogous to violence. It conveys a meaning that is repugnant, but the repugnance stems from the content of the message and not from its form." Noting "[t]he large and liberal interpretation given to freedom of expression indicates that the preferable course is to weigh the various contextual values and factors in s.1 of the *Charter*," the Court noted that s.1 "both guarantees and limits *Charter* rights and freedoms by reference to principles fundamental in a free and democratic society" (*Keegstra*).

The Court stated that the hate propaganda law prohibiting the willful promotion of hatred was a proportional response to Parliament's valid objective of protecting target group members and fostering harmonious social relations in a community dedicated to equality and multiculturalism. Judged not to be an undue impairment of freedom of expression, the Court determined that the section did not suffer from overbreadth or vagueness. The section possessed definitional limits which acted as safeguards to ensure that it will capture only expressive activity which is openly hostile to Parliament's objective. The Court observed that the word "willfully" imports into the offence a stringent standard of *mens rea* which significantly restricts the reach of the prohibition. Moreover, the word "hatred" further reduced the scope of the prohibition to only the most severe and deeply felt form of opprobrium. The section excluded private conservation, was limited to "identifiable groups," and provided defences to the charge.

In that section 2(b) does impair expression, it does not so unduly. Conclusively, the Court held that section 319(2) violated the *Charter*, but was saved under section 1 as a reasonable limit prescribed by law as can be demonstrably justified in a free and democratic society.

OTHER LEGAL RESTRICTIONS CONCERNING HATRED

As noted in Walker's (2018) Parliamentary report on hate speech and freedom of expression, there are several other legal mechanisms that regulate hate speech in Canada (section 3.4). Regulations under the *Broadcasting Act* prohibit any licensee from broadcasting or distributing programming that contains: "any abusive comment or abusive pictorial representation that, when taken in context, tends to or is likely to expose an individual or group or class of individuals to hatred or contempt on the basis of race, national or ethnic origin, color, religion, sex, sexual orientation, age or mental or physical disability" (section 8, *Broadcasting Distribution Regulation*, SOR/97-555). These regulations apply to radio, specialty services, broadcast television and pay television, but internet-based communications do not fit the definition of "broadcasting" (Media Smarts). Canadian Press are not subject to the *Broadcasting Act*; however, they are governed by the Canadian Association of Journalist's Ethics Guidelines. Under the *Immigration Act*, customs officials have the authority to stop hate material from entering Canada, and to refuse entry to individual hate-mongers. This only applies to the movement of physical materials across the border, not sending or accessing content online, but hateful content that is purchased online and shipped to Canada may be confiscated under this legislation (Media Smarts).

THE CANADIAN HUMAN RIGHTS ACT, SECTION 13

In July of 1977, Parliament adopted and gave Royal Assent to the *CHRA*, including a provision dealing with the repeated communication of discriminatory messages by telephone (Rosen, 2000, p. 6). In the *Special Report to Parliament*, the Canadian Human Rights Commission (CHRC, 2009) acknowledged that "section 13 has always been controversial, but particularly so since it was amended in 2001 to include hate on the internet" (p. 1). With the amendment, the impugned section made it a

> discriminatory practice for anyone to communicate by telephone, by a telecommunication undertaking, or by a computer-based communication, including the Internet, any matter repeatedly that is likely to expose anyone to hatred or contempt by reason of the fact that he or she is a member of a particular identifiable group.

Section 13 did not apply to printed publications, unless a print article was posted on an Internet site (Walker, 2018, p. 9). As a subject of controversy, governmental study, reform, and repeal for years (Lunny, 2017, p. 126), section 13 was ultimately repealed in 2013.

Both the *Charter of Rights and Freedoms* and the *CHRA* are rights protecting and enshrining legal mechanisms. There are notable differences, however, that require explanation particularly with respect to hate speech. The *Charter* applies only to the actions of the state; so, for example, the *Charter* protects citizens from actions of the state as in the criminal restriction prohibiting the advocation of genocide, the public incitement of hatred likely leading to a breach of the peace, and the public willful promotion of hatred. In this example, guaranteed protections and balance are, in part, found for the defendant in section 2(b) of the *Charter*, which guarantees freedom of expression. Whereas, the *CHRA* and the provincial human rights provisions apply to other individuals, including corporations, namely, private sector actors. Here a citizen could seek remedy from some individual or some business that, for example, posted a sign denying accommodation to a racialized minority.

The *CHRA* is the principal human rights statute in the federal sector. It applies generally to federal government departments and agencies, Crown corporations and federally regulated businesses. Federally regulated businesses include federal departments, agencies, and Crown corporations; Canada Post; chartered banks; national airlines; interprovincial communications and telephone companies; interprovincial transportation companies; and other federally regulated industries, such as certain mining operations. It prohibits an employer or service provider under federal jurisdiction from carrying out discriminatory practices based on certain prohibited grounds: race, national, or ethnic origin, color, religion, age, sex (including pregnancy and childbirth), sexual orientation, gender identity or expression, marital status, family status, genetic characteristics, disability (including previous or present drug or alcohol dependence), and pardoned conviction (Walker, 2018, p. 8).

Another distinction to be made is that with respect to the restriction on speech, the *Criminal Code* is punitive; whereas

> the purpose of hate speech prohibitions in a human rights system is to rectify discriminatory practices and to compensate the victims of discrimination for the harms they have suffered as a result of that discrimination. (Hamilton & Robinson, 2019, pp. 137–138)

In criminal law, the evidential burden is to adduce sufficient evidence that the accused committed the act with the necessary intent, while the legal burden is to prove these matters beyond a reasonable doubt. Generally speaking, onus lies upon the state in criminal proceedings to prove or disprove any fact and the burden of proof is high. Whereas in human rights law, the intent to discriminate is not a necessary element; what is key is the discriminatory effect or impact of the offending act.

THE FIRST SECTION 13 COMPLAINT: TAYLOR AND THE WESTERN GUARD

The first section 13 complaint and the first case heard by Canadian Human Rights Tribunal (CHRT) was that of long-time fascist and white supremacist, John Ross Taylor, and his Western Guard party for their pre-recorded telephone

message of racist and anti-Semitic invectives (Moon, 2008, pp. 5–6). Found to have breached section 13, Taylor was ordered by the CHRT in July 1979 to cease and desist the pre-recorded phone message. Despite the injunction, the pre-recorded message remained active. In 1981, a Federal Court judge found Taylor and his Western Guard party to be in contempt and imposed a fine of $5,000 and a one-year sentence of imprisonment (Ross, 1994, p. 163). Taylor's sentence was suspended on the condition that he discontinues his discriminatory activities (Moon, 2008, p. 6). He did not, and the sentence was enforced against him. Upon his release from jail, in 1983, Taylor re-established the phone line. For a second time, the CHRC commenced contempt proceedings against him. Using his right to a *Charter* challenge, Taylor argued before the Trial Division of the Federal Court that the section violated section 2(b) of the *Charter* – the right to freedom of expression – and could not be justified under section 1, the limitations provision.

With respect to the constitutional challenge, in a four-to-three decision, the majority ruled that section 13(1) of the *CHRA* did infringe upon section 2(b) of the *Charter*, but that the infringement could be justified under section 1. The Court stated that hate propaganda presents

> a serious threat to society by undermining the dignity and self-worth of target group members and, more generally, contributes to disharmonious relations among various racial, cultural and religious groups, as a result eroding the tolerance and open-mindedness that must flourish in a multicultural society which is committed to the idea of equality.

Speaking to the issue of breadth and vagueness, Chief Justice Dickson noted that "its terms, in particular the phrase 'hatred or contempt', are sufficiently precise and narrow to limit its impact to those expressive activities which are repugnant to Parliament's objective" (p. 895).

Following the reasoning in *Keegstra*, the Court stated that the phrase "hatred or contempt" in the context of section 13(1) refers "only to unusually strong and deep-felt emotions of detestation, calumny and vilification." Dickson remarked that the absence of an intent component raised no problem of minimal impairment when one considers that the objective of the section requires an emphasis upon unintended and discriminatory effects. As in other human rights legislation, an intent to discriminate is not a precondition of a finding of discrimination. Further, even though "the section may impose a slightly broader limit upon freedom of expression than does section 319(2) of the *Criminal Code*," he wrote, "the conciliatory bent of a human rights statute renders such a limit more acceptable than would be the case with a criminal provision" (p. 928). In the dissenting opinion, Justice McLachlin argued that the infringement upon a "fundamental freedom" was deemed to be too great, particularly in that the provision targeted private speech and provided no defences and, in particular, no exemption for truthful statements. With penalty provisions being added in 1998 (para. 54(1)(c), section 54(1.1)) and the internet coming under its purview in 2001 (section 13(2)), these minority opinions would echo subsequent objections calling for the section's reform and ultimate repeal.

THE INTERNET AND HATE SPEECH

Despite the prescience of the Cohen Committee in recognizing that the very media of communication were rapidly evolving, few had the insight to recognize the little-noticed nascence of the internet of the 1960s. By the late 1980s, "the gift" (Kinsella, 2001, p. 416) of the internet allowed right-wing extremists to spread their message across borders, communicate freely and anonymously with one another, and share information on message boards and dedicated websites. According to Kinsella, "the system was unlike anything the hate movement had ever seen before in Canada" (p. 67). From early extremist sites like Marc Lemire's Freedomsite.org and Ingrid Rimland's Zundelsite to contemporary websites and online social networks, hate has managed to grow exponentially and, for the most part, to evade censorship (Hamilton & Robinson, 2019; Kinsella, 2001). Ultimately, it would be Lemire's lengthy appeal battles following a holding by the CHRT that recent amendments of section 13 adding a substantial financial sanction and including the internet within the meaning of "telephonic communications" were unconstitutional, and other cases brought before the CHRT, that lead to significant press coverage and prompted political response ending with section 13's repeal.

Two recent reports have confirmed the growing problem of online hate in Canada and the need for a more coherent and integrated response by concerned stakeholders. A study analyzing the content of right-wing extremism in Canada identified the scope of the problem by noting that Canadian users had a significant presence on over 6,600 extremist channels, pages, groups, and accounts across seven social media platforms (Davey, 2020, p. 5). Similarly, after weeks of governmental testimony and consultation with community and corporate stakeholders, the Housefather Report recommended: tracking of online hate through a national database; collecting data through government and civil society organizations; better training of law enforcement and other legal actors; creating strategies to prevent online hate through education; providing a civil remedy which may entail reinstating some version of the repealed section 13 of the CHRA; formulating a legislative definition of "hate" consistent with Supreme Court jurisprudence; and establishing requirements for online platforms and internet service providers with standards on monitoring and timely removal of posts that would constitute online hatred (Housefather, 2019, pp. 40–42).

CONCLUSION

From its legislative beginnings, Canada "hate speech" law has recognized the importance of technological advancements in mass communication and media and the need for the law to be responsive. Canadian lawmakers and jurists have also held firmly to a balancing of democratic and human rights when regulating extreme vitriolic speech of the highest opprobrium. The challenge going forward will be the government's ability to maintain this balance while responding and adapting to ever-changing platforms and technologies of global communication.

REFERENCES

Butovsky, A. (1985). The Holocaust on trial in Canada. *Patterns of Prejudice*, *19*(3), 34–36.

Canadian Charter of Rights and Freedoms, Part 1 of the Constitution Act, 1982, being Schedule B to the Canada Act 1982 (UK), 1982, c 11.

Canadian Human Rights Act, RSC 1985, c H-6.

Canadian Human Rights Commission. (2009). Special Report to Parliament: Freedom of *expression and freedom from hate in the internet age*. Ottawa: Minister of Public Works and Government.

Clément, D. (2016). *Human rights in Canada: A history*. Waterloo, ON: Wilfrid Laurier University Press.

Cohen, M. (1966). *Report to the Minister of Justice of the Special Committee on Hate Propaganda in Canada (Cohen Report)*. Ottawa, ON: Queen's Printer.

Cohen, M. (1970). The hate propaganda amendments: Reflections on a controversy. *Alberta Law Review*, *9*, 103–117.

Cohen-Almagor, R. (2006). *The scope of tolerance: Studies on the costs of free expression and freedom of the press*. London: Routledge.

Criminal Code of Canada. (1985). RSC 1985, c C-46.

Davey, J. (2020). *An online environment scan of Right-wing extremism in Canada*. London: Institute for Strategic Dialogue.

Hamilton, S., & Robinson, S. (2019). *Law's expression: Communication, law and media in Canada* (2nd ed.). Toronto, ON: LexisNexis Canada.

Housefather, A. (2019). Taking action to end online hate, report to the Standing Committee on justice and human rights. Retrieved from https://www.ourcommons.ca/Content/Committee/421/JUST/Reports/RP10581008/justrp29/justrp29-e.pdf

Johnson, D. (2019). *Hate land: A long, hard look at America's extremist heart*. Amherst: Prometheus Books.

Kallen, E. (1991). Never again: Target group responses to the debate concerning anti-hate propaganda. *Windsor Yearbook of Access to Justice*, *11*, 46–73.

Kallen, E. (2010). *Ethnicity and human rights in Canada*. Oxford: Oxford University Press.

Kaplan, W. (1993). Maxwell Cohen and the Report of the Special Committee on Hate Propaganda. In W. Kaplan & D. McRae (Eds.), *Law, policy, and international justice: Essays in honour of Maxwell Cohen* (pp. 243–274). Montreal, ON: McGill-Queen's University Press.

Kayfetz, B. (1970). The story behind Canada's new anti-hate law. *Patterns of Prejudice*, *4*(3), 5–8.

Kinsella, W. (2001). *Web of hate: Inside Canada's Far Right network*. Toronto, ON: Harper Collins.

Lunny, A. (2017). *Debating Hate crime: Language, legislatures, and the law in Canada*. Vancouver, BC: UBC Press.

MacGuigan, M. (1966). Hate control and freedom of assembly. *Saskatchewan Bar Review*, *31*, 231–250.

Mahoney, K. (2009). Hate speech, equality, and the state of Canadian law. *Wake Forest Law Review*, *44*, 321–352.

McLuhan, M. (1964). *Understanding media: The extensions of man*. New York, NY: McGraw-Hill.

Media Smarts. Online Hate and Canadian Law. Retrieved from https://mediasmarts.ca/digital-media-literacy/digital-issues/online-hate/online-hate-canadian-law

Mill, J. S. (1975). *On liberty*. New York, NY: W.W. Norton & Co.

Moon, R. (2008). *Report to the Canadian Human Rights Commission Concerning Section 13 of the Canadian Human Rights Act and the Regulation of Hate Speech on the Internet*. Ottawa, ON: Canadian Human Rights Commission.

Rosen, P. (2000). *Hate Propaganda (85-6E)*. Ottawa, ON: Library of Parliament (Parliamentary Information and Research Service).

Ross, J. (1994). Hate crimes in Canada: Growing pains with new legislation. In M. Hamm (Ed.), *Hate crime: International perspectives on causes and control* (pp. 151–172). Cincinnati, OH: Anderson.

Russell, N. (1986). Handling hate: Reporting of the Zundel and Keegstra Trials. In N. Russell (Ed.), *Trials and tribulations: An examination of news coverage given Three Prominent Canadian Trials* (pp. 29–61). Regina: University of Regina Press.

Sharpe, R., Swinton, K., & Roach, K. (2002). *The Charter of rights and freedoms* (2nd ed.). Toronto, ON: Irwin Law.

Tarnopolsky, W. (1979). The control of racial discrimination. In R. MacDonald & J. Humphrey (Eds.), *The practice of freedom: Canadian essays on human rights and fundamental freedoms* (p. 295). Toronto, ON: Butterworths.

Tulchinsky, G. (2008). *Canada's Jews: A people's journey*. Toronto, ON: University of Toronto Press.

Walker, J. (2018). *Hate speech and freedom of expression: Legal boundaries in Canada (2018-25-E)*. Ottawa, ON: Library of Parliament (Parliamentary Information and Research Service).

Weimann, G., & Winn, C. (1985). *Hate on trial: The Zundel Affair, the media, public opinion in Canada*. Oakville, ON: Mosaic Press.

Weinrib, L. (1991). Hate promotion in a free and democratic society: *R. v. Keegstra*. *McGill Law Journal*, *36*, 1416–1449.

CASES CITED

Canada (Human Rights Commission) v. *Taylor* [1990] 3 SCR 892.

Lemire v. *Canada (Human Rights Commission)* [2014] FCA 18 (CanLII).

R. v. *Andrews* [1990] 3 S.C.R. 870.

R. v. *Keegstra* [1990] 3 SCR 697.

R. v. *Zundel*, [1992] 2 SCR 731.

Zündel v. *Canada (Attorney General)* [1998] CanLII 8426.

CHAPTER 9

MEDIA LAW, ILLIBERAL DEMOCRACY AND THE COVID-19 PANDEMIC: THE CASE OF HUNGARY

Lucia Bellucci

ABSTRACT

Purpose – *This chapter aims to show how media law strongly contributed to shape in Hungary what has been pictured as a U-turn. This illiberal trend was subsequently strengthened during the Covid-19 pandemic.*

Methodology/Approach – *It considers that law also constitutes and not only orders political and social relationships. Law, including media law, has been in Hungary one of the main factors of change or rather of political-social construction. This chapter therefore moves from the study of positive law and analyzes Hungarian media laws within the theoretical framework of illiberal democracy, drawing from contributions to political science and socio-legal studies.*

Findings – *This chapter demonstrated that media laws have outlined in Hungary a centralized regulatory system with broad powers, which lacks political independence, therefore encouraging self-censorship and limiting freedom of expression and pluralism. These laws contributed to shape the illiberal U-turn occurred in the country before the pandemic, but the coronavirus offered the occasion to reinforce government powers, giving the leeway to rule with no or minimum scrutiny for an indefinite period and further limiting dissent. The analysis enabled to argue that neither the media regulation established during the past decade nor the laws adopted during the Covid-19 pandemic are compatible with a modern democracy.*

Media and Law: Between Free Speech and Censorship
Sociology of Crime, Law and Deviance, Volume 26, 151–167

ISSN: 1521-6136/doi:10.1108/S1521-613620210000026010

Originality/Value – *Based on existing literature, little research has been conducted on the appearance and endurance of non-democratic regimes, and supposedly even less within the context of the coronavirus pandemic which started only a few months ago, compared to the contributions available on democratization processes and democratic consolidation.*

Keywords: Media law; Hungary; illiberal democracy; Covid-19 pandemic; coronavirus; media regulation

INTRODUCTION

The system of checks and balances normally required of modern democracies has been drastically weakened in Hungary and the independence of all crucial institutions of the country undermined. Not only a weakening of procedures and values proper to liberal democratic policies, but also a conscious transformation of fundamental institutions of governance, both formal and informal, in order to over-centralize power has occurred (Hajnal, 2016). This follows the election of a governing coalition, constituted by the Fidesz-Hungarian Civic Union (Fidesz) and Christian Democratic People's Party (KDNP), led by Viktor Mihály Orbán, who became prime minister in 2010 (for the second time after 1998–2002) and was subsequently re-elected in 2018.

In this chapter, we argue that Hungarian media laws, openly providing for political control of the media, have been one of the main tool to enable what János Kornai has defined as a *U-turn* (Kornai, 2015a, p. 1, 2015b, p. 279) with regard to the democratic institutions and the rule of law. Media laws crucially contributed to a democratic backsliding (Sedelmeier, 2014), which indicates an erosion, by the state, of the democratic institutions that constitute the backbone of a pre-existing democracy. We also argue that this phenomenon occurred before the Covid-19 pandemic. The need to manage the emergency caused by the pandemic has then created the conditions to worsen it. Our analysis aims to show that the distinctive features of this system encourage self-censorship and limit freedom of expression and pluralism. Considering that independent media are one of the key features of modern democracies, our analysis will also, more broadly, enable us to address the question whether the Hungarian self-proclaimed illiberal democracy can still be considered a democracy.

Hungary is a case of self-defined, self-proclaimed illiberal democracy. Even though this expression was subsequently little mentioned by official Hungarian sources (see Pap, 2017b, p. 62), Prime Minister Orbán affirmed in 2014 the need to

> break with liberal principles and methods of social organization, and in general with the liberal understanding of society [and that] the new state [under construction] in Hungary is an illiberal state, a non-liberal state. (Orbán, 2014)

He also argued in the same year that:

> Hungarians welcomed illiberal democracy […]. In the Hungarian context, the word liberal has become negative. Liberal democracy has no or very little support in Hungary. [I]t's not true that a democracy can only be liberal. (Simon, 2014)

We will therefore analyze media laws in Hungary in light of two opposite thesis developed within the academic literature on illiberal democracy, with the aim to contribute to the literature on the appearance and endurance of non-democratic regimes, focusing on their developments in times of pandemic. "[In fact] remarkably little research has been undertaken on the emergence or persistence of non-democratic regimes" (Levitsky & Way, 2002, p. 63), compared to the contributions available on democratization processes and democratic consolidation; this all the more so, one imagines, in the context of the Covid-19 pandemic, given that the pandemic began only a few months ago.

Some scholars, including Tamás Csillag and Iván Szelényi, believe that illiberalism creates the conditions for the elimination of democracy, but without necessarily eliminating it, arguing that "the road from democracy to autocracy is paved with the 'stones' of illiberalism" (Csillag & Szelényi, 2015, p. 27). In this sense the word "illiberal" should be understood as referring to a "constitutional and political condition that creates a unique middle ground between a constitutional democracy and autocracy" (Pap, 2017a, p. 162; see also Pap, 2017b, p. 49).

Other authors, such as Steven Levitsky and Lucan Way, argue instead that in some cases "the collapse of one kind of authoritarianism yielded not democracy but a new form of nondemocratic rule" (Levitsky & Way, 2002, p. 63); in the absence of the essential elements of a democracy, it is necessary to refer to non-democratic "regimes" (Levitsky & Way, 2002, p. 63) and, more generally, to non-democratic "actors" (Levitsky & Way, 2002, p. 54). In what the authors identify as "competitive authoritarian regimes" "formal democratic institutions are widely viewed as the principle means of obtaining and exercising political authority" (Levitsky & Way, 2002, p. 52). Although some scholars referred to a "partial or diminished democracy [Levitsky and Way] agree with Juan Linz that they may be better described as a (diminished) form of authoritarianism" (Levitsky & Way, 2002, p. 52; see also Linz, 2000, p. 34).

This chapter attempts to show how law, including media law, constitutes and not only orders political and social relationships (here we draw, in slightly modified form, from the statement contained in Whittington, Kelemen, & Caldeira, 2008, p. 12). In some cases, it not only consolidates a situation but it also becomes its agent (Bellucci, 2018, p. 8). In Hungary, the law is an instrument that allows an illiberal path to remain within a formally legal framework. Media laws in particular "are one of the main tools used to modify the relational binomial state-citizen/law-society, one of the main factors of change or rather of political-social construction" (Bellucci, 2018, p. 9, my translation). Therefore, this chapter moves from the study of positive law. It focuses on a few rules within the rich Hungarian media regulation which greatly contributed to shape a centralized system placed under the control of the political power and shows that the rules adopted during the Covid-19 pandemic have reinforced this control.

This chapter develops its arguments as follows. The first section shows how the National Media and Infocommunications Authority/NMHH (Media Authority) and Media Council have replaced previously existing independent bodies. The second section focuses on the extremely broad range of powers owned by the media regulatory bodies and discusses the case of the independent radio Klubrádió, in order to explain the consequences of the established arbitrary power within

licensing. The third section explores the regulation of the Hungarian public broadcasting, focusing on the centralization of news production. The fourth and final section analyzes how the government powers increased during the Covid-19 pandemic, in particular since the government adopted emergency states and a measure that further limits independent journalism.

THE LACK OF INDEPENDENCE OF THE "NEW" MEDIA INSTITUTIONS IN HUNGARY: A "DIARCHY WITH ONE KING"

This section analyzes media laws in Hungary and the criticism they have sparked, in order to show that these laws have shaped a centralized regulatory system with broad powers and placed it under government control. In Hungary, "new" media institutions have been established. The main institution for media and telecommunication regulation, the Media Authority, was established by the 2010 Act on the modification of certain acts regulating the media and telecommunications (Lengyel, 2010).

According to the original Act CLXXXV of 2010 on Media Services and Mass Media (Media Act), of 31 December (Arts 111(3) and 125(5)), the Media Authority President is appointed by the prime minister for a nine-year term, indefinitely renewable. The President appoints the Media Authority's top managers and exercises the employer's rights over them, including the possibility to terminate their mandate, without being obliged to justify his or her conduct (Arts. 111(2)(c)–(e), 112(1), 113(6), 115(7) and 117(1) of the Media Act).

Within the Media Authority, a Media Council (Art. 109(3) of the Media Act) is established in order to supervise the implementation of the "new" norms on mass media. Its aim of guaranteeing freedom of expression, particularly in light of the Media Act and the Act CIV of 2010 on the Freedom of the Press and the Fundamental Rules on Media Content (Press Freedom Act), seems more declared than real.

When the President of the Authority is appointed, she or he also becomes automatically, *ipso iure*, the candidate for the Media Council's chairperson, whose appointment is subject to parliamentary approval (Art. 125(1) and (3) of the Media Act). The new powers therefore converge on the President.

The mandate of the Media Council President is of the same duration of that of the President of the Authority (Art. 125(7) of the Media Act). The remaining members of the Council are nominated by an ad hoc committee, which is composed of delegates from each parliamentary faction (Art. 124(3) (b) of the Media Act). Votes in this committee are "weighted" in light of the proportion of each faction's representation in Parliament (Art. 124(4) of the Media Act). In practice, the governing coalition selects the remaining members of the Council, using its two-thirds majority within the committee (Brouillette & van Beek, 2012, pp. 2–3; NMHH, 2010). This is the majority that governing coalition (Fidesz-KDNP) holds in Parliament. This two-thirds majority was large enough to enable the government coalition to reshape the

crucial institutions of the country, undermining its checks and balances, and adopt a new constitution, The Basic Law/Fundamental Law of Hungary (The Fundamental Law of Hungary, 2011; see on this topic Bánkuti, Halmai, & Scheppele, 2015, p. 38; Scheppele, 2015, pp. 20 & 43). Reforming media regulation was one of the crucial steps toward the building of a new constitutional order (Bánkuti et al., 2015, p. 40).

The Hungarian system is based upon the duopoly of the Media Authority and the Media Council, which one can describe as a "dual monarchy" (Haraszti, 2011, II para, pt. 2), borrowing an expression used – albeit in a slightly different context – by Miklós Haraszti, former OSCE Representative on Freedom of the Media (the second representative), or a "diarchy with one king" (Bellucci, 2018, p. 64, the translation is mine), under the leadership of a single person appointed by the prime minister. It is without precedent in democratic Europe. In 2011, the Media Act was amended to reinforce the Hungarian media diarchy. Consequently, the mandate of the Media Authority's President (and members of the Media Council) lasts until the new members and President are elected by a two-thirds parliamentary majority at the end of their nine-year terms. If Parliament fails to elect a new President and members at the end of this term, the President and Media Council members in office retain their positions for an indefinite time, until new members are elected (Art. 216(8) of the Media Act).

Although the Media Act formally recognizes the autonomy and independence of the Media Authority and Media Council, media regulation and the related praxis have sparked criticism both in Hungary and internationally. The European Union (EU) has criticized the extension of the competence of the Media Authority and has also expressed concerns about its independence (Kroes 2010; Kroes 2011). The OSCE Representative on Freedom of the Media (Jakubowicz, 2010; Nyman-Metcalf, 2011) and the Commissioner for Human Rights of the Council of Europe (Commissioner for Human Rights, 2011) underlined that Hungarian media regulation lacks pluralism, independence, and impartiality toward political powers. The Commissioner has recommended that the Hungarian government adopt the Council of Europe's recommendations to ensure the independence of media regulatory bodies.

In 2013, amendments to the Media Act were adopted as a result of the pressures from the Secretary General of the Council of Europe on Hungarian authorities (Council of Europe, 2013, paras. 146 &148; Art. 111 A(1) & (3) of the Media Act, consolidated version, 2015). Under these amendments, the President of the Media Authority is appointed by the President of Hungary upon proposal of the prime minister, and his or her appointment is limited to a non-renewable, nine-year term (Council of Europe, 2013, paras. 144 and 148). However, these improvements do not change the structure of Hungarian media's regulatory system (European Parliament, 2012, paras. 38–43) and do not guarantee its independence from political control: the minister of justice's response of 19 December 2012 to a letter of 29 November 2012 from the Secretary General of the Council of Europe clarified that Hungarian authorities soundly rejected the proposal to really separate the Media Council from the Media Authority (Council of Europe, 2013, para. 146).

A CENTRALIZED SYSTEM WITH BROAD COMPETENCES WITHIN ALL MEDIA: ARBITRARY POWER AND THE CASE OF KLUBRÁDIÓ

Hungarian media laws have shaped a centralized system of control with very broad competencies. The Media Authority is responsible for overseeing the sectors of media, telecommunications, and postal services. It has replaced the former agencies regulating media and telecommunications, the National Radio and Television Commission (ORTT) and the National Communications Authority, respectively. It consists of three main entities: the President, the Office, and the Media Council (Brouillette & van Beek, 2012, p. 27). The Authority's powers are specified in the Press Freedom Act and the Media Act, as well as in a series of rules, including the so-called media law "package" (for further details on the rules that shape this "package" see Brouillette & van Beek, 2012, p. 27, note 3).

Critics fear that Hungarian rules will encourage self-censorship (Organization for Security and Co-operation in Europe (OSCE), 2010). This result is even more likely because of the vague language used in Hungary's media laws – language that could have a "chilling effect" on the press. For example, the unusual and broad definition of a "press product" includes not only the printed press, but also online news portals and professional blogs that are managed by commercial enterprises (Art. 1(6) of the Press Freedom Act). Furthermore, Article 17, paragraph 2 of the Press Freedom Act, which has been changed several times, has been recognized for its lack of clarity. By prohibiting content that may offend a "minority" or "any majority," it could limit critical coverage of all groups, undermining the watchdog role of the media in a democracy (Bellucci, 2018, pp. 67–68; The Hungarian Civil Liberties Union (HCLU-TASZ), 2011).

Prior to the Audiovisual Media Services Directive (AVMSD) of 2010 (European Parliament and Council, 2010), which was later amended, the print and online press were bound to provisions of both the Penal and Civil Codes but widely self-regulated. After 2010, the Press Freedom Act and the Media Act, replaced the 1996 Act on Radio and Television Broadcasting, and transposed this directive. Regulations concerning all media, including the print and online press were introduced and criticized at the international level. The Hungarian government, that had justified through the same argument the centralization of competencies within the Media Authority throughout the entire media sector (NMHH, 2010), argued that these regulations were adopted to reply to the demands of media convergence, which had resulted from the digital revolution and that they comply with AVMSD (Brouillette & van Beek, 2012, pp. 30 & 43).

In the EU, the implementation of AVMSD has expanded media regulatory bodies' powers and substituted the traditional approach based on a specific sector with a "technologically neutral" media regulatory model. Nevertheless, the scope of the Hungarian Authority's competences and the extension of its powers are without legal and historical precedent among European media regulatory bodies, since in Hungary a single central authority regulates all media sectors (Brouillette & van Beek, 2012, pp. XV & 11).

László Majtényi, in his capacity as the Director of the Eötvös Károli Institute (founded in 2003 by the Soros Foundation) and former President of the ORTT, has expressed his concerns to the European Parliament, specifically about rules that regulate different media with the same, standardized criteria. He argued that the Media Authority's competence over the press and the new media, including all online publications and a significant portion of blogs, is in violation of most accepted liberal and democratic principles, which are recognized by the jurisprudence both of the European Court of Justice and the Hungarian Constitutional Court (Majtényi, 2011, para. IV, pt. 3).

Despite the fact that the Authority has *de facto* limited its power of control over both print journalism and internet (Article 19 & HCLU-TASZ, 2011; Kovacs, 2012), Dunja Mijatovic has argued in her capacity as the OSCE Representative on Freedom of the Media (the third representative), that this attempt at control results in forms of self-censorship from these media (OSCE, 2010).

The decision delivered by the Constitutional Court on 19 December 2011 (n. 165/2011. *(XII. 20.) AB*, for its summary see HCLU-TASZ, year not mentioned but presumably 2011; Kelemen, 2012; Koltay, 2012; Mérték Médiaelemző Műhely, 2012) partially declared Article 2 of the Press Freedom Act as unconstitutional. More specifically, it annulled the words "and published media." It therefore excluded "print and online press from the scope of application of the Act" (Kelemen, 2012).

The Hungarian regulation concentrates, within the Media Authority, an extremely important array of powers, on a scale that goes from tendering and licensing to issuing sanctions, over broadcasters, public and commercial, digital, and analog. Critics highlight that the Authority controls broadcasters through an allegedly "arbitrary" power of licensing without serious tenders. The law bans broadcasters that have been sanctioned for a "serious breach of obligations" (Art. 55(1)(c) of the Media Act) from participating in future tender offers for a period of five years. The content of the quoted expression is defined by the Media Council. This provision confirms the vagueness of Hungarian media law, and critics point out that it could drive broadcasters to acts of self-censorship in order not to displease the Media Council (Article 19 & HCLU-TASZ, 2011, p. 4).

The case of *Klubrádió*, a popular and independent radio station devoted to political news and debates whose voice is widely acknowledged as having influence on Hungarian public opinion, is representative of the Media Council's power. A Council decision, no. 406/2013 (III. 13.), declared on March 2013 *Klubrádió* the winner of a tender process (Council of the National Media & Infocommunications Authority, 2013, pp. 40–41), after an extenuating judicial journey (see also Gall, 2012). This journey put *Klubrádió* in an extremely difficult legal and financial situation, forcing it to ask for donations from the listeners, who donated in order to allow this radio station to keep going (Article 19, 2011).

In particular, the Media Council invited tenders for an exclusively music radio station precisely when *Klubrádió*'s license was due to be renewed and, in December 2011, it assigned *Klubrádió*'s frequency to an unknown broadcaster called *Autórádió* (Brouillette & van Beek, 2012, p. 92). *Klubrádió* appealed this decision and the Budapest Court of Appeal ordered the Media Council to

review it, because the call for tenders had not complied with current regulations. However, instead of issuing a licence to *Klubrádió*, the Media Council decided to re-examine all tenders it had received (BBC, 2012) and excluded *Klubrádió* from the call for tenders for formal reasons (AFP, 2012). *Klubrádió* had to appeal the Media Council's decision and despite a new judgment in its favor that was filed by the Budapest Court of Appeal, a Media Council decision excluded the broadcaster from the call for tenders for formal reasons. *Klubrádió* appealed even this decision and a new and final judgment in favor of this radio station was filed. It declared that since the Media Council had previously considered the offer valid, it could not later affirm the opposite based on formal aspects.

The result of the judicial battle seemed to have secured Klubrádió at list for a few years. Nevertheless, after the lengthy court battles of the 2010s, *Klubrádió* is still in danger. In September 2020 the Media Council announced that it would revoke the station's rights on frequency 92.2 MHz as of February 2021 for breaches of Hungarian media law and announced a bid open for this frequency for the same month (Gall, 2020; New Europeans.net, 2020).

HUNGARIAN PUBLIC MEDIA AND POLITICAL CONTROL: THE CENTRALIZATION OF NEWS PRODUCTION

Even the regulation of the Hungarian Public Broadcasting Service was modified to guarantee an extremely centralized system, including for what concerns news production. The amendment of 6 July 2010 to Article 61 of the Hungarian Constitution has broadened the definition of the role of public broadcasting to include promoting national and European identity, reinforcing "national cohesion [and satisfying] the needs of national, ethnic, familial and religious communities" (Brouillette & van Beek, 2012, p. 66, text and note 1). Its main aim was to remove the preclusion concerning information monopolies (Haraszti, 2011, II para., pt. 2), replacing the provision that obliged Parliament to pass a law aimed at preventing them with the citizen's right to be provided with "proper" or "adequate" information about public life (Commissioner for Human Rights, 2011, para. 2.1.).

Even the "new" system of governance and news production for Hungarian public service media, introduced by the Act LXXXII of 2010 and the Media Act (Brouillette & van Beek, 2012, pp. 66–67), was placed under the Media Council's control. Formerly independent bodies in charge of supervising public service broadcasting activity have been conflated into a single body, the Public Service Foundation, which is managed by a board of trustees. Six of its members are appointed by the Parliament, while its chairperson and one other member are delegated by the Media Council for a nine-year term. The board of trustees elects the CEO of the public media service provider, determines the terms and conditions of his or her employment contract and remuneration, and may as well put an end to his [or her] employment relationship (Arts. 85(1), 86(1), 86(6), 90(1)(e)–(f) of the Media Act, consolidate version, 2015).

One can read in the former version of the Media Act that the Public Service Foundation oversees the National News Agency (MTI) (see, in particular, Art. 101), that MTI has the exclusive right to produce news programs for Hungarian public broadcasters and it is also in charge of the online news portals and public media, as well as their on-demand media services (Art. 101(4) of the Media Act).

Criticism has been raised over the centralization of news production within the MTI. As the Commissioner for Human Rights has highlighted,

> all public service broadcasting newsmakers are made the employees of a Fund set up under the Media Council […]. These rules make the Head of the Media Council the indirect employer of all journalists of all public service broadcasting. (Commissioner for Human Rights, 2011, para. 2.3, pt. 44)

As the Commissioner has argued, this is inconsistent with standards of the Council of Europe, which are designed to preserve the independence – especially editorial – of public service broadcasting. News production centralization processes have not been unknown in other European countries, and media experts have raised criticism toward them, expressing concerns with regard to media independence from existing political powers (Bellezza & Pollicino, 2012, p. 86) and the protection of a pluralistic public interest programing (Dencik, 2012, p. 87), but the centralization of the Hungarian system overcomes that of any other European country.

More generally, the power of the Media Council (and therefore of the government) over public service broadcasters has prompted strong criticism, in particular with regard to the extent of the Media Council's control over an extremely wide range of activities, from the appointment of broadcasters' directors to the management of funding. The European Parliament has referred to a suppression of "the political and financial independence of public-service media" (European Parliament, 2011) in Hungary, while the Commissioner for Human Rights of the Council of Europe has underscored the risks for the independence of public service broadcasting (Commissioner for Human Rights, 2011). Nevertheless, in 2015 a concentration process took place. As shown in the consolidated version of the Media Act of the same year (Art. 84(1)), the state media companies, including MTI, merged to form a large media outlet, Dunia Media.

THE "STATE OF DANGER," "STATE OF MEDICAL CRISIS," AND A MEASURE LIMITING INDEPENDENT JOURNALISM: REINFORCING GOVERNMENT POWERS DURING THE COVID-19 PANDEMIC

During the ongoing Covid-19 pandemic, measures were taken that increased government powers to an extent that potentially affects the regular functioning of Parliament and limits independent journalism. When the Covid-19 spread, political power and media governance in Hungary were already extremely concentrated and the pandemic has been used as a justification for further centralization and limitation of media independency and freedom.

Because of a new measure, anyone who "distorts" or publishes "false" information on the pandemic can potentially be punished with five years in jail. This provision, written in a vague language like many other media regulations, is likely to justify further control over the media and is a source of concern for independent journalists (Editorial du "Monde", 2020; Walker, 2020c; Walker & Rankin, 2020b). To date, no case has been brought before a court with reference to this rule, but the police have launched more than one hundred investigations and some citizens have been questioned following critical posts published on Facebook (Walker, 2020d; about the control over posts on Facebook see also Walker & Rankin, 2020b). This opened the way to effective control over social media (Bellucci, 2021).

The measure in question was introduced by a law aiming at protecting against coronavirus that, since 2020, March 30, extends the government's emergency powers, suspending the elections and providing the possibility for the government to rule by decree without being bound by the existing laws (International Center for Not For Profit Law (ICNL, n. d.); Kelemen, 2020), during the emergency period. This period is not linked to a precise date and therefore potentially covers an indefinite amount of time. To counter the consequences of the pandemic, the government declared a "state of danger," that is a special state of emergency, which has been introduced on 2020, March 11 (for the chronology of the government decision and measures adopted during the pandemic see ICNL, n. d.; Kelemen, 2020) and is regulated by Article 53 of the Hungarian Constitution.

According to what had been announced by the Hungarian Minister of Justice, Judit Varga (Walker, 2020d), the Hungarian Parliament has adopted a law that, since 2020, June 18, requires the government to terminate the state of danger and provides for the revocation of the extraordinary powers granted by the law aimed at protecting against coronavirus (ICNL, n. d.). Nevertheless, the Parliament has also introduced since June 17, through a law that provided rules on the revocation of the state of danger, a possible other state of emergency, a "state of medical crisis" (ICNL, n. d.), which can be declared unilaterally by the government for a period of six months, renewable by decree; it authorizes the government to suspend existing laws and fundamental rights and to decide by decrees, thus excluding the Parliament (ICNL, n. d.; Thorp, 2020).

Considering also that emergency measures related to the migration crisis, brought in 2016, are still in force (Walker & Rankin, 2020a), the concern had been expressed that even after the revocation of government powers related to the health emergency, the more than 100 decrees adopted since April (or part of them), some of which do not seem to be related to the fight against Covid-19, would remain in force (Walker, 2020d). The rules introduced in the Hungarian legal system since June 17 reinforced this concern. It was noted that the legislative intervention of the Hungarian Parliament does not intend to reinstate the legal framework prior to the coronavirus pandemic, but on the contrary to create "a legal basis for the use of more new, extraordinary and unlimited governmental powers" (Novak, 2020). Although Article 53 of the Hungarian Constitution provides in its paragraph 4 that upon the termination of the state of danger the

decrees adopted by the government during that state shall cease to have effect, the concern of many is that this will not happen (Bellucci, 2021).

A relaxation of control over media and journalists is therefore not to be expected. Most likely, the measures that create the conditions for a tightening of control over the media will also remain in place. The situation therefore poses a real threat to freedom of expression, it adds to the risks for this freedom arising from the fact that the Press Freedom Act (as more generally the other regulations affecting Hungarian media) does not yet contain precise definitions.

Article 6 of the Press Freedom Act was at one point alarming in this regard. It originally provided that:

> [In] exceptionally justified cases, courts or authorities may – in the interest of protecting national security and public order or uncovering or preventing criminal acts – require the media service provider and any person employed by or engaged, in any other legal relationship intended for the performance of work, with the media content provider to reveal the identity of the informant. (para. 3)

Given its lack of precise definitions, this article menaced the protection of sources. The dangers are evident if we consider that the protection of sources is at the root of investigative journalism, allowing media to act as the "watchdog of democracy," and its lack risks promoting self-censorship. The European Court of Human Rights has recognized a journalist's right not to disclose his or her sources, as evidenced in the case of *Roemen and Schmit* v. *Luxembourg*, of 25 February 2003 (European Court of Human Rights, 2003).

The matter of protecting sources is one of the issues that was considered by the aforementioned Constitutional Court's decision of 19 December 2011. The Court "annulled the provisions that would have obliged journalists to reveal their sources in legal proceedings" (Council of Europe, 2013, para. 137). With regard to sources of information, it precisely pointed out two omissions by the legislator: the lack of procedural guarantees for the protection of the sources of information in legal proceedings as well as of proper respect for the sources of information and the duty of confidentiality of lawyers with regard to the duty of the media service providers to furnish data to the Media Authority (Kelemen, 2012).

Article 6 was therefore amended. Paragraph 2 of this article introduced the goal of investigating a crime without making reference to the national security and the public order, but despite the fact that the possibility of ordering source disclosure pertaining to criminal investigation has been limited, Hungarian law still lacks precise definitions. This, and, more generally, the vagueness of the language used in Hungarian media laws, is certain to have a "chilling effect" on investigative journalism (Bellucci, 2018, pp. 102–104; see also the aforementioned considerations about the definition of a "press product" and Article 17, paragraph 2 of the Press Freedom Act).

CONCLUSION

Media laws are part of a larger picture that they also contribute to outline (Bellucci, 2018, p. 159), which includes, for example, the restriction of citizens'

initiative, social rights and civil liberties, and more generally the "emptying" of the checks and balances of modern democracies (Bozóki, 2015, p. 3) previously introduced in Hungary.

Law has been used in Hungary to keep an illiberal trend within the framework of formal legality. Nevertheless, the language used by Hungarian media law is vague and indeterminate; it does not clearly define potential violations (see also Article 19 & HCLU-TASZ, 2011). "It is also populist, sometimes characterized by a paternalistic approach" (Bellucci, 2018, p. 102, the translation is mine).

Furthermore, the Hungarian media regulation is built around the Media Authority and the Media Council. The duality of this leadership under a single person appointed by the prime minister is a unicum in Europe and has raised strong criticism by the international community. Even though we could not cover the enormous investigative and sanctioning powers of the authority, our analysis showed that the Hungarian regulation introduced an extremely centralized system under the leadership of institutions controlled by the governing power. The regulation does not protect independent media from political pressures, as a result these media frequently succumb or risk to succumb to difficult economic conditions. As the Klubrádió case showed, the costs incurred for judicial battles against media institutions may cause these conditions.

The media, as well as more generally the institutions crucial for the organization of any political opposition, do not have the necessary freedom to encourage democratic discussion and their regulation also ensures that the governing coalition maintains control over the Hungarian institutions regardless of any possible changes following future elections. All this while layoffs of independent journalists continue to occur, in line with a previous major reorganization of the public service media that implied the removal of a third of its journalists (Brouillette & van Beek, 2012, p. 68). During the pandemic, for example, in particular on July 22, 2020, Szabolcs Dull, the editor-in-chief of Hungary's biggest independent news website, Index (Human Rights Watch, 2020), was fired a month after he had openly raised alarm over political interference in its operation. An open letter signed in support of Dull by more than 90 journalists of the outlet (Walker, 2020a) and the resignation of more than 70 members of its staff journalists followed this dismissal (Walker, 2020b). Hungarian media can no longer play their role of watchdogs of democracy. Even if one wants to consider democracy only as the effective possibility of holding free and impartial multiparty elections on the basis of certain pre-existing conditions (see Zakaria, 1997, pp. 24–25), "including a certain freedom of expression and, therefore, independence of the media, there would not seem to be any conditions in Hungary" (Bellucci, 2018, p. 160, the translation is mine).

Not only, during the pandemic, the possibility for the media to carry out independent activity was further limited through a new measure, but the pandemic provided an opportunity to introduce a legal base that allows the government to declare an emergency state, thus acquiring extraordinary and illimited powers and excluding the Parliament. Once again, the normative framework is formally legal but profoundly illiberal and disrespectful of democratic values.

This chapter demonstrated that media laws have traced in Hungary a centralized regulatory system with broad powers under government control. The distinctive features of this system encourage self-censorship and compromise freedom of expression and pluralism. These laws strongly contributed to shape the illiberal U-turn occurred in the country. During the Covid-19 pandemic the Hungarian illiberal trend was strengthen.

The pandemic offered the occasion to reinforce government powers: it introduced emergency states and a measure that harnesses independent journalism, giving the leeway to rule with no or minimum scrutiny for an unspecified amount of time and further limiting dissent. It provided the opportunity to rise the exceptionality linked to the moment to a potentially perennial situation (Bellucci, 2021), so that Hungary has been defined as a "a coronavirus autocracy" (Kelemen, 2020). The analysis enabled to affirm that neither the media regulation established during the past decade nor the laws adopted during the Covid-19 pandemic are compatible with a modern democracy. Levitsky and Way's (2002) thesis that considers "a new form of nondemocratic rule" (p. 63) and that, in the absence of the essential elements of a democracy (i.e., even in the absence of independent media) we should refer to non-democratic "actors" (Levitsky & Way, 2002, p. 54), seems more appropriate to describe the essence of the Hungarian U-turn.

REFERENCES

Act CIV of 2010 on the Freedom of the Press and the Fundamental Rules on Media Content (Press Freedom Act). Retrieved from http://mediamonitor.ceu.hu/archive-item/?id=68 (official translations of the original version of November 9, 2010 and amended versions that came into force on August 3, 2011 and July 3, 2012). Accessed on September 26, 2020.

Act CLXXXV of 2010 on Media Services and Mass Media (Media Act), of December 31, 2010 (official translations of the original version of December 31 and amended versions that came into force on August 3, 2011 and July 3, 2012. Retrieved from http://mediamonitor.ceu.hu/archive-item/?id=69. Accessed on September 26, 2020. A consolidated version, as effective from July 1, 2015 can be retrieved from http://hunmedialaw.org/dokumentum/153/Mttv_110803_EN_final.pdf. Accessed on September 26, 2020.

AFP. (2012, July 5). Hongrie: la radio d'opposition Klubradio perd sa fréquence. *L'Express*. Retrieved from www.lexpress.fr/actualites/1/economie/hongrie-la-radio-d'-opposition-klubradio-perd-sa-frequence_1135049.html. Accessed on September 26, 2020.

Article 19. (2011, December 21). Hungary: ARTICLE 19 condemns Media Authority refusal of license to Klubrádió. *Refworld*, Press release. Retrieved from https://www.refworld.org/pdfid/4f266fd92.pdf. Accessed on September 26, 2020.

Article 19 & Hungarian Civil Liberties Union (HCLU-TASZ). (2011, August). Hungarian Media Laws Q&A, Country Report. *Article 19*, London. Retrieved from https://www.article19.org/data/files/medialibrary/2714/11-09-01-REPORT-hungary.pdf. Accessed on September 26, 2020.

Bánkuti, M., Halmai, G., & Scheppele, K. L. (2015). Hungary's illiberal turn: Disabling the constitution. In P. Krasztev & J. Van Til (Eds.), *The Hungarian patient* (pp. 37–46). New York, NY: Central European University Press.

BBC. (2012, March 14). Hungary Orban: Court boost for independent radio. *BBC News Europe*. Retrieved from http://bbc.co.uk/news/world-europe-17367684. Accessed on September 26, 2020.

Bellezza, M., & Pollicino, O. (2012). Expert assessment. In A. Brouillette & J. van Beek (Eds.), *Hungarian Media Laws in Europe: An assessment of the consistency of Hungary's Media Laws with European practices and norms*. Budapest: Center for Media & Communication Studies-CMCS, Central European University. Retrieved from http://homepage.univie.ac.at/katharine.sarikakis/

wp-content/uploads/2012/09/Hungarian_Media_Laws_in_Europe.pdf. Accessed on September 26, 2020.

Bellucci, L. (2018). *La sindrome ungherese in Europa. Media, diritto e democrazia in un'analisi di Law and Politics*. Milano: Giuffrè.

Bellucci, L. (2021). La "democrazia illiberale" in Ungheria tra diritto, narrative politiche e COVID-19: Elementi di riflessione sulla resilienza delle democrazie in tempi di pandemia. *Diritto & Questioni Pubbliche*. Submitted.

Bozóki, A. (2015). Broken democracy, predatory state and national populism. In P. Krasztev & J. Van Til (Eds.), *The Hungarian patient: Social opposition to an illiberal democracy* (pp. 3–36). Budapest: Central European University Press.

Brouillette, A., & van Beek, J. (Eds.). (2012). *Hungarian Media Laws in Europe: An assessment of the consistency of Hungary's Media Laws with European practices and norms*. Budapest: Center for Media & Communication Studies (CMCS), Central European University. Retrieved from http://homepage.univie.ac.at/katharine.sarikakis/wp-content/uploads/2012/09/Hungarian_Media_Laws_in_Europe.pdf. Accessed on September 26, 2020.

Commissioner for Human Rights. (2011, February 25). Opinion of the Commissioner for Human Rights on Hungary's media legislation in light of Council of Europe standards on freedom of the media, CommDH(2011)10, Strasbourg. Retrieved from https://rm.coe.int/16806daac3. Accessed on September 26, 2020.

Council of Europe/Conseil de l'Europe. (2013, June 10). Parliamentary Assembly/Assemblée parlementaire, Committee on the Honoring of Obligations and Commitments by Member States of the Council of Europe (Monitoring Committee), K. Lundgren & J. Fischerová, Request for the opening of a monitoring procedure in respect of Hungary, Report, Doc. 13229, Explanatory memorandum. Retrieved from http://assembly.coe.int/nw/xml/XRef/Xref-DocDetails-en.asp?FileId=19777. Accessed on September 26, 2020.

Council of the National Media and Infocommunications Authority. (2013, May 10). The Klubrádió case: The tender procedures related to the Budapest 95.3 and 92.9 MHz frequencies. Retrieved from http://hunmedialaw.org/dokumentum/96/The_Klubradio_case.pdf. Accessed on September 26, 2020.

Csillag, T., & Szelényi, I. (2015). Drifting from liberal democracy: Traditionalist/neo-conservative ideology of managed illiberal democratic capitalism in Post-communist Europe. *Intersections. East European Journal of Society and Politics*, *1*(1), 18–48. Retrieved from https://intersections.tk.mta.hu/index.php/intersections/article/view/28. Accessed on September 26, 2020.

Dencik, L. (2012). Expert assessment. In A. Brouillette & J. van Beek (Eds.), *Hungarian Media Laws in Europe: An assessment of the consistency of Hungary's Media Laws with European practices and norms*. Budapest: Center for Media & Communication Studies (CMCS), Central European University. Retrieved from http://homepage.univie.ac.at/katharine.sarikakis/wp-content/uploads/2012/09/Hungarian_Media_Laws_in_Europe.pdf. Accessed on September 26, 2020.

Editorial du "Monde." (2020, April 1). Viktor Orban à l'ombre du coronavirus. *Le Monde*. Retrieved from https://www.lemonde.fr/idees/article/2020/04/01/viktor-orban-a-l-ombre-du-coronavirus_6035162_3232.html. Accessed on September 26, 2020.

European Court of Human Rights. (2003, February 25). *Roemen and Schmit v. Luxembourg*, Application no. 51772/99. Retrieved from http://hudoc.echr.coe.int/app/conversion/pdf/?library=ECHR&id=001-60958&filename=001-60958.pdf. Accessed on September 26, 2020.

European Parliament. (2011). European Parliament resolution of 10 March 2011 on media law in Hungary, P7 TA(2011)0094, para. E., [2012] OJ C 199 E/154. Retrieved from https://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//TEXT+TA+P7-TA-2011-0094+0+DOC+XML+V0//EN. Accessed on September 26, 2020.

European Parliament. (2012). Motion for a European Parliament Resolution on the situation of fundamental rights: standards and practices in Hungary (pursuant to the European Parliament resolution of 16 February 2012) (2012/2130(INI)). Retrieved from https://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//TEXT+REPORT+A7-2013-0229+0+DOC+XML+V0//EN. Accessed on September 26, 2020.

European Parliament & Council. (2010). Directive 2010/13/EU of the European Parliament and of the Council of 10 March 2010 on the coordination of certain provisions laid down by law,

regulation or administrative action in Member States concerning the provision of audiovisual media services (Audiovisual Media Services Directive). *OJ* L 95/1.

Gall, L. (2012, August 22). The Clubbing of Klubradio. *Human Rights Watch*. Retrieved from https://www.hrw.org/news/2012/08/22/clubbing-klubradio. Accessed on September 26, 2020.

Gall, L. (2020, September 15). Hungary renews attacks on independent Radio Station. *Human Rights Watch*. Retrieved from https://www.hrw.org/news/2020/09/15/hungary-renews-attacks-independent-radio-station. Accessed on September 26, 2020.

Hajnal, G. (2016). Illiberal or simply unorthodox? Public Administration education in Hungary: A comparative perspective. *Teaching Public Administration*, *34*(2), 205–224.

Haraszti, M. (2011, January 16). Hungary's Media Law Package. A note by Miklos Haraszti, former OSCE representative of freedom of the media. *Statewatch*. Retrieved from www.statewatch.org/news/2011/jan/hungary-haraszti-media-law-package.pdf. Accessed on September 26, 2020.

Human Rights Watch. (2020, July 24). Hungary: Editor's sacking a blow to Press Freedom. *Human Rights Watch*. Retrieved from https://www.hrw.org/news/2020/07/24/hungary-editors-sacking-blow-press-freedom. Accessed on September 26, 2020.

Hungarian Civil Liberties Union (HCLU-TASZ). (2011, January 3). Concerns about media legislation in Hungary (part I). Retrieved from https://tasz.hu/cikkek/concerns-about-media-legislation-in-hungary-part-i-1. Accessed on September 26, 2020.

Hungarian Civil Liberties Union (HCLU-TASZ). (Year not mentioned but presumably 2011). Summary of the decision of the Constitutional Court of Hungary on the Media Laws in 2011 by the Hungarian Civil Liberties Union. Retrieved from https://tasz.hu/files/tasz/imce/2011/hclu_const_court_media_law_dec_brief.pdf. Accessed on September 26, 2020.

International Center for Not-For-Profit Law (ICNL). (Date not available). COVID-19 Civic Freedom Tracker. *Keep Civic Space Healthy*. Retrieved from https://www.icnl.org/covid19tracker/?location=54&issue=&date=&type=. Accessed on September 26, 2020.

Jakubowicz, K. (2010, September). *Analysis and assessment of a package of Hungarian Legislation and Draft Legislation on media and telecommunications*. Study commissioned by the Office of the OSCE Representative on Freedom of the Media, Organization for the Security and Co-operation in Europe, Office of the Representative on Freedom of the Media, Warsaw. Retrieved from http://www.osce.org/fom/71218?download=true. Accessed on September 26, 2020.

Kelemen, K. (2012, January 9). Hungary: The Constitutional Court annulled some provisions of the media laws. *Diritti comparati. Comparare i diritti fondamentali in Europa*. Retrieved from https://www.diritticomparati.it/hungary-the-constitutional-court-annulled-some-provisions-of-the-media-laws/. Accessed on September 26, 2020. Also published on January 16, 2012 in *Media Laws*, Law and Policy of the Media in a Comparative Perspective. Retrieved from www.medialaws.eu/hungary-the-constitutional-court-annulled-some-provisions-of-the-media-laws/. Accessed on September 26, 2020.

Kelemen, R.D. (2020, April 2). Hungary just became a coronavirus autocracy. *The Washington Post*. Retrieved from https://www.washingtonpost.com/politics/2020/04/02/hungary-just-became-coronavirus-autocracy/. Accessed on September 26, 2020.

Koltay, A. (2012, May 30). Hungary's new Media Regulations and the Constitutional Court. *Media, Council of the National Media and Infocommunications Authority*. Retrieved from http://hunmedialaw.org/dokumentum/306/Hun_media_law_JML_article_Koltay.pdf. Accessed on September 26, 2020.

Kornai, J. (2015a). Hungary's U-turn. *Capitalism and Society*, *10*(1), 1–24.

Kornai, J. (2015b). Hungary's U-turn. *Society and Economy*, *37*(3), 279–329.

Kovacs, C. (2012, April 10). *Freedom of expression in Hungary: Risks and perspectives. International conference on freedom of expression in the European Union and Canada/La liberté d'expression dans l'Union européenne* et au Canada, Faculté de droit, Université de Montréal, Montréal. Retrieved from http://www.calendrier.umontreal.ca/droit/?com=detail&eID=103888. Accessed on September 26, 2020.

Kroes, N. (2010, December 23). Letter from Neelie Kroes to Tibor Navracsics. Brussels: European Commission. Retrieved from https://www.europarl.europa.eu/RegData/commissions/libe/proces_verbal/2011/01-17/LIBE_PV(2011)01-17-1_EN.pdf. Accessed on September 26, 2020.

Kroes N. (2011, January 11). Digital Agenda Hungary's New Media Law: Open Hearing on Freedom of the Press in Hungary, European Commission Press Release Database, speech 11/6. Retrieved

from http://europa.eu/rapid/press-release_SPEECH-11-6_en.htm. Accessed on September 26, 2020.

Lengyel, M. (2010). Hungary: A wave of media legislation. *IRIS*, *8*(1/34). Retrieved from http://merlin-int.obs.coe.int/article/5450. Accessed on September 26, 2020.

Levitsky, S., & Way, L. (2002). Elections without democracy: The rise of competitive authoritarianism. *Journal of Democracy*, *13*(2), 51–65. Retrieved from http://muse.jhu.edu/article/17196. Accessed on September 26, 2020.

Linz, J. J. (2000). *Totalitarian and Authoritarian Regimes*. Boulder, CO: Lynne Rienner.

Majtényi, L. (2011, January 11). A criticism of the Hungarian Act on Media Services and Mass Media. Effective January 1, 2011. European Parliament, Open Hearing "Freedom of Press in Hungary," Eötvös Károli Institute, Brussels. Retrieved from http://www.ekint.org/en/media-and-press-freedom/2011-01-11/a-criticism-of-the-hungarian-act-on-media-services-and-mass-media-effective-january-1st-2011. Accessed on September 26, 2020.

Mérték Médiaelemzõ Mũhely. (2012, January 26). Ruling No. 165/2011. (XII. 20.) AB of the Constitutional Court-Summary. *MÉRTÉK*. Retrieved from http://mertek.eu/2012/01/26/ruling-no-1652011-xii-20-ab-of-the-constitutional-court-summary/. Accessed on September 26, 2020.

National Media and Infocommunications Authority (NMHH). (2010, August 18). A new organisation to face new challenges. *Press Release*. Retrieved from http://english.nmhh.hu/cikk/150045/A_new_organisation_to_face_new_challenges. Accessed on February 17, 2017.

New Europeans.net. (2020, September 12). *Klubrádió* to lose licence as Orban continues to trample on press freedom in Hungary. *Press release*. Retrieved from https://neweuropeans.net/it/node/3395. Accessed on September 26, 2020.

Novak, B. (2020, June 16). Hungary moves to end rule by decree, but Orban's powers may stay. *The New York Times*. Retrieved from https://www.nytimes.com/2020/06/16/world/europe/hungary-coronavirus-orban.html. Accessed on September 26, 2020.

Nyman-Metcalf, K. (2011, February 28). *Analysis of the Hungarian media legislation*. Study commissioned by the Office of the OSCE Representative on Freedom of the Media, Organization for the Security and Co-operation in Europe, Office of the Representative on Freedom of the Media, Tallin. Retrieved from http://www.osce.org/fom/75990?download=true. Accessed on September 26, 2020.

Orbán V. (2014). Prime Minister Viktor Orbán's Speech at the 25th Bálványos Summer Free University and Student Camp, July 26, 2014, Tusnádfürdõ (Băile Tuşnad), Romania. *Website of the Hungarian Government* (official English version), July 30, 2014. Retrieved from http://www.kormany.hu/en/the-prime-minister/the-prime-minister-s-speeches/prime-minister-viktor-orban-s-speech-at-the-25th-balvanyos-summer-free-university-and-student-camp. Accessed on September 26, 2020.

Organization for Security and Co-operation in Europe (OSCE). (2010). Hungarian media law further endangers media freedom, says OSCE media freedom representative. *OSCE, Press Release*, Vienna, December 21 or 22 (both dates are mentioned), 2010. Retrieved from www.osce.org/fom/74687. Accessed on September 26, 2020.

Pap, A. L. (2017a). Constitutional identity? The Hungarian model of illiberal democracy. In M. S. Fish, G. Gill, & M. Petrovic (Eds.), *A quarter century of Post-Communism assessed* (pp. 161–186). Cham: Palgrave Macmillan, Springer.

Pap, A. L. (2017b). *Democratic decline in Hungary: Law and society in an illiberal democracy*. London: Taylor & Francis.

Scheppele, K. L. (2015). *The trajectory of democracy — Why Hungary matters: Hearing before the Commission on Security and Cooperation in Europe: March 19, 2013*. Washington, DC: US Government Publishing Office.

Sedelmeier, U. (2014). Anchoring democracy from above? The European Union and democratic backsliding in Hungary and Romania after accession. *Journal of Common Market Studies*, *52*(1), 105–121.

Simon, Z. (2014, December 15). Hungary on path to shed Junk Grade and Shield Forint, Orban says. *Bloomberg*. Retrieved from https://www.bloomberg.com/news/articles/2014-12-15/hungary-on-path-to-shed-junk-grade-and-shield-forint-orban-says. Accessed on September 26, 2020.

The Fundamental Law of Hungary. (2011, April 25). (Official version in English). Retrieved from http://hunconcourt.hu/rules/fundamental-law. Accessed on September 26, 2020.

Thorp, N. (2020, June 16). Coronavirus: Hungary votes to end Viktor Orban emergency powers. *BBC*. Retrieved from https://www.bbc.com/news/world-europe-53062177. Accessed on September 26, 2020.

Walker, S. (2020a, July 22). Editor-in-chief fired at Hungary's leading independent news site. *The Guardian*. Retrieved from https://www.theguardian.com/world/2020/jul/22/editor-in-chief-fired-hungary-leading-independent-news-site-szabolcs-dull-index. Accessed on September 26, 2020.

Walker, S. (2020b, July 24). Hungarian journalists resigne en masse after claiming of political interference. *The Guardian*. Retrieved from https://www.theguardian.com/world/2020/jul/24/hungarian-journalists-resign-en-masse-after-claims-of-political-interference. Accessed on September 26, 2020.

Walker, S. (2020c, April 3). Hungarian journalists fear coronavirus law may be used to jail them. *The Guardian*. Retrieved from https://www.theguardian.com/world/2020/apr/03/hungarian-journalists-fear-coronavirus-law-may-be-used-to-jail-them. Accessed on September 26, 2020.

Walker, S. (2020d, May 26). Hungarian government to end Orbán's rule-by-decree legislation. *The Guardian*. Retrieved from https://www.theguardian.com/world/2020/may/26/hungarian-government-to-end-orbans-rule-by-decree-legislation-emergency-coronavirus?CMP=Share_iOSApp_Other. Accessed on September 26, 2020.

Walker, S., & Rankin, J. (2020a, March 30). Hungary passes law that will let Orbán rule by decree. *The Guardian*. Retrieved from https://www.theguardian.com/world/2020/mar/30/hungary-jail-for-coronavirus-misinformation-viktor-orban. Accessed on September 26, 2020.

Walker, S., & Rankin, J. (2020b, May 13). Hungary's coronavirus laws prompt new showdown with Brussels. *The Guardian*. Retrieved from https://www.theguardian.com/world/2020/may/13/hungary-viktor-orban-coronavirus-laws-prompt-new-showdown-with-brussels. Accessed on September 26, 2020.

Whittington, K. E., Kelemen, R. D., & Caldeira, G. A. (Eds.). (2008). *The Oxford handbook of law and politics*. New York, NY: Oxford University Press.

Zakaria, F. (1997). The rise of illiberal democracy. *Foreign Affairs*, *76*(6), 22–43.

CHAPTER 10

STIRRING UP STRIFE: THE CENSORSHIP OF COMMUNIST PUBLICATIONS IN LATE COLONIAL INDIA

Devika Sethi

ABSTRACT

Purpose – *Public debates about censorship laws largely focus on their desirability and the limits set on free speech. From a historical perspective, however, the logic and contradictions inherent in these laws' implementation, as well as their evasion, also merit attention. This chapter places at the heart of its investigation the General Communist Notification (1932) in British India which prohibited specific kinds of Communist publications from import and circulation, even more so in a context of mass anti-colonial nationalism.*

Methodology/Approach – *Using government and intelligence agencies' archival records, intercepted documents of the Communist Party of India, legislative debates and memoirs, this chapter illustrates the censorship of Communist literature in India at two levels: one, it sketches a broad picture of the mode and extent of the censorship of Communist literature in late colonial India (c. 1925–1947). Two, by excavating debates and processes around the treatment to be accorded to books of two British Communist writers, John Strachey and R. P. Dutt, it reveals the constraints and dilemmas of censorship of Communist literature. While doing so, it brings both Indian and British voices to the fore.*

Findings – *This investigation provides valuable insights into the operation of laws related to specific genres of publications, provides an assessment of the success of censorship measures, and highlights the repercussions of their failure.*

Media and Law: Between Free Speech and Censorship
Sociology of Crime, Law and Deviance, Volume 26, 169–183

ISSN: 1521-6136/doi:10.1108/S1521-613620210000026011

Originality/Value – *By illustrating the limited success of censorship measures, as well as the dilemmas of censors and debates among them, this chapter urges for a more nuanced and multidimensional understanding of the operation of censorship, particularly in politically fraught contexts.*

Keywords: Communism; British India; Rajani Palme Dutt; censorship; Communist Party of India; free speech

INTRODUCTION: THE CASE AGAINST COMMUNIST LITERATURE

In 1951, four years after India ceased to be a British colony, the American Ambassador to India observed that bookshops in that country were full of Communist publications.[1] The next year, the Communist Party of India (CPI) emerged as the largest opposition party in Parliament; five years later, the Indian state of Kerala had elected the first democratically elected Communist government in the world. The circulation of Communist publications in India during the late colonial period (*c*. 1925–1947) was not so unhindered. The CPI was established in 1925, banned by the Government of India (GoI) in 1934, and not legalized till 1942. Even so, laws and official notifications were stymied not just by evasive publishers and readers, but also by the intricacies of implementation in a colonial context. This chapter excavates that history. It examines the problems and paradoxes that marked the implementation of the General Communist Notification (GCN) of 1932. It does so by highlighting internal debates within the colonial state, questions raised about this specific form of censorship by British and Indian lawmakers, and evasion by readers. This investigation thus examines not the making of a law, but its implementation and resistance to it in three tumultuous decades in modern Indian history.

In twentieth century India, the colonial situation, proximity to Soviet Russia, the beginnings of mass anti-colonialism and the presence of Indian revolutionaries abroad led to a dynamic – and from the point of view of the colonial state – a dangerous cocktail. Movements espousing mass-based nationalism, revolutionary terrorism, and Communism all engaged with each other, and often shared personnel, ideas, and methods. Gerald Barrier, the pre-eminent historian of censorship in twentieth century colonial India, states that Bolshevik documents began circulating in India in the immediate aftermath of the First World War, and that the Communist International (Comintern), in its capacity as the international center of the world Communist movement, directly targeted the destruction of the British empire as a "premier goal." Publications were dispatched to India by Moscow trained Indian revolutionaries, as well as from Indians such as the Berlin based M. N. Roy, active in the global Communist movement and an associate of Vladimir Lenin.[2]

The GoI on its part adopted a multipronged strategy to keep Communist literature and influences out: postal cordons, customs acts, specific and general laws and notifications targeting Communist literature, and "conspiracy" trials, for instance, the ones at Peshawar (1922–1923), Cawnpore/Kanpur (1924), and

the most famous one at Meerut (1929–1933) where 33 Communist labor leaders (including three British Communists) were put on a trial that lasted over four years and were given harsh sentences.[3] British intelligence officials seemed to simultaneously despise and fear Communists in India. On the one hand, they harbored a low opinion of Indian Communists as being "greedy opportunists," and unlikely to succeed on account of Mahatma Gandhi's domination over the Indian masses.[4] On the other, they were aware that the Communist movement had a global scope, whereas the appeal of sedition was limited to India itself.[5] Their fear of unfavorable comparisons with Soviet Russia was such that in 1934 the English translation of Nobel prize winning Indian public intellectual Rabindranath Tagore's letters from Soviet Russia (which he visited in 1930) were banned in India for their admiration of that country.[6]

For the colonial state, the link between what Indians read and how they subsequently acted was a clear one. This was particularly so when the publications in question urged revolutionary change or violent methods. In January 1933, Judge Yorke sentenced 27 of the 30 accused in the Meerut conspiracy case to punishments ranging from transportation for life to rigorous imprisonment for three years. In his judgment he referred to the plan of the Comintern to overthrow existing governments throughout the world. Their methods, in his view, included encouragement of strikes, as well as "the conduct of propaganda by speeches and leaflets."[7] The appellate judgment of the Allahabad High Court in August 1933 reduced the sentences of the accused, and acquitted nine of them, but the judges nevertheless took a serious view of the conspiracy. In their view:

> […] when a seditious appeal is made to illiterate and ignorant workers and peasants, when organized propaganda is carried out among them, particularly during strikes, and when inflammatory speeches are made, it is not possible to look upon such conduct merely with contempt on the ground of its futility; such a course of action might well lead to a serious general strike, a widespread breach of the public peace, and bloodshed.[8]

Colonial assessments about India and stereotypes about its supposedly gullible population had a bearing on the risk-assessment of controversial literature, as far as state personnel were concerned. This was also a form of paternalism, where the colonial state claimed to know what was best for its subject population, and aimed to protect them from ideas whose violent import they could supposedly not understand in full measure. This judgment lays bare for us the position of colonial officials, and explains the necessity, in their view, for strict surveillance and control over Communist literature.

THE LEGAL FRAMEWORK: THE OPERATION OF THE GENERAL COMMUNIST NOTIFICATION (1932)

The CPI was founded in 1925, but it was not until 1932 that the GoI issued a general ban on the import of Communist publications to the country via the GCN. It applied to publications that emanated from three sources: the Comintern, any organization "affiliated to or controlled by or connected with" it, or from any

person who held office in any such organization. It also applied to documents that contained "substantial reproductions" of the matter banned under the notification.[9] The GCN banned (technically, prohibited) these works from import into India under the terms of the Sea Customs Act. Gerald Barrier estimates that between 1932 and 1934 as many as 150 of the total 175 publications seized while entering India pertained to Communism.[10] Communist themselves believed that the Sea Customs Act was a conspiracy to keep Communist literature out of India in the immediate aftermath of the First World War, when large-scale censorship measures current during war-time could no longer have been applied.[11] Much of the Communist literature in India was imported, and therefore the GoI initially deemed this strategy to be an effective one.

However, as banning a publication in India required proscription by the local government, a book that was prohibited entry under the Sea Customs Act was not necessarily – technically speaking – banned in India. This distinction between prohibited literature (i.e., imported publications, which fell under the purview of the GCN or the Sea Customs Act) and proscribed literature (typically publications printed within India, banned by specific notifications, or for falling afoul of provisions of the Indian Penal Code) was a crucial one when it came to practical matters of implementing the law, as we shall see later.

Senior officials of the Home Department of the GoI examined and read suspicious books. Books by certain authors on the GoI's radar were all intercepted and examined, though not prohibited automatically. If deemed "harmless," they were released.[12] This required coordination between the DIB, the Home department of the GoI, Post Office and Customs officials, as well as provincial departments, including police. The system was subject to various pressures: differences of opinion among officials, lack of coordination among departments, and questions regarding the surveillance and control of publications in the British Parliament as well as by Indian legislators.

The Intelligence Bureau took pains to emphasize that the GCN was not applied to all literature about Communism, or by all Communists. In 1938, the Deputy DIB, P. C. Bamford, assured the Home Member R. M. Maxwell that the GCN was not applied indiscriminately, and enclosed two lists (which he admitted were not comprehensive). One contained books that were technically covered by the GCN but were still allowed entry after examination, while the other contained a list of books on Communism that were considered harmless.

The first list – that of examined, then passed, books, 17 in all – consisted of titles such as all editions of Karl Marx's *Das Capital,* Marx's *Critique of Political Economy*, Maxim Gorki's *Last Plays*, Maurice Dobb's *Marxism Today*, all scientific publications by USSR's Academy of Science and even a biography called *Genghis Khan* by Ralph Fox. Seven of the 17 books listed either were either USSR publications, or relating to that country, including titles like *Scenes from Soviet Life* by W. P. & Z. K. Coates.[13]

The second list (the "harmless" one) included 26 books with titles decidedly more action-oriented than the books on the first list, for example, P. Kropotkin's *An Appeal to the Young*, Rudolf Rocker's *Anarcho Syndicalism*, Harold Laski's *Communism*, L. Kameneff's *Dictatorship of the Proletariat*, H. W. Lee's *Evolution*

of Revolution, Edgar Snow's *Red Star Over China*, and R. Osborne's *The Psychology of Reaction*.[14] Bamford also explained which books on Communism were considered harmless: serious works written by individuals neither connected with Communist organizations, nor published by firms that published on behalf of Comintern. The GCN, added Bamford, helped prevent Comintern sponsored Communist literature from reaching India, and also served to filter correspondence. This was because, as he put it, "The printed word ranks high among the prescribed weapons of communist propagandism…." Bamford stated that the GCN had been applied with ever increasing liberality in the last two years, and that people in India were not really deprived of "harmless books on Communism."[15]

Even as the GoI appointed itself the confident arbiter of harmful and harmless books on Communism, the operation of the GCN was marked by inconsistencies and contradictions. Case studies of the books of two authors, John Strachey and R. P. Dutt, serve to demonstrate the complexities and dilemmas involved in the application of censorship laws and provisions to control print media in late colonial India.

CENSORS' DEBATES AND DILEMMAS: THE CASE OF JOHN STRACHEY

In 1937, elections were held in Indian provinces and the Indian National Congress (INC) – the GoI's chief nationalist opponent – assumed power in eight of the 11 provinces of British India. A leading light of the INC at this time was Jawaharlal Nehru, who had socialist sympathies, unlike many other Congressmen.[16] In these circumstances, the GoI intercepted a letter between a functionary of the All Indian Congress Committee and a British publisher, Victor Gollancz, regarding a tie up so as to publish left literature in India under the auspices of a "Left Book Club." The GoI did not welcome this development, more so when it emerged that the plan had the full support of Jawaharlal Nehru.[17] Gollancz had earlier come to the notice of the GoI for despatching "objectionable" literature to India, and had, by his own admission, "always absolutely refused" to submit publications to the India Office in London prior to sending them, as was recommended to publishers.[18]

An official of the Home department conceded that the publication of such literature in India would be difficult to control, as only one imported copy would be required to start off the process of replication and circulation throughout the country. INC ruled provinces were unlikely to use provisions of the Indian Press Act of 1937, which in any case did not cover Communist propaganda.[19] The Home Secretary's views were expressed more succinctly: "Dangerous if successful, but I doubt if funds will be forthcoming."[20] The Home Member was less optimistic: he too considered the plan one of "disturbing possibilities," but feared that funds would eventually be arranged.[21] The CPI and INC had a complex relationship: periods of amity between them were interspersed with periods of mutual derision and suspicion. During the latter, the CPI disparaged the INC for its bourgeois methods and aims, in line with directives from Comintern collaborating or not

with "bourgeois nationalist" parties and forces. An alliance between them was therefore a source of unease for the colonial state.

In any case, British colonial bureaucrats themselves – coming from a free country where they were used to holding and expressing divergent views – themselves rarely agreed on the risks posed or not by specific books. One such debate shows us how colonial discourse on the issue of censorship generally and of Communist publications particularly was itself a variegated creature. This debate – between the Home Secretary and the DIB – was over John Strachey's *The Theory and Practice of Socialism* (1936). Strachey was a British Communist and an economics expert, who also served in 1935 as Treasurer of the British Anti-War Council, the branch of an organization funded by the Comintern. The book was already banned in India under the GCN, and it was this ban that was challenged in 1938 by K. B. Menon, the Secretary of the Indian Civil Liberties Union, who suggested that the GoI consider lifting the ban.[22] Home Secretary J. A. Thorne – after carefully reading the book – held the book to be "a serious and sober book which, neither for its matter not for its tone, can properly be banned."[23] In fact, he further recommended that import of the book also be allowed because people outside India were likely to form a negative impression of India's administration if "sober" works on Communism were excluded, and because:

> […] some resentment is caused among reasonable people in India by their being deprived of the opportunity of reading what intelligent and reasonable people in England are interested in.[24]

Thorne also suggested that the GCN be used more like a filter rather than a wall through which nothing would pass. The decision to unban was, however, subject to approval by the DIB P. C. Bamford. The latter, in his official notings and discussions with the Home Secretary, was in favor of the book remaining under ban for the following reasons: the author was associated with two bodies (the Left Book Club and the Writers' International) which were connected to the Communist International[25]; the book preached Communism; and because "there is no real demand in India, even from the Congress Party, that restrictions on Communist propaganda should be relaxed."[26] Bamford was unwilling, therefore, to compromise on the issue of Strachey, whose three other books were also withheld in India: *What are we to do?, The Coming Struggle for Power*, and *The Nature of Capitalist Crisis*. Bamford listed several reasons for his intransigence: Strachey was well reputed as an authority by Indian Communists, his books were used by the Party's study circles, and "attempts to smuggle in Communist literature almost invariably include a 'Strachey.'" Bamford acknowledged that although Strachey's books may not be more harmful to the "average person" than the books that were allowed, nevertheless:

> […] they preach communism and class-distinction with such emphasis as to mean class-hatred, and are therefore capable of causing harm on the immature minds of Indian students who would be the main readers of them and who are incapable of distinguishing between Communism as a political theory and the ultimate and violent stage of armed revolution. I would stress the point that books advocating communism are likely to do much more harm in India than in England. I suggest that the Government of India should not be influenced by unfavourable estimates of the intelligence and intention of the administration … formed by people abroad who probably know little about local conditions. It is the estimate of people in India that counts…[27]

In the difference of opinion that developed between the DIB and the Home Secretary (who favored selective application of the notification, and expressed the view that Strachey's books may have been withheld without sufficient reason),[28] the Home Member H. D. Craik sided with Bamford. As the settled policy of the GoI had been, he stated, "to upset the Communist schemes of Propaganda in India as much as possible," because in his opinion Strachey's emphasis on coercion was "dangerous," and because the book would definitely help the Communist aim of "creating class consciousness and antagonism," the Home Member favored barring the book from India, and it was his view that eventually prevailed. The view that Indians readers could be as discerning as British ones was trumped by the view that held them to be incapable of such maturity and judgment. This view was to be reiterated in another case.

CHANGING TIMES, CHANGING STRATEGIES: THE CASE OF R. P. DUTT

Rajani Palme Dutt (1896–1974) was a leading light of the Communist Party of Great Britain (CPGB). Born to an Indian father who had migrated to Britain, and a Swedish mother, Dutt also happened to be the grand-nephew of R. C. Dutt, one of the pioneers of the economic history of India written from a nationalist perspective. As a student at Oxford during the First World War, he declared himself to be a conscientious objector, and was imprisoned. He later founded the *Labour Monthly* journal in 1921, and continued to edit it for 50 years. In the 1930s he befriended nationalist leader and the future first Prime Minister of India, Jawaharlal Nehru, and influenced his thinking. Although he had been writing about India since long, he first visited the country only in 1946. As a member of the Comintern, he was involved in mentoring the CPI, and this made him a suspicious figure for the GoI.

In 1938, an (Indian) legislator asked the (British) Home Member of the GoI if, and why, R. P. Dutt's two books – *Fascism and Social Revolution* and *World Politics, 1918–1936* – were banned in India (since 1934 and 1936, respectively).[29] R. M. Maxwell replied that both books were banned as they fell under the scope of the GCN of 1932. Other legislators then reminded the Home Member that these and many other books banned by the GoI were recommended for study in schools and universities in England. In other words, Indians too could read these books outside India. The Home Member admitted that the books were banned because of the author's affiliation to the Comintern rather than the content of the books. When asked if books that propagated Fascist ideas were also banned, the Home Member replied that "We have not reached that stage."

One legislator shared the information that R. P. Dutt's books were available at Higginbothams (a government authorized railway bookseller) in 1936 and 1937. Yet another Indian legislator wanted to know who in Government read these books, and whether "any human brain was applied to this book." When a third legislator suggested that as the banning of these books led to "wrong ideas about Communism being formed in this country," and urged the Government to

reconsider, the Home Member was closed to the suggestion, and said that this logic could also apply to the importation of opium.[30] This discussion reveals two important aspects about the censorship of Communist literature in this period. One, we see the rule of "colonial difference" operating clearly, where the metropolis and the colony were subjected to entirely different rules, despite the claims of "liberal imperialism."[31] Two, in the operation of the GCN, it was not as much the content of a specific book that was under scrutiny, but its source, or the political and institutional affiliation of its author.

R. P. Dutt's books evoked different reactions from members of the colonial bureaucracy. Whereas the DIB, V. T. Bayley, denounced the books as "Communist propaganda of the virulent type" by an "expert propagandist," the Home Secretary, J. A. Thorne, after reading the two books, considered them "pretty harmless."[32] The Home Member opined that the GCN did not actually prohibit the import of Communist "literature" or "propaganda" as such, only documents falling under specific clauses, whereupon the DIB commented that R. P. Dutt's work did fall in these clauses, as he was an office holder of the CPGB, controlled by the Comintern. He cited a report from London stating that Dutt was in fact "second-in-command of the CPGB, if not, as is believed in some quarters, the real leader of the Party."[33] In such circumstances, the GoI did not favor lifting the ban on his books. In other words, books on Communism by theoreticians may well have been deemed safe, but never those by an active practitioner of the ideology. Though colonial officials had varied opinions of the risk posed by Dutt's writings (attesting to the subjective nature of censorship), it was finally the view of Intelligence officials that prevailed.

Dutt again came to the notice of the GoI in 1940 for his book *India Today*, published by British publisher Victor Gollancz in a Left Book Club edition. The book critiqued the economic exploitation of India by the British from the mid-eighteenth century onwards. In his review of this book in 1940 DIB V. T. Bayley termed it

> a tiresomely long treatise with massive documentation in support of the usual theme of "population, poverty, and the 'drain'" … over 500 pages of unusually bitter hatred of British rule….

Calling it Indian Communists' *vade mecum*, or reference handbook, Bayley stated that the book stressed poverty in India, without mentioning conditions before the arrival of the British or "the factors in Hindu tradition and social institutions that contribute so largely to the poverty and backwardness of India." Bayley made much of R. P. Dutt – and Marx – having never visited India, and accused Dutt of quoting official sources torn out of context. Dutt contemplated, according to Bayley, revolutionary violence without thinking of the sacrifice of human life and happiness that it involved. Although the book was already banned in India, Bayley wanted it banned by a separate notification, because:

> In intelligent hands a book of this kind carries its own obvious refutation despite the writer's obvious ability and the many half-truths he has assembled. But in the hands of impressionable Indian youth it is capable of causing a whole lot of mischief, whilst its library of "facts" and glibly expressed anti-imperialist arguments are just the ammunition the various extremists in this country want for their subversive attacks and "anti-war" tirades…[34]

Another Home department official wanted the book banned because it emanated from an "extreme" left point of view, and advocated Soviet democracy, end of imperial domination, agrarian revolution, and liquidation of landlordism, among other things.[35] Anticipating that after the ban on the import of the book (via the GCN) attempts would be made to publish local editions, the IB informed local intelligence departments that these local editions too would contravene orders under the Defence of India Rules (special war-time provisions) banning publication of any matter in any document that was itself prohibited entry into British India.[36] Despite all these precautions, two months after the book was banned (in November 1940) by a specific notification,[37] bayley noted that Communist leaders in Bombay had managed to smuggle copies of the book, and had already printed what they called "the most valuable parts of the book" in three pamphlets that had been secretly printed and distributed from Bombay (Mumbai).[38]

The book was viewed by the CPI very differently: for them Dutt's formulation that the agrarian revolution was inevitably linked with the national revolution was particularly valuable, and they thought it contained important "historical and statistical material" that "had to be made available to the Party and the public." Dutt's earlier works on India had been variously hailed by CPI leaders as "the best and the most popular exposition of the programme of our party," as "the textbook of the Communists in India," and "the best Marxist analysis of the Indian politics then available. It inspires the pioneers of our party and laid its ideological foundation."[39]

Less than a year after the ban on *India Today,* circumstances had changed, as had the GoI's relationship with the CPI. In fact the CPI had a chequered relationship both with British colonialism and the INC. For a large part of its existence, the CPI was a banned political party. Germany's attack on Russia in June 1941 during the course of the Second World War led to a change in the position of the CPI toward the GoI, as they declared the hitherto "imperialist war" (which they were bound to oppose) to now be – with Soviet Russia joining the Allied powers – a "People's War" which they were bound to support.[40] In other words, the CPI's legalized status came at the cost of supporting British imperialism in a war-time context. This cooperative relationship between the CPI and GoI led to changes in the way the GoI dealt with their publications as well.[41] When, in 1943, the IB reviewed Dutt's *Guide to the Problem of India* (an abridged version of *India Today* with a few additional chapters, published in 1942), it found that the new chapters were objectionable as well, for even as they castigated the INC for threatening civil disobedience, they also referred to Government "provocation." Entry into India of *India Today* had been banned in 1940, and that of the *Guide* also banned, as it was a "substantial reproduction."

After news reached the GoI of the CPI's plans to republish the *Guide* in India itself, the IB urged that a specific notification in reference of this book be issued.[42] However, Additional Secretary of the Home department, Richard Tottenham, was reluctant to use the war-time Defence of India Rules against it as it would simply give the book and its author publicity, and there was some doubt about which rule to apply in any case.[43] He suggested:

> Would it not be possible to treat the Communist people in Bombay as reasonable beings – tell them that the entry of this book has been banned and that we do not want them to republish it – and see what happened [*sic*.]. They have plenty of other literature to get out and not too much paper![44]

This is what was done, and the General Secretary of the CPI, P. C. Joshi, was advised accordingly.[45] By end November 1943, the IB reported that the CPI had given up the idea of reprinting this particular book. "This commonsense approach seems to have worked," commented a satisfied Home department official.[46] In a context of war-time truce with the CPI, the GoI was thus able to use persuasion as opposed to legal measures to get the outcome it desired. Dictates of practicality, not unchanging ideological principles, determined censorship policy, more so in a war-time context.

It is also insightful to view the situation from the other side, that of the CPI. As revealed in their internally circulated correspondence, the Party had extra-ideological reasons for their desire to publish Marxist literature. In an internal party document in 1943, the Central Committee of the CPI listed three reasons why they wanted to publish an "Indian edition of Marxian classics in English":

> The hunger for reading Marxian classics is so great that unless we fulfil it any publisher can run away with publishing a book and make profits.
>
> It will fulfil a great political need.
>
> It will bring us cash.[47]

As the negotiations over the R. P. Dutt episode showed, however, political considerations and a war-time alliance with the colonial state prevailed over ideological – and monetary – ones in the end.

EFFICIENCY AND EVASION

How successful were colonial legal provisions in meeting their professed aims? By the late 1930s, there was general agreement in the Intelligence department of the GoI regarding two points: first, written propaganda of Communist and other varieties had the capacity to evoke action; and second, the GoI was not very successful in hindering the free circulation of texts that were deemed dangerous. In 1938, a senior IB official attempted to provide a numerical estimate of GoI's efficiency in keeping out prohibited (imported) literature. In his opinion, "With the means available I doubt that we can expect our efforts to keep put prohibited literature to be, on an average, more than 75% efficient." This was because the time available to Customs to search incoming packet post was "strictly limited," and neither could they search all packets or even all postal bags. He stated that it was difficult to estimate the amount of prohibited matter that passed through, some of which would be held up at the post office, but not all.[48]

Despite provisions such as the GCN at their disposal, in 1939 the IB reported that the CPI, at this time a banned organization, was using the print media to their utmost advantage. Newspapers and pamphlets in English and other languages were circulated at inexpensive or "popular" prices, and there was:

> […] no doubt that the present increase in industrial and peasant unrest is attributable largely to the cumulative effect of this propaganda, and the freedom of press and platform in the hands of individuals whose sole purpose is to preach class-hatred and stir up strife.[49]

In addition to indigenous propaganda, the report stated that foreign Communist propaganda was in demand and very much a source of inspiration. The report conceded that this could not "altogether be excluded in spite of the considerable seizures in every incoming foreign mail under the provisions of the Sea Customs Act." The failure to seize these meant that foreign editions of Marxist and Leninist writings were printed locally for circulation, serving as "practical hand-books of revolution." The result was that "…every element of the left is thoroughly conversant with the doctrines and practices of revolutionary communism which govern their whole outlook."

Regular publications inspired by foreign propaganda were published in several Indian languages and bore rousing titles such as *Azad* ("Independent"; Sindhi), *Kirti Lehr* ("Wave of Victory"; Gurmukhi, Urdu, Hindi), *Naya Hindustan* ("New India"; Hindi), *Inquilab* ("Revolution"; Hindi), *Chingari* ("Spark"; Urdu), *Lal Jhanda* ("Red Flag"; Urdu and Hindi), *Karanti* ("Revolution"; Marathi), and *Comrade* (English). In addition to these, there were also inexpensive pamphlets advocating Communism, which came "complete with forewords by leading Indian communists on how best to translate theory into practice." The report noted a press source as saying that these pamphlets, "rather than the bigger newspapers dealing with news as such, are influencing mass opinion and deflecting it more and more to the left."[50]

In 1935, an Indian legislator, P. N. Sapru, argued in the Council of State that Communist books ought not to be proscribed. Referring specifically to three books (Leon Trotsky's *History of the Russian Revolution*, Ralph Fox's *Lenin*, and John Strachey's *Coming Struggle for Power*), he argued that:

> These books are not likely to be read by unintelligent men, but by intelligent men who have given a thought to these grave questions of economic reconstruction …. If you want to deal with communism in this stringent manner, you will have to shut up your universities. You will have to deport your teachers of philosophy, your teachers of economics, your teachers of politics, your teachers of history. You will then have peace, but it will be the peace of the grave. This is not the way in which you will be able to deal with communism. I do not accept the communistic philosophy. I think it is fundamentally wrong as a philosophy. But I do say that the communist has a right to express his own opinions so long as he does not commit any overt act.[51]

It is not surprising that in a colonial context this defence of classic liberal principles came from the colonized, and not from the colonizer.

P. N. Sapru may not have been an admirer of Communist principles, but several Indian students and even the occasional British civil servant were passionately devoted to them, and served as agents to smuggle and disseminate it in India. Indian writer Sajjad Zaheer, a committed Communist who had studied at Oxford university, explained the procedure of smuggling proscribed publications from England to India in the 1930s thus:

> Sometimes eminent Indians would come to England, eminent in the sense that they were not suspected by the British government … we would give them a parcel of books or a little handbag consisting of Marxist literature and ask them to deliver it to so and so in India.

One such Indian was Prof. S. Radhakrishnan, Professor of Eastern Religion and Ethics at Oxford University. Zaheer and his friends requested him to carry some books and told him they were of Lenin and Marx. The future first President of India thus became complicit in the bringing of proscribed literature to the country.[52]

An even more startling example was that of Michael Carritt, an Oxford educated ICS officer in Bengal in the 1930s, who began reading left literature in the early 1930s. He knew he was breaking the law: books by Marx and Lenin, Harold Laski and G. D. H. Cole, which he read, were all banned under the GCN and as a British civil servant he was well aware of the fact. Furthermore, he knew that there was "vigilant censorship of all incoming mail from Europe in order to exclude such seditious stuff"; even as, later in his career, he signed police orders for the interception of such parcels, as a young officer he thought it was a "safe bet that parcels addressed to a District Officer would slip through the fishing net of a suspicious police."[53] After making contact with Ben Bradley (a British Communist convicted in the Meerut conspiracy case), he received instructions and money to make contact with the banned CPI, to publish an illegal news-sheet, and to carry Comintern literature to India. During the course of his activities, he was "surprised by the outstanding inefficiency of the methods and detective work of the Police Intelligence Bureau in Bengal."[54]

CONCLUSION

British officials in India were not blind to the operation of the principle of colonial difference. Banning of books, including Communist ones, emanating from Great Britain necessitated close coordination with the India Office in London, from where the Secretary of State for India oversaw the functioning of the GoI. On the other hand, the GoI – and more so the IB – were very conscious of their status as "the final judges of what we can and cannot admit to India," and unwilling to concede this right to the India Office. The reason for this was clearly enunciated by an official who stated that:

> The India Office view is inevitably influences by the principles classified in Great Britain in respect of freedom of expression of opinion; principles which it is not practicable to extend to India.[55]

Colonial discourse around Communist publications stressed on the authors' intentions, affiliations, and agenda, but also focused on their Indian readers' supposed gullibility, ignorance, and susceptibility to propaganda. It was not so much *what* a book said, as *to whom it said it to*, that determined its fate.

One historian has contended that although the CPI did not play a major role in the transfer of power from colonial to Indian hands, "its intellectual and political achievements were by no means insignificant."[56] The failures of colonial censorship were at least partly responsible. The British estimated that the party had a membership of 5,000 in 1942, and 30,000 by mid-1945, in addition to 2,50,000 trade union members' following. Irfan Habib attributes this growth in part to the circulation of Marxist literature, more so after the ban on it was removed during

the Second World War. By the end of that war, the CPI published 10 weekly journals, one (*People's War*) with a weekly circulation of 25–30,000 copies.[57] Although historian of imperial intelligence activities Richard Popplewell credits the failure of the Comintern itself in India to the success of British intelligence,[58] it could equally be argued that the popularity of Communist ideas themselves could at least partly be attributed to the inability of the law to keep a check on Communist publications.

NOTES

1. Loy Henderson expressed concern about Communist propaganda in India, and mentioned journals that attacked America "all the time." He also noticed the absence of US publicity material. Note by Jawaharlal Nehru, the then Prime Minister of India, regarding his meeting with the American Ambassador, 15 September 1951. Reproduced in Gopal (1994, p. 629).
2. Barrier (1974, p. 92).
3. Barrier (1974, p. 96). Roy was an Indian revolutionary who had escaped the country in 1915 and spent time in the United States, Mexico, and Russia. He was the main organizer of the Comintern's activities in India. For a discussion on the ban on the party, see Manzer (2004).
4. Petrie (1972 [1927]). Petrie wrote this book in his capacity as Director of the Intelligence Bureau (DIB), GoI.
5. Manjapra (2010).
6. Bairathi (1987, p. 18).
7. *Moral and Material Progress Report 1932–33* (1934, pp. 48–49).
8. *Moral and Material Progress Report 1932–33* (1934, pp. 49–50).
9. GoI Finance Department (Customs) Notification No. 61, 10 September 1932.
10. Barrier (1974, p. 126).
11. Adhikari (1971, p. 69).
12. Question and supplementary questions asked and answered in the Legislative Assembly in India on 1938, September 20. Note by R. B. Elwin, 14 September 1938. GoI Home Political f. No. 22/67, 1938. This and all archival sources cited in this chapter were consulted at the National Archives of India, New Delhi.
13. List A, Appendix I accompanying note by P. C. Bamford, 23 May 1938, GoI Home Political f. No. 27/4/38. Reproduced in Chatterji (1999, pp. 1008–1009).
14. List B, Appendix II accompanying note by P. C. Bamford, 23 May 1938. Reproduced in Chatterjee (1999, pp. 1009–1010).
15. Note by P. C. Bamford, 23 May 1938. Reproduced in Chatterjee (1999, p. 1006).
16. For an account of the relationship between Congress ministries (1937–1939) and the Communist movement, see Sethi (2019).
17. Note by DIB V. T. Bayley, 4 December 1937. Reproduced in Chatterjee (1999, pp. 1010–1012).
18. Gollancz's statement quoted as an extract from an intercepted letter between him and Mahmuduzaffar Khan in an IB note, 7 January 1938. Reproduced in Chatterjee (1999, pp. 1010–1012).
19. Note by H. S. Stephenson, 3 December 1937. Reproduced in Chatterjee (1999, pp. 1010–1012).
20. Note by R. M. Maxwell, 7 December 1937. Reproduced in Chatterjee (1999, pp. 1010–1012).
21. Note by H. D. Craik, undated. Reproduced in Chatterjee (1999, pp. 1010–1012).
22. IB note, 12 July 1938 on John Strachey. Reproduced in Chatterjee (1999, p. 1008).
23. Note by J. A. Thorne, 25 April 1938. Reproduced in Chatterjee (1999, p. 1004).
24. Note by J. A. Thorne, 14 May 1938. Reproduced in Chatterjee (1999, p. 1005).

25. Note by P. C. Bamford, 13 May 1938. Reproduced in Chatterjee (1999, p. 1004).
26. Views of the DIB in conversation with the Home Secretary J. A. Thorne, as reported by the latter in note, 14 May 1938. Reproduced in Chatterjee (1999, p. 1005).
27. Bamford also stated that the general silence of Congress governments on this issue, except that of Madras, signified that they did not seriously object to the use of the notification. Note by P. C. Bamford, 23 May 1938. Reproduced in Chatterjee (1999, p. 1007).
28. Notes by J. A. Thorne, 14 May 1938 and 26 May 1938. Reproduced in Chatterjee (1999, p. 1005 and p. 1007), respectively.
29. Question asked and answered by Mrs. K. R. B. Subbarayan on 16 August 1938. GoI Home Political f. No. 22/36, 1938.
30. Extracts from Legislative Assembly debate on 16 August 1938 in GoI Home Political f. No. 22/36, 1938.
31. For an explication of "colonial difference," see Chatterjee (1993). Robert Darnton has pointed to the inconsistencies of "liberal imperialism" in Darnton (2001).
32. Note by J. A. Thorne, 24 July 1938. GoI Home Political f. No. 22/36, 1938.
33. Note by V. T. Bayley, 8 August 1938. GoI Home Political f. No. 22/36, 1938.
34. Note by V. T. Bayley, 13 November 1940. GoI Home Political (I) f. No. 41/16, 1940.
35. Note by I. S. (signature unclear), 18 November 1940. R. Tottenham agreed in his note, 18 November 1940. GoI Home Political (I) f. No. 41/16, 1940.
36. IB Circular Memo No. 28/Int/39, 17 December 1940. GoI Home Political (I) f. No. 41/16, 1940.
37. Finance department (Customs) Notification No. 60, 30 November 1940. GoI Home Political (I) f. No. 41/16, 1940.
38. Note by V. T. Bayley, 27 January 1940. GoI Home Political (I) f. No. 41/16, 1940.
39. Extract from intercepted Communist Central Party Letter No. 2, 9 December 1940, on "Internal Politics." Sent by Madras CID to the GoI with Letter No. 202/C, 11 January 1941. GoI Home Political (I) f. No. 41/16, 1940.
40. For a detailed account of this volte face, see Manzer (2007).
41. An account of this by a critic of the party is that of Masani (1951). A later account by a well-known Indian historian sympathetic to the party can be found in Habib (1998).
42. Note by DIB, 2 August 1943. GoI Home Political f. No. 7/16, 1943.
43. One provision prohibited publication, sale and distribution of documents but not their printing, while another was unsuitable for the proscription of individual books. GoI Home Political f. No. 7/16, 1943.
44. Note by R. Tottenham, 17 August 1943. GoI Home Political f. No. 7/16, 1943.
45. Note by G. C. Ryan, Assistant Director IB, 27 October 1943. GoI Home Political f. No. 7/16, 1943.
46. DIB U.O. No. SA/518 PCJ, 30 November 1943. Home department official (signature illegible) note, 1 November 1943. GoI Home Political f. No. 7/16, 1943.
47. CPI C.C. Political Circular No. 99/43 titled "On the People's Publishing House," 23 June 1943. This 5 page typed circular was appended to an IB note, 2 August 1943. GoI Home Political f. No. 7/16, 1943.
48. Note by P. C. Bamford, 31 August 1938. GoI Home Political f. No. 22/36, 1938.
49. "A Note on Communist Activity in India," 15 February 1939. Reproduced in Chatterjee (1999, pp. 163–168).
50. "A Note on Communist Activity in India," 15 February 1939. Reproduced in Chatterjee (1999, pp. 163–168).
51. Council of State Debates, Vol. II, No. 5, 23 September 1935.
52. Sajjad Zaheer, interviewed by H. D. Sharma on 4 December 1969, Nehru Memorial Museum and Library (New Delhi), Oral History Transcripts, p. 32. Zaheer was an Oxford educated Indian writer and committed Communist.
53. Carritt (1986, p. 126).
54. Carritt (1986, pp. 129–130).
55. Note by J. M. Ewart, 17 November 1936. GoI Home Political f. No. 41/13, 1936.
56. Chandavarkar (1997).
57. Habib (1998, p. 23).
58. Popplewell (1995).

REFERENCES

Adhikari, G. (Ed.). (1971). *Documents of the history of the Communist Party of India* (Vol. 1, 1917–1922). New Delhi: People's Publishing House.

Bairathi, S. (1987). *Communism and nationalism in India: A study in inter-relationship, 1919–1947*. New Delhi: Anamika Prakashan.

Barrier, G. (1974). *Banned: Controversial literature and political control in British India, 1907–1947*. Columbia, MO: University of Missouri Press.

Carritt, M. (1986). *A Mole in the Crown: Memoirs of a British Official in India who worked with the Communist Underground in the 1930s*. Calcutta: Rupa.

Chandavarkar, R. (1997, Spring). From Communism to "Social Democracy": The rise and resilience of Communist parties in India, 1920–1995. *Science and Society*, *61*(1), 99–106.

Chatterjee, P. (1993). *The nation and its fragments: Colonial and postcolonial histories*. Princeton, NJ: Princeton University Press.

Chatterji, B. (Ed.). (1999). *Towards freedom: Documents for the movement of independence in India (1938, part I)*. New Delhi: Indian Council for Historical Research and Oxford University Press.

Darnton, R. (2001). Literary surveillance in the British Raj: The contradictions of liberal imperialism. *Book History*, *4*, 133–176.

Gopal, S. (Ed.). (1994). *Selected works of Jawaharlal Nehru* (Series 2, Vol. 16, Part II). New Delhi: Jawaharlal Nehru Memorial Fund.

Habib, I. (1998, May–June). The left and the national movement. *Social Scientist*, *26*(5/6), 3–33.

Manjapra, K. (2010). Communist internationalism and transcolonial recognition. In S. Bose & K. Manjapra (Eds.), *Cosmopolitan thought zones: South Asia and the global circulation of ideas* (pp. 129–177). New York, NY: Palgrave Macmillan.

Manzer, H. (2004). British measures against Indian Communists, 1934–37. *Proceedings of the Indian History Congress*, *65*, 514–521.

Manzer, H. (2007, November–December). Communist party policy during the imperialist war (1939–41). *Social Scientist*, *35*(11/12), 55–62.

Masani, M. R. (1951, March). The Communist Party in India. *Pacific Affairs*, *24*(1), 18–38.

Moral and Material Progress Report 1932–33. (1934). London: India Office.

Petrie, D. (1972 [1927]). *Communism in India, 1924–27*. Calcutta: S Ghattak.

Popplewell, R. J. (1995). *Intelligence and imperial defence: British intelligence and the defence of the Indian empire, 1904–1924*. London: Frank Cass.

Sethi, D. (2019). *War over words: Censorship in India, 1930–60*. Cambridge: Cambridge University Press.

INDEX