ROUTLEDGE FOCUS

SOCIAL MEDIA AND HATE

Shakuntala Banaji and Ramnath Bhat

ROUTLEDGE Focus

Social Media and Hate

Using expert interviews and focus groups, this book investigates the theoretical and practical intersection of misinformation and social media hate in contemporary societies.

Social Media and Hate argues that these phenomena, and the extreme violence and discrimination they initiate against targeted groups, are connected to the socio-political contexts, values and behaviours of users of social media platforms such as Facebook, TikTok, ShareChat, Instagram and WhatsApp. The argument moves from a theoretical discussion of the practices and consequences of sectarian hatred, through a methodological evaluation of quantitative and qualitative studies on this topic, to four qualitative case studies of social media hate, and its effects on groups, individuals and wider politics in India, Brazil, Myanmar and the UK. The technical, ideological and networked similarities and connections between social media hate against people of African and Asian descent, indigenous communities, Muslims, Dalits, dissenters, feminists, LGBTQIA+ communities, Rohingya and immigrants across the four contexts is highlighted, stressing the need for an equally systematic political response.

This is an insightful text for scholars and academics in the fields of Cultural Studies, Community Psychology, Education, Journalism, Media and Communication Studies, Political Science, Social Anthropology, Social Psychology, and Sociology.

Shakuntala Banaji is Professor of Media, Culture and Social Change in the Department of Media and Communications at the London School of Economics and Political Science.

Ramnath Bhat is postdoctoral fellow at the International Centre for Advanced Studies in New Delhi and visiting fellow in the Department of Media and Communications at the London School of Economics and Political Science.

Routledge Focus on Communication and Society
Series Editor: James Curran

Routledge Focus on Communication and Society offers both established and early-career academics the flexibility to publish cutting-edge analysis on topical issues, research on new media or in-depth case studies within the broad field of media, communication and cultural studies. Its main concerns are whether the media empower or fail to empower popular forces in society; media organisations and public policy; and the political and social consequences of the media.

Bad News from Venezuela
Alan MacLeod

Reporting China on the Rise
Yuan Zeng

Alternative Right-Wing Media
Kristoffer Holt

Disinformation and Manipulation in Digital Media
Information Pathologies
Eileen Culloty and Jane Suiter

Social Media and Hate
Shakuntala Banaji and Ramnath Bhat

For more information about this series, please visit: www.routledge.com/series/SE0130

Social Media and Hate

Shakuntala Banaji and Ramnath Bhat

Routledge
Taylor & Francis Group
LONDON AND NEW YORK

First published 2022
by Routledge
4 Park Square, Milton Park, Abingdon, Oxon OX14 4RN

and by Routledge
605 Third Avenue, New York, NY 10158

Routledge is an imprint of the Taylor & Francis Group, an informa business

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

Library of Congress Cataloging-in-Publication Data
Names: Banaji, Shakuntala, 1971– author. | Bhat, Ramnath, author.
Title: Social media and hate/Shakuntala Banaji and Ramnath Bhat.
Description: Milton Park, Abingdon, Oxon; New York, NY: Routledge, 2022. | Includes bibliographical references and index.
Identifiers: LCCN 2021047804 (print) | LCCN 2021047805 (ebook)
Subjects: LCSH: Social media—Psychological aspects. | Social media and society. | Online hate speech.
Classification: LCC HM742. B35 2022 (print) | LCC HM742 (ebook) | DDC 302.23/1—dc23/eng/20211001
LC record available at https://lccn.loc.gov/2021047804
LC ebook record available at https://lccn.loc.gov/2021047805

ISBN: 978-0-367-53727-2 (hbk)
ISBN: 978-0-367-53726-5 (pbk)
ISBN: 978-1-003-08307-8 (ebk)

DOI: 10.4324/9781003083078

Kamala, my mother, was a proponent of Buddhism, and a Vipassana teacher late in her life. It is she who led by example and convinced me that it is possible to believe in and work towards a more hopeful future. She was sometime sceptical of my choices but always believed in me. It is to her and her spirit that we dedicate this book.

Ramnath Bhat

Contents

Acknowledgements

This book has been made possible by the generous support of multiple individuals and groups in Brazil, India, Myanmar and the UK, including but not confined to the team of dedicated researchers who have undertaken fieldwork with us since 2018. We can never thank our interviewees and participants enough for their courage and integrity, and the ways in which they laid their experiences and wisdom open for scrutiny in order to enable a way forward. We would also like to acknowledge the dedication of our families and colleagues who supported us when the weight of the hate we encountered became unbearably heavy, and, especially Prof. Robin Mansell, who read and commented on draft chapters. We are in your debt.

Tables and figures

Trigger warning

Please be aware that this book discusses sensitive and traumatic subjects from genocide to racism, misogyny and Islamophobia as well as the effects these have had on individuals and communities; we detail these through evidence, both explicit and recounted. Readers might find these triggering for a number of reasons, particularly if you have also experienced discrimination, dehumanisation or online hate.

1 Introduction

On 23 February 2020, organised far right vigilante mobs targeted Muslims in the northeast of India's capital Delhi. Even as Delhi's assembly elections began on 8 February 2020, multiple BJP and other far right Hindutva leaders held rallies and live-streamed videos inciting violence against Muslims and Dalits. During the ensuing pogrom, vigilantes streamed and posted videos of violence that were widely shared, celebrated and defended.[1] Facebook whistle-blower Sophie Zhang has shown that Facebook employees had red-flagged several accounts as instances of Coordinated Inauthentic Behaviour (CIB). However, predictably – and *deliberately* – Facebook failed to act. When it did, much later, it was already too late. The mutilated and tortured bodies of mainly Muslim victims lined the streets and were stuffed into the drains of northeast Delhi. False narratives about the causes of the violence circulated on WhatsApp and via other social media and were amplified by mainstream media. Accounts that were overtly violating Facebook's Terms of Service by inciting violence were linked to prominent political leaders from the ruling party – the BJP.[2] Facebook as a company had decided where its loyalties lay.

This episode helps illustrate two key approaches that we bring to our work on social media and hate and to this book. First, we insist on the need to locate discrimination, incitement and hate speech historically within specific socio-political, economic and cultural contexts. For example, the 2020 Delhi pogrom cannot be fully apprehended without understanding the preceding local resistance led primarily by Muslim women (epitomised by the women of Shaheen Bagh[3]) against the discriminatory 2019 Citizenship Amendment Act. The anti-Muslim violence was a culmination of the BJP's campaign to polarise Delhi by spreading propaganda and disinformation. Approximately two months after the pogrom, Facebook invested nearly six billion US dollars to acquire just under a 10% stake in Jio Platforms[4] (a tech subsidiary of Reliance Industries owned by Mukesh Ambani, PM Modi's close associate and the fifth richest man in the world). With contexts such

DOI: 10.4324/9781003083078-1

as these informing our analysis of Brazil, India, Myanmar and the UK, we seek to provide a sense of what Massey (2005) calls 'power geometries' – relations of power that result in societies arranged in different kinds of hierarchies. Studying hateful content production and ways of reducing it without attention to power geometries is a self-defeating endeavour. The ongoing trauma which underlies the sedimentation and rearrangement of these geometries through discrimination and violence are key subjects in our chapters.

In this introductory chapter, we present a concise, historicised and critical review of research on social media and hate. The chapter also incorporates a critical review of methods used to define, delimit and understand these phenomena and possible ways to ameliorate them. We elaborate a theoretical framework that broadens the scope of investigation into online and offline far right activity, abuse, threat, discrimination, prejudice and dehumanisation in peer-to-peer networks, apps, and platforms as well as on cross-platform applications (described collectively as social media). We do so by discussing a) user practices, attitudes and experiences; b) technological and social infrastructures within which social media operate; c) relationships between social media and other forms of media and communication (face-to-face, broadcast, print and so on) and d) a historicised account of socio-political contexts that create the conditions where social media activity can legitimise, contribute to, or be used to organise targeted, extreme and persistent state discrimination, social discrimination and citizen-on-citizen violence.

When we refer to infrastructures, we mean technical and cultural systems that create institutionalised structures, or a system of interlinked materials and ideas, binding individuals and groups to specific forms of conduct and subject positions (Bhat, 2020; Larkin, 2008, p. 6; Star, 1999, p. 330). These infrastructures include market conditions, as well as the national and international legal and regulatory regimes to which they are subject, and the social milieu within which users produce, receive, share and act on media content circulated with the aim of denigrating, discriminating against, dehumanising, threatening and violating individuals and communities. We argue that individuals' and groups' experiences of themselves and others are formed in part through historical processes and in part through iterative engagement with communication infrastructures. Intersections of identity (caste, race, gender and so on) strongly inflect the nature and outcomes of these experiences. Thus, through an enhanced form of 'listening' which draws on Spivak's (1988) theorisation of subaltern voice, our theoretical framework studies the dialectical relationship between communication infrastructure/technological affordances and the phenomenology of embodied, intersectional subjectivity in the context of hateful communication. Such a phenomenology includes not only a sense of self

and other (Merleau-Ponty, 2012) but also a sense of the world produced by the use of media (Banaji, 2017; Gray, 2020). For those at the receiving end of discrimination, violence and hate, it includes trauma, loss and/or the theorisation of their group experiences in ways that enable conscientised resistance and practice (Freire, 2000). Throughout the book we argue that analysis of social media use in the context of increasing incivility, bullying, authoritarianism, political violence and polarisation benefits greatly from the interrogation of the networked infrastructures – the ecosystems – within which social media use occurs.

While we were researching and writing this book, several interviewees questioned our decision to write about *online* hate speech when vigilantism and physical violence might appear to be more urgent priorities. The 2020 Delhi pogrom, the anti-Rohingya violence and the coup in Myanmar, the racist, homophobic and transphobic attacks in the UK and the violent suppression of feminists, gay people, Indigenous and Afro-Brazilian communities challenging Bolsonaro cropped up repeatedly during our research. This suggests that, as Kishonna Gray has demonstrated with regard to digital games (2020) and Ruha Benjamin (2019) has demonstrated with regard to coding, algorithms and AI, rather than constructing an artificial binary between the offline and the online or the digital and the real, these domains are inseparable, conceptually and materially.

Online and offline discrimination, harassment and violence are part of the same constellation and act on each other on the local, national and international levels (subject, we argue, to different power geometries). Whether one is critical of the disproportionate attention given to the 'online' or convinced about the disruptive potential of the internet to change all aspects of human life, a common tendency is to frame 'the Internet' as a fundamentally ahistorical phenomenon that acts on and affects society, but not vice versa (Morozov, 2013). Our approach does not deny the speed and specificity of particular forms of online harassment, dehumanisation, incitement and hate speech, but we follow Banaji and Buckingham (2013) in arguing that 'the online' is itself shaped by and part of an individual and collective psychic and politico-historical experience that is always also 'offline'.

Whether this is termed a dialectic or not depends upon the amount of agency (Banaji, 2017) one attributes to sociotechnical systems and infrastructures such as the Internet. The fast-paced heartbeats and physical anxiety that many people we spoke to evince on seeing rape threats aimed at them or their children on TikTok, Messenger or Instagram, the depression and anger they experience when their private and personal lives are targeted by trolls or doxers because of their liberatory stances on issues of identity or their concern for circulating factual evidence, cannot be disconnected from the persistent discriminatory comments aimed at them, their children

and their communities. Nor can hate online be dissociated from the street harassment, stalking, physical intimidation, police brutality, legal injustice and social exclusion that many social justice activists, women, LGBTQIA+, disabled, Indigenous and minority ethnic or religious citizens, refugees and asylum-seekers, recount as the quotidian backdrop to their social media use (Awan, 2016; Elareshi, 2019; Felmlee et al., 2018). The intensity of these experiences is often further enhanced by intersections of identities which provide convenient targets for the politics of the far right that has swept across the globe since 2014.

For these reasons, our research centres the experiences and views of social media users whose communities are directly impacted by online hate. This focus allows us to trace how harmful content including legally provable hate speech, hate crimes, threatening content, inciting content and discriminatory disinformation on social media, emerge via an interplay of social and technological infrastructures and, equally important, via an ideological nexus between social media and other forms of communication, including face-to-face communication, urban and rural spatial and material practices, and mainstream broadcast and print media.

Methodologies underpinning our study

While we draw on political economy traditions to explain corporate decisions, the main focus of our book is on audiences, the distribution of information and everyday media practices and cultures that are historically, ideologically and contextually located (cf. Banaji, 2011; Parks & Starosielski, 2015). Between 2018 and 2021, we and our research collaborators Zico Al-Ghabban, Marina Navarro Lins, Nihal Passanha and Letyar Tun interviewed more than 100 individuals and conducted 20 focus group discussions and 15 expert interviews in Brazil, India, Myanmar and the UK. This book is based on in-depth analysis of these transcripts and a background textual analysis of 3000 social media posts drawn from Facebook, WhatsApp, Instagram, Twitter and Tik-Tok. Some of this material was provided by our interviewees. Other materials come from online research for this book and for our WhatsApp Vigilantes research (2019).

Our research methods are qualitative and grounded in post-structuralist and interpretivist epistemological traditions that pay attention to shifting patterns in people's expressions and understandings of their own and others' identity, as well as to the gaps between individual and collective memories, lived experiences and recorded events. Such approaches require considerable self-reflexivity and attention to situated knowledge production (Gupta & Ferguson, 1997; Haraway, 1988; Visveshwaran, 1996). We acknowledged with our interviewees our own positionality, vulnerability and privilege. Apart from the sensitive nature of the topic at hand, the second half of our

research was conducted amidst a global pandemic, where we and our informants were facing illness, bereavement and other collective or individual traumas. Given these circumstances our interviews and analysis were founded on empathy, allowing interviewees to speak at length by creating a shared narrative space (Douglas, 1985; Zinn, 1979). Further, our interviews were designed and carried out with a sensitivity to trauma-informed contexts in order to ensure that our research did not reproduce trauma that had already been experienced (Favaro et al., 1999; Dyregrov et al., 2000; Thomas et al., 2019). In line with a desire to respect the autonomy of our participants as experts and co-constructors of the research, we used pseudonyms or real names where requested. They – and we – fully recognise the painful modalities between these choices of partial erasure or increased visibility linked to increased risk. As our chapters show, however, at no point did we wish our voices to be heard above theirs. Thus, including extended excepts from interviews and focus groups was a deliberate methodological choice.

Given the myriad possibilities of a phenomenological approach, we examine how social media-related technological and social changes are intertwined in each of our four case study countries – Myanmar, Brazil, UK and India. In doing so, we do not claim the nation-state as the sole valid unit of such analysis (Amelina et al., 2012). Arising from a post-war context, this kind of methodological nationalism assumes the existence of universal categories that can then be 'tested' in various countries. Our framework treats every empirical investigation in a different society as constitutive of the universal category – which by definition must remain an unfinished project. In other words, as Kuan-Hsing Chen puts it,

> to do area analysis is not simply to study the object of analysis through a process of constant inter-referencing. . . [rather], relativizing the understanding of the self as well as the object of the study is a precondition for arriving at different understandings of the self, the Other and world history.
>
> (2010, p. 253)

Legal instruments and international principles addressing hate speech

Although hateful propaganda, discriminatory disinformation and hate speech existed long before the Internet, the amplification of particular forms of hate on social media deserves special attention. As Alkiviadou argues:

> Firstly, the sheer number of users of such networks on a global scale results in the need to pay particular attention to this digital vehicle. Secondly, social networks are used by individual users but also by

> organised and semi-organised groups to promote hateful rhetoric and target the victims of such rhetoric. Thirdly, social networks come with some kind of content regulation which must be assessed for purposes of ascertaining whether or not and, if so, the extent to which this regulation contributes to the effective tackling of online hate.
>
> (2019, p. 20)

No universally accepted definition of hate speech exists in international law, even though both hateful content and its consequences have been all too clear, particularly during the 1930s and 40s in Europe, and with the spread of social media and smart devices since 2000. The International Covenant on Civil and Political Rights (ICCPR) adopted by the United Nations in 1966, specifically Article 20,[5] states that 'any advocacy of national, racial, or religious hatred that constitutes incitement to discrimination, hostility or violence shall be prohibited by law'. Amongst other legal instruments and international principles related to hate speech, The International Convention on the Elimination of all forms of Racial Discrimination (ICERD) came into effect in 1969, limiting its definition of hate speech to speech about race and ethnicity, and, in another crucial difference, expanding the scope of liability in terms of disseminating hate speech, in contrast to the ICCPR which limited liability to *proof of intent to cause harm*.

In the Genocide Convention of 1951[6] hate speech is limited to public incitement of genocide, and the Convention on Elimination of all forms of Discrimination Against Women (CEDAW) of 1981[7] focuses on discrimination and violence against women. In 1997, the Council of Europe's Committee of Ministers in its Recommendation[8] No. R (97) 20, defined hate speech as

> all forms of expression which spread, incite, promote or justify racial hatred, xenophobia, anti-Semitism or other forms of hatred based on intolerance, including: intolerance expressed by aggressive nationalism and ethnocentrism, discrimination and hostility against minorities, migrants and people of immigrant origin.

Amongst many definitions, this is the one that appears to encompass the widest range of constituents and circumstances and is the one we return to as we proceed.

Unfortunately, by and large, the law has not been able to address hate speech against the most vulnerable. Indeed, '[t]he places where the law does not go to redress harm have tended to be the places where women, children, people of colour and poor people live' (Matsuda, 1989, p. 2322). In Australia, the United States and Canada, this inability of the law to bring legal

sanctions against hate speech and its attendant discriminatory or violent material effects historically has stemmed from a commitment to preserving an extreme and problematic version of freedom of speech. We argue that this has resulted not only in physical harm but also in longstanding psychic trauma to multiple communities across these countries.

The emphasis of the North American approach lies in assessing speech in terms of the likelihood that 'speech acts' will result in *clear danger to life and property*. However, in other parts of the world, speech is regulated not just regarding the likelihood of harm but also in terms of whether the intrinsic content is objectionable (Gagliardone et al., 2015; Pohjonen & Udupa, 2017). In a review of regulatory and legal approaches to hate speech in Latin America, Hernández (2011) argues that racist speech about Afro-descendants is ubiquitous. Sickeningly, they are commonly likened to animals and, in particular, to monkeys. Hernandez argues that

> [t]hese perspectives about Afro-descendants are so embedded in the social fiber of Latin American societies, that Afrodescendants' subordinated status in society is viewed as natural and logical [while] . . . the historical notion that "racism does not exist" in Latin America disinclines those unaffected by hate speech to acknowledge the harms it causes marginalized groups.
>
> (ibid.: 820)

Since those 'unaffected' by racist hate in Latin America tend to be white or white-passing and from non-Indigenous populations, the ties between struggles for Indigenous rights, racial justice and against hate speech are linked.

Given the dominant (read: white) groups' access to multiple channels of representation via politics, media and religion – and their use of these channels for the repeated derogatory positioning of Afro-descendants in the social hierarchy, this racist hierarchy itself has become naturalised. The resulting inequality is further legitimised to different degrees in different Latin American nations by the adoption of essentialist assumptions (Hall, 1997) about the subordinate status of minoritized Indigenous and Afro-descendant groups in education, workplaces and justice systems. Showing existing regulatory approaches to be profoundly inadequate, Hernandez argues for the need to bring about legislation that can specifically address hate speech:

> Because of its great symbolic power, a ban on hate speech can easily become a symbol that is an end in and of itself rather than part and parcel of an overarching policy against racism. It is thus centrally important to enact hate speech legislation that focuses on its anti-discrimination role

> rather than viewing it as an antidefamation inspired law or simply as a dignitary harm. Incorporating civil as well as criminal code provisions would also enhance the anti-discrimination role of hate speech legislation. Restricting hate speech legislation to the criminal code context, as is done in many jurisdictions, may limit its efficacy for a number of reasons. Entrusting the enforcement of the criminal law to public authority risks having the law undermined by the complacent inaction of public officials who may harbor the same racial bias as the agents of hate speech. This is a particular danger in Latin America, where police officers are consistently found to discourage Afro-descendants from filing racial discrimination complaints, and are often the perpetrators of discrimination and violence themselves.
>
> (ibid., p. 829)

Here, we ask our readers to note the role and implication of the police in racist violence and in suppressing redress for those who have experienced it – a phenomenon which will be seen repeatedly in the chapters on Brazil, India and the UK. Similarly, a recent report released by the International Dalit Solidarity Network (Shanmugavelan, 2021) argues that both offline and online caste-hate speech needs to be grounded in the historical contexts of an Indian subcontinent shaped by caste hierarchies.

Shanmugavelan rightly notes that the dominant castes in India and the diaspora within and outside institutions perform 'castelessness' that serves to conceal the brutal history of caste oppression and provides fertile ground for caste-pride and caste-hatred. Although caste-based discrimination and the daily occurrence of caste-hate speech in everyday life in mainstream media and social media has been acknowledged by various international bodies, and there is some limited but largely symbolic support for anti-caste activities evinced by social media companies, there is no clear set of legal principles preventing caste-hate speech. Thus, Shanmugavelan argues that 'it is essential that caste-hate speech is recognised as . . . a distinctive form of hate speech – and that Dalits are included in actions to mitigate caste-hate speech online and offline, at every level' (ibid: 27).

In a 2015 report on hate speech and incitement to hatred against minorities in the media, UN Special Rapporteur on Minority Issues, Rita Izsak,[9] emphasised the need to distinguish between different types of expressions: a) expressions that constitute an offence that can be prosecuted criminally; b) expressions that may justify restriction and civil sanctions; and c) expressions that raise concerns about tolerance, civility and respect for others. In other words, Izsak views hate speech on a much wider spectrum than the current narrowly defined legal category. Izsak goes on to argue, correctly in our view, that non-legal and social responses to hate speech should be given

as much attention and discussion as legal responses. With this in mind, we move to discussing scholarly efforts to define hateful, violent and discriminatory speech.

Conceptual approaches to and empirical research on online hate

'Naming' and its discontents

In an attempt to rescue free speech from the encroachment of ill-conceived and misused hate speech legislation, Susan Benesch proposes the concept of 'dangerous speech' arguing that 'when an act of speech has a reasonable chance of catalysing or amplifying violence by one group against another, given the circumstances in which it was made or disseminated, it is dangerous speech' (2013, p. 1). Benesch provides five variables to determine the *degree of dangerousness* involved: (i) the speaker, who is much more likely to commit successful incitement if he or she has some form of pre-existing influence or authority over an audience; (ii) the audience, the more fearful it is, the more vulnerable it is to incitement; (iii) the speech act itself, by way of the use of certain rhetorical devices, such as the 'accusation in a mirror' strategy, persuading the audience that they are going to be attacked; (iv) the social and historical context and (v) the mode of dissemination.

Benesch's work brings scholarly attention to the social and historical contexts within which hateful and discriminatory communication takes place as well as the distribution and infrastructural aspects which impact speech acts. However, her focus on speech with a reasonable chance of catalysing or amplifying *violence* excludes the amplification of discrimination and structural inequality which is often the aim and result of what we have been calling 'hate'. When discussing racism in the US context, legal and critical race scholar Patricia Williams (1987, p. 129) argues for a view of ' . . . racism as a crime, an offense so deeply painful and assaultive as to constitute something I call "spirit murder" '. Benesch's re-labelling of 'hate speech' as 'dangerous speech' and her definition, while genuinely useful in identifying speech acts geared towards lynching, pogroms and genocide, does not account for the collective trauma and psychic harms of constant belittling, maligning, insulting and exposure to humiliation and abuse in cases where violence is not necessarily imminent. Nor does it address how deep-seated prejudices attendant upon such dehumanisation influence discriminatory practices in carceral systems, housing, land acquisition, employment, schooling, higher education, the culture industries and so on.

Other authors who object to the 'reified category' or 'thick concept' of hate speech are Pohjonen and Udupa (2017), who propose the concept of

'extreme speech', arguing correctly that hate speech is not a binary (hate/not hate) but lies along a spectrum. However, in their legitimate concern to protect the free speech of actors working to critique widely accepted social and religious practices, the authors appear to divert attention from the plight of minoritised communities at the receiving end of discrimination and violence whipped to a frenzy by hateful communication. Critical satire, jokes and open critique aimed at oppressive practices and aimed to draw attention to injustice, unfairness or imbalance and inequity (between and within communities), are *not* on the spectrum of hate speech, even if they are disingenuously mis-characterised as such by malicious politico-legal or religious regimes.

Moving away from the concept of hate speech simply because hate speech laws are deployed *in bad faith* against legitimate actors engaging in critique or dissent has further implications for the recipients and targets of hateful communications. In the context of structural discrimination and psychic harms, scholarly work on disablist hate speech reminds us that fear of impairment is projected onto the 'other'. As Burch (2018) argues

> [t]he use of 'parasites' as a means of identifying and marking out disability is supported by the relationship between welfare and employment, to which the first is presented as inferior to the latter. Making this connection, one Reddit user argues that 'you are a parasite on the productive class', thus confirming that the disabled figure is not only unproductive, but burdensome to those who are productive.
>
> (ibid., p. 401)

Benesch's work serves as a useful framework in theorising another widely held and common instance of hate speech known as 'Islamophobia' (Allen, 2010). There is scope to use the framework to interrogate the actors involved, their influence, status and legitimacy in particular societies and their motivations in specific social and historical contexts.

Quantitative textual analysis can yield interesting results about the virality, spread and ubiquity of Islamophobic tropes across specified online populations. Aguilera-Carnerero and Azeez (2016) analyse more than 10,000 tweets with the hashtag #jihad to show how Islamophobia has spread globally via the misrepresentation of Muslims and Islam in the post-9/11 mediascape. They report two important findings. First,

> [W]hen used by Islamophobes the meaning of the word 'jihad' becomes associated with ideas of 'violence' and 'war'. From the data, we could not even state that they are talking about a 'holy war' because that would imply an ulterior religious motivation, but many of the tweets contain

> information only about assorted felonies and misdemeanours. It is not a unique phenomenon that a religious term transcends the religious lexical field and becomes part of the daily vocabulary of any language (for example the terms 'apocalypse' or 'purgatory' from Christianity), but most of them retain their original meaning or a part of it. In this sense, the process the word 'jihad' undergoes is a different one as there has been a lexical conversion that serves the speaker's intention.
>
> (Aguilera-Carnerero & Azeez, 2016, p. 30)

And second: 'Far from countering any of the clichés previously attributed to Muslims and Islam the corpus, on the contrary, reinforces and expands existing negative stereotypes' (Ibid: 30). Drawing on literature referring to the core ideology of orientalism leading to Islamophobia at both the institutional and interpersonal levels (c.f Abu-Sway, 2005; Said, 1978), the authors show that online Islamophobia is deeply connected to misrepresentations and stereotyping of Muslims by western media since the early 2000s. For instance, in a study about representation of British Muslims in nearly 1,000 UK newspapers from 2000–2008, most of the coverage was found to focused on Muslims as threats in terms of terrorism, or differing values, or both, in terms of Muslim extremism (Moore et al., 2008).

Using British legal definitions of hate speech as expressions of hatred toward someone on account of that person's colour, race, disability, nationality (including citizenship), ethnic or national origin, religion, gender reassignment, or sexual orientation, a parallel study conducted on Islamophobia using a qualitative analysis of 100 different social media pages, posts and comments found nearly 500 instances of online hate speech directed against Muslims (Awan, 2016). Word cloud frequency was deployed to examine key words depicting Muslims in an overtly prejudicial way (such as 'Paki' or 'Muzrats'). Awan posits a typology as a starting point for a framework to analyse Islamophobia as expressed by users on Facebook. In line with our findings, his typology includes 'opportunists' who post hate speech and incite violence against Muslims immediately after incidents such as those involving Daesh; 'deceptive' users creating fear by posting about false events to intensify hate against Muslims; 'fantasists' who set up Facebook pages to fantasise over Muslim deaths, often making direct threats to Muslims; and finally, 'systematic producers and distributors' of Islamophobic content.

In a similar vein, studying hate speech in Kenya, Busolo and Ngigi (2018, p. 43)

> interrogate the prevalence and development of hate speech over time, investigate the perpetrators of hate speech and the targeted groups,

> critically analyse the consequences of hate speech, dissect the freedom of speech vs. the protection from hate speech, highlight various challenges in curbing hate speech and reflect on strategies and methods of curbing hate speech being used by various agencies.

Citing McGonagle, they argue that

> [t]here are different types of hate speech perpetrators. There are offenders by conviction . . . people with clear intention of engaging in hate speech. On the other hand, incidentalists are people who may post information without thinking about the consequences, but when legal or social repercussions arise, they tend to be shocked.
>
> (Ibid: 45)

Another study analyses comments on popular Slovenian news websites to examine the unique factors that motivate different kinds of perpetrators of hate speech (Erjavec & Kovačič, 2012). The authors argue that some users are 'soldiers' who typically belong to political parties and systematically vilify users identifying with the opposition, while yet other users serve as 'watchdogs' who use hate speech to draw attention to social problems. The term 'soldiers' resonates with some of our own previous work.

We conducted research in four states of India to investigate the role of WhatsApp in vigilante mob violence against minority communities in the 2018–19 period (Banaji & Bhat, 2019). Using focus group discussions and in-depth interviews with nearly 300 users and analysing over 1000 WhatsApp messages, we found that privileged users – often Hindu and upper caste, with sympathies for the ruling Bharatiya Janata Party (BJP) – justified their sharing of hate speech against Christians and Muslims, Pakistanis and others on the basis of nationalism, civic duty and the credibility of the person who forwarded the misinformation. Importantly, these powerful and privileged groups were found to be the ones most often involved in systematic production and distribution of hateful disinformation and had significant technical digital literacy skills as well as high levels of formal education. Some members of vulnerable minorities, on the other hand, in WhatsApp groups for professional or family reasons, became complicit in spreading hate by forwarding the messages they received to appear compliant, often without having the time or energy to engage fully with or even read through the hundreds of banal prejudiced and derogatory memes, GIFs, morphed images, quotes or false stats. The targets and 'victims' of these messages (amongst these rural or poor urban women, Dalits, Adivasis, Muslims, Christians and young political activists on the left, feminist or LGBTQIA+ scenes) while heterogenous, and often lacking in social capital

were often far more alert to, critical of and attuned to discriminatory speech and practices in everyday life, and consequently either less able or less willing to pass on political misinformation received online. We return to some of these testimonies in Chapter 4.

Maya Mirchandani (2018) in her paper on Hindu majoritarian online hate speech targeted at Muslims in India builds on the notion of offence and 'hate spin' as theorised by Cherian George (2016). She summarises his argument:

> The Islamic far right in countries such as Pakistan, Indonesia and the Maldives, the Christian far right in the US and Western Europe, the Buddhist far right in Myanmar, and the Hindu far right in India, are feeding on people's sentiments of being "offended" . . . . Cherian George makes the case that political groups selectively mobilise genuine religious devotion to manufacture both *offense* and a sense of being offended- or *offendedness*.
>
> (Mirchandani, 2018, p. 2)

She also summarises George's perspective on the media 'caught up' in the amplification of *offendedness*:

> [T]he main objective of hate speech is met when the support base is widened, a divisive narrative is created, and people are mobilised around a political agenda. The media, meanwhile, are caught in reporting incidents when they happen, or else inadvertently serving as a vehicle for politicians who use hate speech as a tool for identity politics. In the process, the media often lose sight of the manufactured quality of hate spin, especially where the line between hate speech and free speech are blurred.
>
> (ibid.: 3)

While George's arguments corroborate findings in some of the contexts studied here, they suggest far too meagre a view of the role of mainstream media in contributing to discrimination and violence linked to hate speech. Mainstream media outlets have political, ideological and economic links with far-right groups which then influence the ways in which mainstream media discourses retain an intertextual relationship with social media and interpersonal (online and offline) discourses. For instance, mainstream media discourses naturalise the validation of majoritarian anxieties – the narrative of false victimhood – while hate speech on closed messaging apps such as WhatsApp 'picks up the baton' by more explicitly targeting minoritised communities. Further, George seems to imply that it is

offendedness or *offence* alone that is manufactured, but when amplified by media becomes hate spin. More often than not, however, existing hatred entwines with political subjectivities and affects about deserving and undeserving citizens – such as feelings and discourses of disgust, contempt, racism, misogyny, homophobia and so on – and masquerades as *offendedness* which the mainstream media uncritically or complicitly, amplifies.

Mirchandani's own work (2018) bears out this more concerning thesis. She posits a theoretical framework that works in two parts to explain the emergence of majoritarian violence and hate speech in India. The first draws on Appadurai's (2006) *Fear of Small Numbers*, wherein Appadurai outlines the notion of 'predatory identities', premised on the extinction of proximate social categories that emerge especially out of pairs that have often experienced long histories of contact, mixing and stereotyping of each other. The second part of Mirchandani's framework deals with when and what turns majoritarianism from affective and discriminatory to actively violent. Building on scholarship that distinguishes between radicalism (which could comprise wide-ranging hostilities to the political status quo) and radicalisation (which includes various 'push' factors that turn individuals and groups towards violence for a cause and against a group defined as 'others'), coupled with Arendt's (1963) notion of the 'banality of evil', Mirchandani argues that these existing realities of violent predatory identities need to be taken into account by various actors working on counter-terrorism and prevention of violence. While this is a fair and practical point, of course, there is also further room for anxiety given that many of those involved in law enforcement, the justice system and counter-terrorism often subscribe unreflexively to just such predatory identities while attributing *radicalisation* only to the groups defined as 'other'.

Although Mirchandani's framework is useful in thinking about hate speech, there are at least two broader questions. First, if Hindutva in India and the diaspora can be seen as an ideology based on emerging 'predatory identities', how are we to reconcile this with pre-existing caste-based divisions? From a critical caste perspective, the 'majority' dominant castes (across religions) actually constitute less than 30% of the Indian population, whereas Dalit, Bahujan and Adivasi groups constitute more than 65% (Aloysius, 1997; Ambedkar, 1989). Second, are all individuals and groups equally susceptible to radicalisation and in the same ways? If not, what explains the differences? Both of these questions indicate the need for theoretical definitions of hate and radicalisation to be situated within specific historical and political contexts.

A typological approach to hateful communication

There is scholarship on the potential targets of hate speech and other forms of violent discrimination and dehumanisation. Surveying nearly 1000 young adults in the US, Costello et al. (2017, p. 588) argue that online hate speech

> differs from cyberstalking or cyberbullying in that hate materials express hatred or degrading attitudes *toward a collective* instead of an individual in isolation. Thus, hate materials express extreme attitudes devaluing others because of their religion, race, national origin, sexual orientation, gender, gender identity, ethnicity, or some other characteristic that defines a group.

This is particularly important at a time when it has become commonplace to dismiss online aggression as an ubiquitous feature of 'online culture' that is said to affect *all* social media users equally if posts are made about contentious topics; merely being visible online can invite invasions of privacy.

Our work suggests that attempts by oppressed and minoritised communities to critique and resist their oppression are frequently miscategorised in nations/institutions where the powerful majority racial, ethnic or political group has cultivated a sense of victimhood against a set of 'others'. Subaltern resistance is censored or sanctioned *as hate speech* simply because it is aimed at a collective (albeit an oppressive one with a history of political aggression and discrimination). To complicate matters further, the majoritarian community who are now oppressors might once have belonged to a community who were themselves subjected to injustice and oppression and the minoritised community which is collectively subjected to violence and/or discrimination may itself subject some of its own members (women, LGBTQIA+, non-conformists) to extensive and historically embedded forms of discrimination and violence.

Some of this complexity can be witnessed in North American and European contexts in the chilling effects of the use of charges of antisemitism levelled at those critiquing the violence of the Israeli state and settlers. In other cases, in India, for instance, oppressor communities appeal to laws on free speech to support their right to malign others publicly while also levelling charges of religious hatred and offense against (secular or Muslim) comedians who draw attention to the lack of care for human life which motivates Hindu or upper caste 'cow-protection' lynch mobs. Both of these forms of *malign censorship* are rampant across the globe and remain powerful hypocrisies cultivated by supporters of North American and British conservatism/Republicanism and the alt-right.

The complexity and importance of offline environments in the generation of political hate, propaganda, disinformation and misinformation that is knowingly and systematically targeted at particular groups and individuals online are typically ignored by those who assume the separateness of these domains. Costello et al.'s premier assumption that '[t]he extent to which individuals' online activities bring them into virtual contact with motivated offenders affects their likelihood of victimization' (2017, p. 589) is partly banal and partly misleading. Their assertion is that:

> [u]sing a modified version of RAT [Routine Activity Theory], we find robust evidence that online habits, such as utilizing numerous SNS platforms and visiting hostile online environments, are related to being targeting by hate online. Indeed, SNS usage has the strongest relationship with being targeted by hate; avid users being nearly 6 times as likely to be targeted.
>
> (2017, p. 597)

This conclusion has problematic repercussions. While the intention might be to support interventions that could reduce the extent to which people are targeted with hate online, its refusal or inability to consider the socio-political factors connecting recipients and producers of online hate means that interventions will target the wrong factors and are more likely to fail.

Bikhu Parekh's (2012) work on conceptualising hate speech emphasises three key characteristics of hate speech: Directed against easily identifiable individual(s) based on an arbitrary and normatively irrelevant feature; stigmatising the target group by ascribing to it qualities widely regarded as highly undesirable; and a target group is viewed as an undesirable presence and as a legitimate object of hostility. Building on this, Gelber and McNamara (2016) interviewed 101 individuals from indigenous and minority ethnic communities in Australia. They note the distinction between the *constitutive* and *consequential* harms of hate speech (Maitra & McGowan, 2012) to observe a *spectrum of hate speech* such as verbal and symbolic epithets, exclusion, negative stereotyping, transmission of racism, threatening and harassing behaviour and so on. Based on reported examples recollected from the interviewees, the authors find a wide range of constitutive and consequential harms such as: Feelings of being hurt and upset, a resulting fear, fear leading to a sense of paralysis, disempowerment, withdrawal from spaces which offer opportunities for redress, silencing and/or being rendered mute, silence as an avoidance tactic, feeling dehumanised and violated, feeling anger and frustration and deciding to dis-identify from their own identities as a protective mechanism.

Scholarship that draws on the voices and lived experiences of those who experience hate in the context of discrimination and violence (Sethi, 2018) rather than on regulatory approaches, policies, laws, the hate speech itself or those who propagate or amplify hate speech is an important addition to the research literature. Gelber & McNamara explain that their methodology favoured a bottom-up understanding of the harms of hate speech which allowed for a much more capacious understanding than merely focusing on the threat of immediate violence:

> reflections shared by the interviewees confirm that public racism in Australia occurs in face-to-face encounters and general circulation in targeted communities. These two types of hate speech were not experienced as qualitatively different in terms of seriousness or harmfulness. [Indeed. . .] public hate speech is frequently experienced as an attack on worth and dignity. As discussed by critical race scholar Delgado (1993), harms which are non-physical and do not fall under immediate danger are . . . enduring and not ephemeral.
>
> (2016, p. 336)

As we move through the cases in this book, we will develop further the notion of hate speech as an attack on worth and dignity but also on the right to have rights, and argue that different intensities and modalities of hate have similar seriousness and harmfulness.

Analytical framework for this book: Theory and typology

Theoretical positioning

As research in the fields of cultural studies, media and communications, political economy, Science and Technology Studies and infrastructure studies has shown, technological developments do not only act upon society. Technological developments are also *imagined*, *engineered, enacted* and *acted upon* in various ways, shaping communicative innovations as well as the ways in which they operate in specific circumstances (Castells, 1996; MacKenzie & Wajcman, 1985; Mansell, 2012; Williams, 1974; Winner, 1980). After the International Telecommunication Union (ITU) framed telecommunications as a major 'engine for economic growth' in its 1984 Maitland Report (Chakravartty, 2004), countries in the Global South liberalised their telecommunications sectors and attracted large investments towards wireless infrastructures. Alongside smartphones becoming cheaper as a result of a push in Chinese manufacturing in the early 2000s, these

countries invested in increased internet penetration, hoping to see growth in their GDPs. This political economic context is crucial in understanding how 'the Internet' has emerged as an agglomeration of material practices, infrastructures and discourses.

Our theoretical framework aims to disrupt any simplistic binarism that assumes a separation of online hate from offline history and politics. Both scholarly and corporate-technological attempts to discuss online hate include a disavowal and underplaying of the social, political and historical contexts that explain the specific circumstances under which some groups have systematically dominated other groups thereby supplying the grounds on which hate can be manufactured, rationalised, legitimised and normalised. Because of this disavowal of historical context (except, ironically, with regard to the historical spread of digital tools and technologies), hate speech and the individuals and groups involved (as perpetrators or recipients) of hateful communications appear to be treated as interchangeable actors, equal in the eyes of theory-making, law-making and policymaking. Our framework, on the contrary, treats online hate as an ecosystem consisting of various political (corporate, government and party ideological interests), technological (media, infrastructures, algorithms, AI) and social (identities, inequalities, histories of oppression and struggle) aspects. Rather than focusing on the speech itself, we rely on situating users' experiences in a historical and contextual setting to emphasise the systemic ways in which elements are interrelated to comprise what we call an ecosystem of online hate.

Infrastructures can be theorised as technical and cultural systems that create institutionalised structures and bind people together towards specific subject positions. Drawing on insights from a critical political economy of media and communications, we argue that these technical and cultural systems are themselves subject to power flows, including ones induced by economic relations between social media companies, the state and domestic/international large corporate entities. Our case studies illustrate this dynamic clearly, be it the economic dominance of the Myanmar military regime controlling Internet infrastructure or Meta/Facebook's consistent prioritisation of profits at the cost of widespread racist violence and genocide.

A clear pattern emerges from an analysis of the literature on online (and offline) hate speech, related concepts and issues such as violence and disinformation. Overall, efforts to address and define hate speech seem to be haunted by a concern to preserve the notional concept of free speech, while actually preserving *some people's freedom to express hate* at the expense of *other's freedom to live and thrive*. Whether these outcomes arise from regulatory approaches or the policies and terms of service of social media companies, whether it is explicitly mentioned or implicitly guiding research,

hate speech and free speech are often treated as abstract objects framed in binary opposition to each other, thus making it the apparent duty of different actors (scholarly, political, legal or corporate) to balance the two conceptually and in practice.

Barring a few exceptions, research approaches to hate speech seldom analyse the identities of those who face hate. The *intrinsic content* of communication that is objectionable or harmful seems to serve as a sufficient basis for working to prevent hate speech or for disavowing the harm it has done and is doing in favour of an abstract notion of free speech. Listening to the voices of those who directly face hateful communication and its attendant discrimination and violence can help acknowledge the legitimacy and theoretical weight that should be accorded to *affect* and *lived experience* in understanding the consequences of socio-political hate as a tool of power. *Listening* as both a theoretical and methodological framework can expand our collective understanding of what 'harm' means in the context of hate online. Such an expansion to include the ways in which local marginalised populations theorise their own experiences of being othered, excluded and dehumanised, and are silenced or conscientised into action, has political, legal and psycho-therapeutic implications. In addition to the steps being taken at present, such an expanded understanding of hate speech – which includes its embedding in histories of othering and the creation of difference in order to gain or maintain social position, economic profit and political power – may eventually open up new ways of countering the discrimination and violence attendant upon and surrounding hateful communication.

Our framework draws upon a variety of literatures that highlight people's experiences and affects around discrimination and hate faced on and offline. While there are phenomenological distinctions to be made between experiences and parasocial relationships in virtual environments (always nested within the material world) and those solely in the material world, interconnections between these spheres bear deeper examination. Social media and the online world has been mythologised in popular discourse as a disruptive technology that leaves everything in its wake irreversibly changed. Multiple studies refer to a 'digital age', to lives 'lived online' and to 'online worlds'. In our theoretical framing, what we perceive as 'the Internet' is governed by sets of protocols controlled by specific institutions (such as those allocating domain names and those negotiating intermediary rights). These protocols exert power, and influence the ways in which we use digital technologies (Galloway, 2004). Social media service providers generally prioritise and act faster when troubling incidents take place in close proximity to their parent corporations' geographical or imagined communities, ignoring or mischaracterising incidents of Internet-enabled

discrimination across much of the Global South and against disenfranchised groups in the Global North. The ways in which technology acts upon society and the ways in which technological developments themselves are socially shaped give rise to complex strategies and responses, complicated by states' and corporations' political and economic motivations. Given the complexity of these dynamic processes acting upon each other, infrastructure and (the phenomena of) online hate can be difficult to stabilise as the foundational objects of our research. We have chosen instead to locate our research in the phenomenological life of those who experience online hate as part of *who they are* and *what they do*. Listening attentively to interviewees' accounts of hateful experience and embodied subjectivity enables us to deconstruct the peculiar and specific ways in which history and online ecosystems intersects with hate. Deconstructing the foundational objects of research opens up new ways of investigating the relationship or dialectic between online and offline phenomena. Our framework therefore allows us to propose an inclusive definition of online hate and a typology that is attentive to both contexts and intersectional readings of users' identities.

Defining social media hate and typology of hateful content

Emerging from our reading of the literature and our analysis of data presented in Chapters 2–5, we opt for an inclusive definition of what has been called 'online or social media hate' as

> online content which demeans, dehumanises, stereotypes, perpetuates or legitimises discrimination against, initiates or legitimises violence against individuals or groups based on protected characteristics such as: Social class, caste, race, religion, ethnicity, gender, sex, sexual orientation, disability, neurodiversity, age, language, body size and political orientation.

Based on this inclusive definition and on evidence from past and current research,[10] we propose a typology of social media hate that illustrates the complexity and diversity of the problems that need to be addressed.

This typology foregrounds three interlinked insights: First, a sense of the spectrum of hateful content linked to collective identities that circulates on social media; second, a spectrum of potential actors linked to collective identities who engage in and perpetrate online hate; and third, a spectrum of actors linked to identities who are most likely to be the targets and intended recipients of such hate.

Table 1.1 Typology of Social Media Hate, Perpetrators and Recipients

TYPES OF HATEFUL CONTENT	**Racist content** including but not confined to anti-Black, anti-Asian, anti-Indigenous, antisemitic, Islamophobic, anti-Dalit and casteist denigration, disinformation, misinformation, stereotypes, slurs (often disguised as jokes or questions), direct personalised denigration (sometimes disguised as intellectual engagement or false praise), abuse, threats and still or moving images of killings and lynchings. **Sexist and misogynist content** (often aimed at a subset of women based on an intersection of sexual, caste, racial or religious identity) including but not confined to sexist disinformation, misinformation, jokes, rape jokes, rape threats, pornography, objectification, slut-shaming, victim-blaming, personalised denigration, body shaming, indirect group denigration (sometimes disguised as mansplaining or apparent 'intellectual' challenge), patriarchal religious edicts, private images made public, morphed images and deep fakes. **Xenophobic and anti-immigrant content** including but not confined to denigrating and even genocidal comments about wars, losses in wars, the superiority of particular nations, races and ethnicities over others, slurs, jokes, morphed still and moving images containing disinformation and misinformation, incitement to violence against refugees and asylum seekers, victim-blaming and images of the dying or dead. **Homophobic, transphobic and biphobic content** including but not confined to denigrating or genocidal comments about all members of these groups, denigrating stereotypes, slurs, misinformation and disinformation, morphed images displaying what are assumed to be degrading sexual positions, body-shaming, sex-shaming, dead-naming, transmisogyny, allegations of being predators, false association with paedophilia, as well as direct and indirect threats of violence, rape and death. **Classist content** including but not confined to denigrating comments, classist labels, slurs, associations of particular religious, ethnic, racial or caste characteristics with particular class backgrounds, open or disguised snobbery and denigration of working-class tastes. **Anti-fatness and body shaming** often occurring at the intersection of another aspect of identity such as gender, race or sexual orientation **Ageism** occasionally targeted at older people/the elderly but primarily aimed at "teenagers" and "young people" including intellectual denigration, slurs, demeaning stereotypes, generalisation from incidents of public disorder, and direct, personalised abuse.

(*Continued*)

Table 1.1 (Continued)

	Ablism including derogatory slurs around particular mental health conditions, misinformation about long term illnesses, physical conditions or disabilities and learning difficulties, suspicion of claims around neurodiversity and disability, morphed images targeting particular groups (sometimes at the intersection of another aspect of identity), threats, abuse, denigration and genocidal comments.
	Anti-democratic and anti-justice content aimed widely at critics of conservative or illiberal politics and at dissidents to the state and within particular movements, including but not confined to abuse, death-threats, rape threats, slurs, morphed images, deep-fakes, sexualised and racialised threats and slurs, false accusations of corruption or nepotism, allegations of being paid supporters, visual association with vilified public figures, images with nooses and other weapons to indicate death threats and celebration of pain, incarceration or torture.
TYPES of HATE PERPE-TRATOR/ ACTOR identities	**Organised state-linked groups/actors** (paid and unpaid) These producers and spreaders of disinformation and hateful content are usually working for the government and/or ruling party where this is rightwing and/or far right in ideology, with a bouquet of socially and economically authoritarian goals. They operate both online and offline, with protection from the state. In countries with weak liberal or leftist governments, these online actors sometimes work for the main rightwing opposition party to oppose the government.
	Organised non-state groups/actors (paid and unpaid) These groups and actors are usually working on behalf of or think they are working on behalf of the government and/or ruling party or on behalf of a racial or religious supremacist ideology, with a bouquet of socially and economically authoritarian goals. They operate both online and offline, with considerable power and legitimacy. In countries with weak liberal or leftist governments, these online actors sometimes work for the main rightwing opposition party.
	Unorganised non-state actors united by prejudices or by presumed caste/religious/ethnical or racial identity, usually digitally literate individuals acting independently (occasionally left and/or liberal politically but usually with an affinity for conservative, rightwing or far right ideas and systems)
	Opportunist grifters who troll/spread misinformation to increase their fame, following and/or finances – these are usually high-profile people, or those who once held left/ liberal values and are now publicly performing their rightwing allegiance for economic or political gain

	Disruptive libertarians who regularly troll, flame or dox with a mission to disrupt, destabilise, unsettle political consensus on specific issues (such as masking, vaccines, medical advice) or mock particular individuals (who support causes that they oppose) but who want to remain under the radar
	Digital stalkers who intentionally target individual social media handles for ideological or personal reasons (desire to patronise, wish to get a reaction), or through a sense of spurned affection/loyalty which may spill over into violence
	Intermittent trolls, malicious users, inexperienced users, those piling-on through peer-pressure or out of fear of being targeted themselves
TYPES OF RECIPIENT /Context/ Identities	**Women** (of specific races, castes, classes, sexualities and religions) in the public sphere. Within this category Black women with leftist/pro-democracy views, and Muslim women are most visibly subjected to violent forms of misogyny and hate online
	Women (of specific races, castes, classes, sexualities and religions) in the quasi-private sphere
	Men (of specific races, castes, classes, sexualities and religions) in the public sphere. Within this category trans men, Black men with leftist/pro-democracy views, and Muslim, Dalit and Adivasi men are most visibly subjected to violent forms of racism, casteism and hate online
	Men (of specific races, castes, classes, sexualities and religions) in the quasi-private sphere
	Trans individuals and LGBTQIA+ groups (male, female, non-binary, agender, gender-fluid) and particularly those who are visible at the intersection of another protected characteristic, with the highest violence aimed at Black and Brown transwomen
	Women (of all races, castes, classes, sexualities and religions) in the public sphere
	Women (of all races, castes, classes, sexualities and religions) in the quasi-private sphere
	Minoritised groups by virtue of religious, ethnic, racial, sexual, class, caste, disability, body image – and in particular those groups who are publicly vocal about rights
	Refugees, asylum seekers and economic migrants
	Young people, particularly those who are vocal on justice issues
	Traveller communities, Roma
	Working class activists, **trade unionists** and economic rights groups
	Investigative journalists, scholars, rights activists and groups in the spheres of disability, human and civil rights, feminism, the environment and anti-capitalism.

(*Continued*)

Table 1.1 (Continued)

	Pro-democracy fact-checkers, lawyers and judges **Medical professionals/experts** who comment/campaign on health policy **Political dissidents,** community organisers and activists, with leftist/social justice views **Celebrities, entertainers,** actors, artists, writers and other high profile professionals who express support for decolonial causes (such as Palestine, Kashmir) **Individual men from racial or religious majority communities** who are targeted for their ally-work/ outspokenness against prejudice/critique of rightwing/ authoritarian practices and/or have been accused of public wrong-doing.

Conclusion

This chapter has traced a pathway through a series of literatures on legal and scholarly definitions of hate or hate speech, on methodological studies and typologies of hate and its recipients to lay out our theoretical framework and an original typology of social media hate. In Chapter 2 we explore the ways in which existing ecosystems of hate online have been tackled by corporate, national and international laws and policies. Asking who is culpable, we explain how and when the levels of hate reach genocidal levels as they did in the case of the Rohingya in Myanmar. The failure to act against hateful content is linked to *a refusal on the part of many in power to acknowledge and act* on the acute danger faced by target communities tied to the value (or lack thereof) that those in power accord to the lives being ruined and lost in comparison to diplomatic and profit-related goals. A nexus of soft and hard prejudice in government chambers, courts, media houses and boardrooms is shown to encourage discrimination, hate and abuse to flourish in the streets and online, and to do its violent work unimpeded. Chapters 3, 4 and 5 analyse and illuminate the links between contemporary political power, ideology, media infrastructures, social media hate, recipient identities and historical trajectories of discrimination, violence and dispossession in Brazil, India and the UK respectively, while Chapter 6 sums up our conclusions.

In this book we argue that in every country there is an urgent need to get a sense of which groups have been oppressed by which other groups through which means and serving whose interests. Alongside this it is essential to comprehend how hegemony has been constructed in discursive and material terms. This investigation requires an intersectional and historicised consideration of cultural and social practices relating to systemic and sociocultural

discrimination, hatred and violence. For instance, the caste system as a graded system of inequality, using endogamy juxtaposed on exogamy and as a basis for the division of labour and labourers (Ambedkar, 1968) necessarily intersects with gender and class relations across South Asia (Banaji, 2017). A critical anti-caste perspective, like one nested in Black liberation and critical race theory, could go a long way in explaining the power geometries of hate speech in India and casteist speech abroad. This explanation works not only in terms of identifying and countering explicitly casteist hate speech as Shanmugavelan (2021) argues, but also to explain the historical reasons why chauvinist Hindus have othered Indian Muslims and Christians in insidious ways to divert attention from caste-based violence and discrimination. Similarly, unpacking the complicated history and relationship between colonialism, slavery, race, class and skin colour, especially in regard to contemporary anti-Blackness (cf. Telles, 2006) is crucial to understand the matrix of power-relations which generates much contemporary hate speech, online harassment and violence in Brazil and the UK. Our theoretical framework emphasising the centrality of history, political economy, infrastructure, listening, intersectionality and conscientisation also leads us to attend to online hate speech as systematic, intentional and predictable – an integral facet of the contemporary *creation of difference* which strengthens authoritarianism.

Notes

1 Sagar. (2021, March 1). Delhi violence unmasked. *The Caravan*. https://caravanmagazine.in/politics/part-one-how-rss-bjp-members-invoked-hindu-identity-to-mobilise-hindutva-mobs-at-maujpur
2 Julia Carrie Wong & Hannah Ellis-Petersen. (2021, April 15). Facebook planned to remove fake accounts in India – until it realized a BJP politician was involved. *The Guardian*. www.theguardian.com/technology/2021/apr/15/facebook-india-bjp-fake-accounts
3 Arvind Chhabra. (2020, January 4). Shaheen Bagh: The women occupying Delhi street against citizenship law. *BBC*. www.bbc.co.uk/news/world-asia-india-50902909
4 David Fischer and Ajit Mohan. (2020, April 21). Facebook invests $5.7 Billion in India's Jio platforms. *Facebook*. https://about.fb.com/news/2020/04/facebook-invests-in-jio/
5 See www.ohchr.org/en/professionalinterest/pages/ccpr.aspx
6 See www.ohchr.org/en/professionalinterest/pages/crimeofgenocide.aspx
7 See www.ohchr.org/EN/ProfessionalInterest/Pages/CEDAW.aspx
8 See https://search.coe.int/cm/Pages/result_details.aspx?ObjectID=0900001680505d5b
9 See https://documents-dds-ny.un.org/doc/UNDOC/GEN/G15/000/32/PDF/G1500032.pdf?OpenElement
10 Particularly our report WhatsApp Vigilantes, and the typologies proposed by Gelber and McNamara, and Awan.

References

Abu-Sway, M. (2005). Islamophobia: Meaning, manifestations, causes. *Palestine-Israel Journal of Politics, Economics and Culture*, *12*(2), 1–15.

Aguilera-Carnerero, C., & Azeez, A. H. (2016). "Islamonausea, not Islamophobia": The many faces of cyber hate speech. *Journal of Arab & Muslim Media Research*, *9*(1), 21–40.

Alkiviadou, N. (2019). Hate speech on social media networks: Towards a regulatory framework? *Information and Communications Technology Law*, *28*(1), 19–35.

Allen, C. (2010). *Islamophobia*. Ashgate.

Aloysius, G. (1997). *Nationalism without a nation in India*. Oxford University Press.

Ambedkar, B. R. (1968). *Annihilation of caste with a reply to Mahatma Gandhi; and castes in India: Their mechanism, genesis, and development*. Bheem Patrika Publications.

Ambedkar, B. R. (1989). From millions to fractions. In *Dr. Babasaheb Ambedkar: Writings and speeches* (Vol. 5, pp. 229–246). Education Department, Government of Maharashtra.

Amelina, A., Nergiz, D., Faist, T., & Schiller, N. G. (2012). *Beyond methodological nationalism: Research methodologies for cross-border studies*. Routledge.

Appadurai, A. (2006). *Fear of small numbers: An essay on the geography of anger*. Duke University Press.

Arendt, H. (1963). *Eichmann in Jerusalem: A report on the banality of evil*. Faber and Faber.

Awan, I. (2016). Islamophobia on social media: A qualitative analysis of the Facebook's walls of hate. *International Journal of Cyber Criminology*, *10*(1), 1–21.

Banaji, S. (Ed.). (2011). *South Asian media cultures: Audiences, representations, contexts*. Anthem.

Banaji, S. and Buckingham, D. (2013). The Civic Web: Young people, the internet and civic participation. Massachusetts, MA: MIT Press.

Banaji, S. (2017). *Children and media in India: Narratives of class, agency and social change*. Routledge.

Banaji, S., & Bhat, R. (2019). *WhatsApp vigilantes: An exploration of citizen reception and circulation of WhatsApp misinformation linked to mob violence in India*. The London School of Economics and Political Science.

Benesch, S. (2013). *Dangerous speech: A proposal to prevent group violence*. United States Holocaust Memorial Museum.

Benjamin, R. (2019). *Race after technology*. Polity.

Bhat, R. (2020). *The politics of internet infrastructure: Communication policy, governmentality and subjectivation in Chhattisgarh, India* [PhD thesis, The London School of Economics and Political Science].

Burch, L. (2018). "You are a parasite on the productive classes": Online disablist hate speech in austere times. *Disability and Society*, *33*(3), 392–415.

Busolo, D. N., & Ngigi, S. (2018). Understanding hate speech in Kenya. *New Media and Mass Communication*, *70*, 43–49.

Castells, M. (1996). *The rise of the network society: The information age: Economy, society and culture*. Blackwell Publishers.

Chakravartty, P. (2004). Telecom, national development and the Indian state: A post-colonial critique. *Media, Culture & Society*, *26*(2), 227–249.

Chen, K. H. (2010). *Asia as method: Toward deimperialization*. Duke University Press.

Cherian, G. (2016). *Hate spin: The manufacture of religious offense and its threat to democracy*. MIT Press.

Costello, M., Hawdon, J., & Ratliff, T. N. (2017). Confronting online extremism: The effect of self-help, collective efficacy, and guardianship on being a target for hate speech. *Social Science Computer Review*, *35*(5), 587–605.

Delgado, R. (1993). If he hollers let him go: Regulating racist speech on campus. In M. J. Matsuda, C. Lawrence, R. Delgado, & K. Crenshaw (Eds.), *Words that wound: Critical race theory: Assaultive speech, and the first amendment* (pp. 53–88). Westview Press.

Douglas, J. (1985). *Creative interviewing*. Sage.

Dyregrov, K., Dyregrov, A., & Raundalen, M. (2000). Refugee families' experience of research participation. *Journal of Traumatic Stress*, *13*(3), 413–426.

Elareshi, M. (2019). University students' awareness of social media use and hate speech in Jordan. *International Journal of Cyber Criminology*, *13*(2), 548–563.

Erjavec, K., & Kovačič, M. P. (2012). "You don't understand, this is a new war!" Analysis of hate speech in news web sites' comments. *Mass Communication and Society*, *15*(6), 899–920.

Favaro, A., Maiorani, M., Colombo, G., & Santonostaso, P. (1999). Traumatic experiences, posttraumatic stress disorder, and dissociative symptoms in a group of refugees from former Yugoslavia. *The Journal of Nervous and Mental Disease*, *187*(5), 306–308.

Felmlee, D., Rodis, P. I., & Francisco, S. C. (2018). What a B!tch!: Cyber aggression toward women of color. In *Gender and the media: Women's places* (pp. 105–123). Walden University.

Freire, P. (2000[1980]). *Pedagogy of the oppressed*. Bloomsbury.

Gagliardone, I., Gal, D., Alves, T., & Martinez, G. (2015). *Countering online hate speech*. UNESCO.

Galloway, A. R. (2004). *Protocol: How control exists after decentralization*. MIT Press.

Gelber, K., & McNamara, L. (2016). Evidencing the harms of hate speech. *Social Identities*, *22*(3), 324–341.

Gray, K. (2020). *Intersectional tech: Black users in digital gaming*. LSU.

Gupta, A., & Ferguson, J. (1997). *Anthropological locations: Boundaries and grounds of a field science*. University of California Press.

Haraway, D. (1988). Situated knowledges: The science question in feminism and the privilege of partial perspective. *Feminist Studies*, *14*(3), 575.

Hernández, T. K. (2011). Hate speech and the language of racism in Latin America: A lens for reconsidering global hate speech restrictions and legislation

models. *University of Pennsylvania Journal of International Economic Law*, *32*(3), 805–842.
Larkin, B. (2008). *Signal and noise: Media, infrastructure, and urban culture in Nigeria*. Duke University Press.
MacKenzie, D., & Wajcman, J. (1985). *The social shaping of technology*. Open University Press.
Maitra, I., & McGowan, M. (2012). Introduction and overview. In I. Maitra & M. McGowan (Eds.), *Speech and harm: Controversies over free speech* (pp. 1–23). Oxford University Press.
Mansell, R. (2012). *Imagining the Internet: Communication, innovation and governance.* Oxford University Press.
Massey, D. (2005). *For space*. Sage.
Matsuda, M. J. (1989). Public response to racist speech: Considering the victim's story. *Michigan Law Review*, *87*(8), 2320–2381.
Merleau-Ponty, M. (2012[1945]). *The phenomenology of perception*. Routledge.
Mirchandani, M. (2018). *Digital hatred, real violence: Majoritarian radicalisation and social media in India*. ORF Occasional Paper.
Moore, K., Mason, P., & Lewis, J. (2008). *Images of Islam in the UK: The representation of British Muslims in the national print news media 2000–2008*. Cardiff University.
Morozov, E. (2013). *To save everything, click here: Technology, solutionism, and the urge to fix problems that don't exist*. Allen Lane.
Parekh, B. (2012). Is there a case for banning hate speech? In M. Herz & P. Molnar (Eds.), *The content and context of hate speech: Rethinking regulation and responses* (pp. 37–56). Cambridge University Press.
Parks, L., & Starosielski, N. (2015). *Signal traffic: Critical studies of media infrastructures*. University of Illinois Press.
Pohjonen, M., & Udupa, S. (2017). Extreme speech online: An anthropological critique of hate speech debates. *International Journal of Communication*, *11*, 1173–1191.
Said, E. (1978). *Orientalism*. Penguin.
Sethi, A. (Ed.). (2018). *American hate: Survivors speak out*. The New Press.
Shanmugavelan, M. (2021). *Caste-hate speech: Addressing hate-speech based on work and descent*. DSN.
Spivak, G. (1988). Can the subaltern speak? In C. Nelson & L. Grossberg (Eds.), *Marxism and the interpretation of culture* (pp. 271–313). Macmillan.
Starr, S. (1999). The Ethnography of Infrastructure. *American Behavioral Scientist, 43*(3), 377–391.
Telles, E. (2006). *Race in another America: The significance of skin color in Brazil*. Princeton University Press.
Thomas, M. S., Crosby, S., & Vanderhaar, J. (2019). Trauma-informed practices in schools across two decades: An interdisciplinary review of research: *Review of Research in Education*, *43*(1), 422–452.
Visveshwaran, K. (1996). *Fictions of feminist ethnography*. Oxford University Press.

Williams, P. (1987). Spirit-murdering the messenger: The discourse of fingerpointing as the law's response to racism. *University of Miami Law Review*, *42*(1), 127–157.

Williams, R. (1974). *Television: Technology and cultural form*. Routledge.

Winner, L. (1980). Do artifacts have politics? *Daedalus*, *109*(1), 121–136.

Zinn, M. B. (1979). Field research in minority communities: Ethical, methodological and political observations by an insider. *Social Problems*, *27*(2), 209–219.

2 When hate-speech policies and procedures fail

The case of the Rohingya in Myanmar

Introduction

Since 2005, social media companies have grown rapidly in terms of numbers of users and in terms of economic value. Pursuing markets, corporate owners have paid insufficient attention to preventing discrimination and violence targeted at vulnerable groups in society. In the introductory chapter we argued that the debate around hate speech online and incitement to violence has been framed around the preservation of an abstract principle of free speech and a failure to acknowledge the different forms of dehumanisation, harassment, discrimination and violence that groups and individuals experience daily. In this chapter, we critically evaluate some of the safeguarding policies agreed by social media companies – either as guidelines, approaches or terms of service for users. Most of these documents list a range of problematic content. Our analysis is restricted to guidelines available in the public domain. In addition, social media companies have internal documents for content moderators, which are not accessible to the public.[1]

Overall, social media companies have adopted approaches to discriminatory and toxic speech which combine big data, artificial intelligence and algorithms. Most social media companies operate across a fine balance between contradictory objectives: On the one hand, restricting the ease with which information can be circulated (especially bulk circulation) amongst users and, on the other hand, promoting both the number of overall users (often linked to valuation of the company) and the ease of information circulation (often linked to greater Average Revenue Per User or ARPU). This dynamic is a result of the political economy of social media wherein surplus value is generated through complex mechanisms such as commodification of audiences, monopolisation and concentration, advertising and so on (c.f. Arvidsson & Colleoni, 2012; Dahlberg, 2015; Fuchs, 2010; Rigi & Prey, 2015). The primary mode of achieving this balance is through innovation on affordances (Gibson, 1982, p. 403). This approach is based on

DOI: 10.4324/9781003083078-2

'investigating the empirical question of embodied human practices in real time situated interaction involving technologies' (Faraj & Azad, 2012; Hutchby, 2003, p. 582).

An advantage of affordance approaches is the attention they pay to the ways in which the materiality of communication technologies and infrastructures intersects discrimination and violence. Affordance-based approaches also maintain a relational perspective on the use of technologies, thereby bringing attention to *users* and *usage*, rather than technologies in a vacuum. For example, after significant evidence of WhatsApp misuse for bulk circulation of disinformation in India, WhatsApp introduced features to curb bulk distribution. These restrictions included limiting the number of messages that can be forwarded at a single time, providing identifying information on forwards (including messages forwarded many times) and so on. While WhatsApp has not released country-specific data on the impact of such affordance-based interventions, large-scale disinformation, abuse, discrimination and other forms of physical and symbolic violence continue in India and across the globe.

The affordance-based approach has shortcomings too. First, it fails to consider the complex historical contexts within which heterogenous users have unequal access to power. Some users have power *to* and power *over* others while others experience life as individuals or groups who may resist but are frequently *othered* and *acted upon*. The lack of a political and historical perspective obscures the complex ways in which communication technologies and infrastructures are themselves caught up in contests over power and dialectically related to intersecting and changing identities. Situating the emergence of social media use within a neoliberal economy and culture, against the backdrop of race, caste, religious or gender conflict, adds precision to sociotechnical decision-making by pointing more specifically to factors which enhance or inhibit equity and justice online.

While these approaches are most common for dealing with discrimination and violence on a global scale, even these approaches are internally inconsistent. Market pressures and geopolitical allegiances lead to uneven ways in which violence has been handled in different parts of the world. In order to illustrate this, we take the case of discriminatory, violent and hateful content in Myanmar. Apart from sporadic attention in 2018 to pogroms targeted at Rohingya Muslims and more recently in 2021 with the military coup, Myanmar has largely remained invisible to social media companies when it comes to designing policies about online violence and hate. The murder of and displacement of Rohingyas into neighbouring countries has further widened the spread of disinformation, discrimination and hate against the Rohingya in mainstream and social media. Social media companies have largely failed to stem the hate in which they remain complicit.

Social Media corporations' responses to hateful content: Policies, human moderation and machine learning

Facebook's Community Standards[2] (also applicable to Instagram) define hate speech as a 'direct attack against people based on what we call protected characteristics: race, ethnicity, national origin, disability, religious affiliation, caste, sexual orientation, sex, gender identity and serious disease'. Attacks are defined as 'violent or dehumanising speech, harmful stereotypes, statements of inferiority, expressions of contempt, disgust or dismissal, cursing and calls for exclusion or segregation'. Twitter's policy,[3] meanwhile, claims that the company will not allow users to 'promote violence against or directly attack or threaten other people based on race, ethnicity, national origin, caste, sexual orientation, gender, gender identity, religious affiliation, age, disability or serious disease'. YouTube's policy[4] on hate speech too claims that it will remove content promoting violence or hatred against individuals or groups based on a set of attributes similar to the ones mentioned above.

The constellation of social media also includes cross-platform apps, peer-to-peer and peer-to-group messaging services, such as Facebook-owned WhatsApp and Facebook Messenger, Viber, Telegram and Signal that have gained in popularity since 2017. These messaging services and apps enable private groups on a fully encrypted architecture that makes it difficult to detect hate speech via content moderation unless encryption is compromised, or until screenshots are made public on other platforms. This dilemma over de-encryption (and the compromising of individual privacy) epitomises how violence, hate and inciting communication are addressed in jurisprudence in the form of a tension between preserving commitments to freedom of speech and expression while preventing hate speech and further hate crimes.

Telegram was founded by Pavel Durov, who also founded a Facebook-like social media company called Vkontakte in Russia in 2006. When Durov refused to divulge sensitive data to the Russian government in 2012, the government acquired 48% of Vkontakte. Under pressure to preserve privacy, Durov created Telegram as a more secure service. Durov recounts how after the Snowden revelations, especially around the complicity of technology and social media companies in cooperating with the US government on surveillance of citizens, he was inspired to take Telegram public.[5] Telegram's Terms of Service[6] document has nothing to say specifically about incitement or hate speech, except in the prescription that users agree not to 'promote violence on publicly viewable Telegram channels, bots' and so on.

Two convergent events over the course of late 2020 and early 2021 have drawn attention to links between policies on hateful content and corporate

actions: The Capitol riots in Washington D.C. followed by the suspension of accounts on Facebook and Twitter, including the accounts of then American president Donald Trump; and changes to WhatsApp's privacy policy. These led to a surge in users for Telegram and Signal which were seen as more secure.[7] WhatsApp's changes were ostensibly to share data from WhatsApp business accounts, but their failure to reassure users led to a mass switching of accounts to Signal and Telegram, including by Hindu supremacists, white supremacists and other far right users.

WhatsApp's Terms of Service[8] claim to prevent users from using WhatsApp in ways that are 'illegal, obscene, defamatory, threatening, intimidating, harassing, hateful, racially or ethnically offensive, or instigate or encourage conduct that would be illegal or otherwise inappropriate, including promoting violent crimes'. Working with data that is unencrypted and available to companies – such as phone number, Display Picture (DP), status updates, geolocation and correlation of this data to those users' content on Facebook and/or Instagram that is unencrypted and easily accessible – policy and technical teams at WhatsApp are exploring options for identifying users propagating hate speech. Another avenue to curb the overflow of antisemitic, Islamophobic, anti-Black, homophobic, misogynist and otherwise racist expression has been to attempt to reduce the usage of unauthorised 'clone' apps that circumvent preventative measures – such as limits on forwarding – taken by the owners of these applications.

Given the scale of their operations, social media companies deploy Artificial Intelligence (AI) not only to respond to reports of hate speech by users but also to identify and act on communication potentially designated hateful before users report it. On a webpage titled 'Safety in India', WhatsApp, for instance, claims[9] to have developed 'advanced learning technology to identify and ban accounts engaging in bulk or automated messaging and bans more than two million accounts from WhatsApp per month, 75% of them without a recent user report'. Note, however, that WhatsApp under the aegis of parent company Facebook has restricted itself to removing accounts associated with a specific type of coordinated activity rather than committing to tackling casteist, Islamophobic, racist, misogynist, and/or homophobic content *per se*. This deliberate choice is even more troubling if it is considered that it is perfectly possible for bulk and/or automated messaging to be used for non-commercial, pro-social and non-threatening purposes such as to combat common misinformation and to send out vaccination reminders.

The use of algorithms to organise and prioritise content on various social media platforms and apps has prompted significant research. Much of this work acknowledges the increasing prominence of algorithms in all kinds of social relations with real-world consequences including large-scale violence and discrimination (Beer, 2009; Edelman & Luca, 2014; Lewis,

2018). Several scholars in the fields of Science and Technology Studies and social media studies, as well as ethical AI teams at Google and elsewhere, have explored ways in which algorithmic operations (and the biases coded into them) increase discrimination and can be held more accountable (Eubanks, 2018; Gebru et al., 2017; Noble, 2018). Options for ameliorating the situation include small-scale observation, reverse engineering, 'scraping audits' and so on (Bucher, 2012; Rieder et al., 2018; Sandvig et al., 2014). The reasons and complications that reduce the possibility of using technical means to limit hate speech and a deeper exploration of the ideologies and predilections which shape the genesis and role of algorithms are important but fall outside the scope of this book. It should be noted, however, that the technicalities of algorithms notwithstanding, there have been some other efforts, mainly in the USA, to put commercial and moral pressure on social media companies to take down hate speech more widely and effectively.

While there is little doubt that social media companies will continue altering algorithms in order to increase their efficiency in detecting and removing anything that could legally be proven to be hate speech on their platforms or which is called hate speech by powerful entities such as governments, it is worth noting that in the end these companies are heavily constrained from acting on hate speech since these actions will increase 'friction' of usage. Examples include Instagram enabling users to filter comments and Twitter enabling users to hide replies. This friction – which decreases the speed and ease of usability – has consequences for the growth of these companies. Friction reduces unique subscribers and potentially loses users to other platforms and apps. Their valuation is tied to the number of users and, concomitantly, to a monetisation of user data.

Bearing in mind the contradictions between the ethical impulses of the employees and the commercial interests of the shareholders and owners, it is unlikely that owners will undertake any radical action to curb the variety of dehumanising and discriminatory communications on their apps and platforms without market pressure. Indeed, researchers working at Google's AI ethics unit, Timnit Gebru and Margaret Mitchell, were removed by Google in late 2020 and early 2021 for planning to publish research highlighting flaws in Google's AI technology. Stalled, Gebru refused to retract, and sent internal emails to colleagues maintaining that Google was silencing marginalised voices.[10] Google's actions are indicative of less than stellar employment practices, and of a decision to downplay and drag their feet on evidence of digitally enhanced discrimination.

Another case, that of Facebook's whistle-blower Sophie Zhang, who wrote a long memo after she quit Facebook in September 2020, is also illustrative of the corporate tech world's response to online hate. Speaking to Buzzfeed soon after her resignation[11]in 2020, Zhang revealed that

she had been a data scientist for Facebook's Site Integrity fake engagement team, looking into what Facebook calls 'Coordinated Inauthentic Behavior' (CIB), in multiple countries including Brazil, India, Myanmar and the UK. CIB encompasses fake pages, bots, troll farms and so on. In many of these countries, much manipulation was traceable to users affiliated to the ruling political parties during elections. In some cases, Zhang dealt with pages that clearly propagated disinformation against specific groups based on their ethnicity, religious identity, sexual orientation, political affiliation, including disinformation related to Covid-19. While it is not easy to draw a causal link between the content (including hate speech) published on these pages and the physical consequences that followed for vulnerable groups in these countries, Zhang said that she felt that she had 'blood on her hands'.

It is worth noting here that the problem was *not* Facebook's ability to identify and flag CIB and associated hate speech, but rather two other issues arising from large-scale CIB. First, for Facebook, CIB although a universal problem, was to be dealt with through an internal evaluation process that prioritised addressing incidents of hate speech in Europe, North America and Australia-New Zealand rather than in 'far away' countries such as Azerbaijan or Honduras. While it is possible to posit profit-based cynicism as an explanation for the lack of action taken over countries such as Azerbaijan or Honduras due to their relatively small user base or their marginality in the mediated presentation of the world, it is another matter for Facebook to ignore CIB and hate speech in vast countries such as India and Brazil. The *Wall Street Journal* broke stories[12] in August 2020 with evidence that senior officials at Facebook India, such as the Public Policy director Ankhi Das, deliberately ignored warnings from Facebook employees about Islamophobic hate speech by politicians of the far-right ruling BJP. By October 2020 Ankhi Das had left Facebook to save face for the company, but the intentions of Facebook were clear when she was replaced by another Hindutva supporter, Shivnath Thukral, to lead its public policy unit. Thukral worked for the BJP's election campaign in 2014 when Narendra Modi became Prime Minister. TIME magazine has since reported[13] that Thukral too failed to remove hate speech against Bangladeshi Muslims posted by Assamese BJP leader Shiladitya Dev.

Facebook's audience size puts India at the top with 320 million users in 2021.[14] By comparison, the United States is second with only 190 million users. Brazil too has 190 million users, and when Zhang alerted her colleagues to CIB in India and Brazil these reports were ignored, not because these countries were deemed fiscally unimportant – rather, users who were connected to the identified CIB were closely affiliated to the parties in power in these countries. Brazil has over 120 million monthly active WhatsApp users with 98% of smartphone users surveyed reported to be users.

India has approximately 400 million users making it by far the largest market for WhatsApp.[15] The parties in power in Brazil and India have, through their leaders' acts and speeches, as well as various politicians affiliated to these parties, repeatedly and brazenly flouted social media policies on hate speech and incitement to violence.[16] Indeed, Facebook and others have not just 'looked the other way' in order to continue benefiting from their large user base in these countries, but have actively recruited high ranking employees with sympathies for the far right regimes and their ideologies. It is clear then, that when policies fail, they do not fail simply because the policies are flawed or moderation teams poorly trained. Sometimes it seems clear that the highest echelons in companies never intended those policies to jeopardise their political and economic relationships with authoritarian governments.

Social media companies are capable of ideological course correction when it comes to hate speech when subjected to pressure. For example, after Minneapolis police officer Derek Chauvin murdered George Floyd on the 25 May 2020, groups advocating on behalf of Black communities put pressure on social media companies to improve their policies. In June 2020, a campaign called Stop Hate for Profit lobbied large corporations to stop advertising on Facebook. Unilever's announcement that it would stop advertising on Facebook within the United States prompted Mark Zuckerberg to declare that Facebook would overhaul its algorithms detecting hate speech.[17] The swiftness of this *volte face* following years of equivocation indicates that the politics and ideological commitments of corporate directors play a major role in sustaining or ameliorating the stereotyping, dehumanisation, degradation, misrepresentation and violent or threatening abuse that characterise the online experiences of individuals from ethnic, racial, religious and sexual minorities, and disabled communities, worldwide.

The campaign to reduce advertising expenditure – which amounts to a call for political boycott – is important. Advertising is the main source (97%) of revenue for Facebook, increasing by 21% from $69.7 billion in 2019 to $84.2 billion in 2020.[18] In July 2020, Nick Clegg (Vice President of Global Affairs and Communications) asserted disingenuously that Facebook does not profit from hate. Faced with the threat of a boycott by advertisers, Facebook tripled to 35,000 the employees working on safety and security, and reported that their algorithm assessed 95% of hate speech reports in less than 24 hours, faster than YouTube and Twitter. This suggests that when Facebook's profits are threatened, they have the capacity[19] and the will to improve the identification and take down of hateful communications. How these takedowns are then circumvented by the far right and by users with racist and/or other prejudiced opinions, and how the policies are used to

feed into rightwing narratives of victimhood, are beyond the scope of this book, but provide fascinating material for further analysis.[20]

In July 2020 the National Association for the Advancement of Colored People (NAACP) as part of a civil rights petition[21] to Facebook demanded a list of ten changes to bring about accountability (including civil rights infrastructures, independent audits and refunds to advertisers); decency (including take downs of hateful content and automatic internal flagging of hateful content in private groups) and support (increasing personnel and expert teams at Facebook). At a pace much slower than that seen after Unilever's pressure, Facebook announced in December 2020 that they would embark on an overhaul of their algorithms to detect hate speech. Titled 'Worst of Worst' (WOW), *The Washington Post* accessed documents that showed that the project[22] would in effect prioritise hate speech against groups that have historically faced discrimination, especially by virtue of race. There is little doubt that the crucial role of social media in mobilising Black communities after the murder of George Floyd has resulted in the same groups now pressing those very social media companies to undertake systemic and long-lasting changes regarding racialised hate speech. Whether this purposeful pursuit of a less violent and damaging online discourse will have beneficial long-term effects or be by-passed by ever more tech-savvy supremacist and conservative systemic lobbying and infiltration remains to be seen. However, tech corporations continue to be wary of the ever-stronger rightwing backlash.

There is much evidence that hate speech laws and other policies around online offense tend to be used by those with power against less powerful groups and lose their efficacy when brought into play by the communities who fought to implement them. As we will discuss later in this book, this twisting of corporate policies against the most vulnerable is clear in India where Islamophobic and anti-Christian content and disinformation circulate unimpeded while critiques of Hindutva fascism are taken down and their users sanctioned, in Sri Lanka where Islamophobic and anti-Tamil sentiments have thrived online and in Myanmar where anti-Rohingya postings stayed up long after they were reported but anti-government political actors are censored and imprisoned. This is also clear in reports that aim to document state crimes by Israeli authorities and settlers against Palestinians and of posts that thoughtfully aim to discuss differences between anti-Zionism and antisemitism. We now discuss these issues in relation to the case of Myanmar.

Myanmar: A predictable genocide

With at least eight distinct major ethnic groups, Myanmar in South-East Asia has a population of over 50 million, and shares land borders with

China, India, Bangladesh, Thailand and Laos. Although the constitution mentions freedom of religion, in practice, the majority religion is Buddhism; with Christianity, Islam and Hinduism along with Indigenous faiths practiced to a lesser extent. The region has a more or less continuous history of struggle between 'centre' (Bamar) and periphery (Shan, Mon and Rakhine groups) that was suppressed with British colonial rule in 1885 (Aung-Thwin, 1985; Taylor, 1987; Than, 2005). The colonial separation of 'Burma proper' from the frontier areas with non-Bamar ethnic groups ('Scheduled Areas') enabled Bamar nationalists to construct an 'imagined' Myanmar unified by a melding of disparate ethnic-based sovereign entities (Cady, 1965; Selth, 1986). The rise of Major-General Aung San as a leader of the Anti-Fascist People's Freedom League (AFPFL) for national independence and his subsequent assassination in 1947 saw the creation of a mythic figure that still plays a role in contemporary politics (Guyot, 1989; Taylor, 1987). With Aung-San's assassination, the military regime produced a routinised form of violent insurgency and counterinsurgency, 'a knitting together of networks of violence that constituted a tenuous but productive form of state-building' (Callahan, 2005, p. 115). Revolts from the other ethnic groups such as the Mon, Pao and Kachin prompted a military coup in 1962 (Smith, 1991). Border skirmishes and negotiations doubly benefitted the military leaders – the Tatmadaw – by helping to justify coercion in the face of armed insurgencies, and by allowing the junta to profit from both licit and illicit trades in profitable jade, timber and opium, while retaining control of sectors from banking and transport to telecommunications (Meehan, 2011; Selth, 1986).

Between 1988 and 2010, Myanmar's GDP grew from $12.6 bn USD to $45.4 bn, imports rose from $246 mn to $4.8 bn, exports rose from $167 mn to $8.7 bn and foreign investment rose from $4 mn to $8.3 bn, a sizeable growth primarily backed by China's involvement in Myanmar (Jones, 2014). A general election was held in 2010 and the military-backed Union Solidarity and Development Party (USDP) swept the board while the National League for Democracy was declared illegal. However, soon after the elections, politician Aung San Suu Kyi, daughter of assassinated Major-General Aung San, was released and various reforms began to be introduced such as a reform in labour laws allowing for unions and strikes, a relaxation of press censorship and amnesties on hundreds of political prisoners.

Even while the state continued to suppress minority dissent brutally, the military regime's dominance seemed to end with the formation of the Socialist Republic of the Union of Burma in 1974 under a new constitution. The democratic movement leading up to the coup of 1988 was largely comprised of Bamar intellectuals along with a coalition of students, ex-military dissidents and older political leaders (Guyot, 1989; Smith, 1991). In 1990,

the military junta – known as the State Law and Order Restoration Council (SLORC) – held an election in which Aung San's daughter Aung San Suu Kyi leading the National League for Democracy (NLD) won 392 seats and, along with the Shan NLD, cornered 90% of the seats. However, in 1990, the military ensured that the victors would only be responsible for a new constitution. Upon Suu Kyi's release in mid-1995, the two parties remained in deadlock, despite continued dialogue, with coercion and violence continuing (Smith, 1991). In the 2015 and 2020 elections, the NLD won with NLD's Htin Kyaw elected as the first non-military president since the military coup of 1962 with Suu Kyi as state counsellor. However, in February 2021 the *Tatmadaw* declared a state of emergency, jailed senior NLD leaders including the president and state counsellor, and imposed martial law in many areas. As we write, mass protests have broken out, and protestors have been brutally repressed.

'Othering' as a tool of authoritarian governance: 'We have been trained since we were children to be racists'

The process of 'othering' (Hall, 1996, pp. 4–5) links discursive practices of identity construction to stereotyping, exclusion and violation. Some minority ethnic groups were incorporated into the *Tatmadaw's* nation-building project. Others have been excluded and systematically targeted. Although several ethnic groups have struggled for decades against the *Tatmadaw*, the Rohingyas have suffered the most (Clarke et al., 2019) by being subject to a diverse set of strategies such as 'lawfare' and 'spacio-cide' (MacLean, 2019). These legal and governmental strategies legitimise the exclusion of Rohingyas from territories where they have lived for generations. Such strategies have most commonly been used, as Appadurai argues,[23] by 'paranoid' sovereignties and/or 'predatory' states against 'biominorities', in other words, populations whose difference (based on ethnicity, religion, race and so on) from the national majorities is perceived as a bodily threat to the national ethos (c.f. Hanafi, 2009; Yiftachel, 2005).

One of the first major acts of 'lawfare' against the Rohingya was 'Operation Dragon King' in 1978, which saw foreigners screened and illegal immigrants expelled. Across successive ethnic cleansing operations over the next two decades, roughly 550,000 Rohingyas fled to neighbouring countries, especially Bangladesh, though some 250,000 gradually returned. This exodus was a response to the Burmese Junta's abuse of international principles pertaining to citizenship. For example, the 1982 citizenship law, updating the 1947 constitution, explicitly tied ancestry to territory to come up with three categories of citizenship: Full citizens (descendants of residents who lived in Myanmar before 1823); associate citizens (people who acquired

citizenship under the 1948 union citizenship law); and naturalised citizens (people who have been residing in Myanmar before 1948 but had failed to apply for citizenship under the 1948 union citizenship law). All three categories of citizenship only applied to Myanmar's 'national ethnic races' first published in 1960 naming the major groups of Bamar, Kayin, Shan, Kachin, Rakhine and so on, but not to Rohingyas. A new census was conducted in 1983 and results of 135 ethnic groups revealed only in 1990, and again Rohingyas did not figure since Bamars constructed Rohingyas as Bengalis who migrated to Myanmar after 1823 and not an Indigenous population[24] (Cheesman, 2017; Cheng Guan, 2007; Ferguson, 2015; Kipgen, 2019; Kyaw, 2015).

During the early 1990s, Rohingyas were forcibly evicted so that 'model homes' could be built on their lands, with the labour of displaced Rohingyas. On 16 October 1992, the then Special Rapporteur on Freedom of Religion or Belief informed the Government that he had received information that[25]:

> since late 1989, the Rohingya citizens of Myanmar . . . have been subjected to persecution based on their religious beliefs involving extrajudicial executions, torture, arbitrary detention, forced disappearances, intimidation, gang-rape, forced labour, robbery, setting of fire to homes, eviction, land confiscation and population resettlement as well as the systematic destruction of towns and mosques.

In 2012, three Rohingya men received death sentences for the gang-rape of a Rakhine woman. A week later, ten non-Rohingya Muslims were lynched, and after a stridently effective hate speech campaign across mainstream and alternative media, mass violence followed, targeting Muslims in Rakhine, with full complicity from the state. More than 10,000 homes were destroyed and 140,000 Rohingyas were displaced, most of whom live in temporary camps set up in Rakhine. Similar to Israeli practices of settler colonialism, the *Tatmadaw* battalions, the Buddhist Rakhine communities and 'entrepreneurs' seeking fresh markets combined to drive Rohingyas from their lands and homes in Rakhine. How Facebook came to play a role in furthering the ethnic cleansing of the Rohingya and strengthening the *Tatmadaw* in Myanmar goes to the heart of our concerns in this book.

Before 2021, incidents ranging from communal clashes based on rumours, incendiary political speeches and announcements, misuse of old laws and the introduction of new laws targeting Rohingyas as well as other political dissenters were increasing. Superficial reforms and the more active role played by politically-savvy Suu Kyi had led international organisations and the UN to ignore her Bamar chauvinism (Davis, 2021; Lee, 2014). However, in 2015 Myanmar's parliament approved a set of discriminatory

laws collectively dubbed the Race and Religion Discrimination bills. Submitted to Parliament in 2013 by Ashin Wirathu, then leader of the far-right Buddhist and nationalist organisation Association for the Protection of Race and Religion (also known as *Ma Ba Tha* within Myanmar), these bills, now laws,[26] allow regional officials to establish 36-month birth spacing for specific communities (Rohingya), compel Buddhists and other groups to obtain official approval to marry partners from another faith, prohibit Muslim couples from having more than two children and impose monogamy to target Muslims who are often framed as sexual deviants. Before his account was blocked in 2018, Wirathu used Facebook unimpeded to spread disinformation and advocate violence against Rohingyas (Fink, 2018; Whitten-Woodring et al., 2020). More than two thirds of Myanmar's Rohingya population have fled.

The Rohingyas' suffering has commonly been framed as a 'humanitarian crisis' (Ullah, 2011). Since the mid-1990s, Rohingya groups failed to mobilise collectively, failed to revoke the military ban on Muslim organisations such as the Rohingya Students Union or the Rohingya Youth League and/or split into factions such as the Rohingya Solidarity Organisation and the Arakan Rohingya Islamic Front. Without protection, the 2016–17 ethnic cleansing led more than 600,000 Rohingyas to flee to Bangladesh. Rohingya men are routinely demonised and killed, while women and children are disproportionately targeted with sexual violence and torture. Subsequently, denial of healthcare and other basic infrastructural facilities in camps fuelled disease and death. Much of this has been common knowledge for years. Severe travel restrictions, systematic denial of land use for agriculture, extortion and bribery have contributed to a 'slow genocide' that seeks to erase Rohingya identity and culture from Myanmar (Amnesty International, 2017; Anwary, 2021; Houtman, 1999; MacLean, 2019; Mahmood et al., 2017; McCarthy & Menager, 2017; Ware & Laoutides, 2019; Zarni & Cowley, 2014).

Communications in Myanmar are tightly controlled by the *Tatmadaw* through prohibitive pricing and centralised control over communicative infrastructure. SIM cards and mobile phones were introduced by the Myanmar Post and Telephone department in 2000 at a cost of $5000. The Myanmar Computer Science Development Law criminalised unregistered access to the internet with a maximum 15-year prison sentence. Pre-2014, draconian legislations such as the Burma Wireless Telegram Act, Printers and Publishers Registration Act and the Myanmar Computer Science Development Law were in place (Sablosky, 2021). Post-2014 saw a rapid proliferation of legacy broadcast and print media, with increased internet penetration through foreign investment, a proliferation of cheaply available SIM cards and mobile smartphones and drastic increase in social media use (Davis,

2021; Farrelly & Win, 2016; Renshaw, 2013). In 2016–17, Facebook tried to penetrate emerging markets like India and Myanmar through Free Basics, wherein users would get Facebook pre-installed and not pay for data while using Facebook. The program failed to take off in India, but succeeded in Myanmar. Thenceforth, Myanmar relied on Facebook as the primary news interface (Whitten-Woodring et al., 2020, pp. 414–415). State media, meanwhile, was busy spreading hatred towards Rohingyas via misrepresentations as 'Bengali', 'foreigners', 'terrorists' and so on (Lee, 2019).

While Facebook provided some independent information sources for minority groups, Facebook did nothing to take down communications inciting violence even by its own definitions and standards. In 2012, President Thein Sein's spokesperson Zaw Htay posted the following message on his Facebook page:

> Rohingya terrorists as members of the RSO are crossing the border into Myanmar with weapons . . . Our troops have received the news in advance so they will completely destroy [the Rohingya]. It can be assumed that the troops are already destroying [the Rohingya]. We don't want to hear any humanitarian or human rights excuses. We don't want to hear your moral superiority, or so-called peace and loving kindness. Go and look at Buthidaung, Maungdaw areas in Rakhine State. Our ethnic people are in constant fear in their own land. I feel very bitter about this. This is our country. This is our land. I'm talking to you, national parties, MPs, civil societies, who are always opposing the President and the Government.
>
> [Detailed findings of the independent fact-finding mission in Myanmar, 2019, p. 167]

The Rohingya community is also well aware of the popular discourse on Facebook (Whitten-Woodring et al., 2020, p. 418):

> . . . the coverage of local media is not fair. What local media portrayed was that the Rohingya set their own homes on fire, like the same thing you would see on Facebook. Most people shared such news on Facebook and the whole country believed that the Rohingya just set their homes on fire.

Even as Suu Kyi herself tried to appease Buddhist nationalists in 2016 by asking the UN not to use the word Rohingya (presumably preferring the disinformation term 'Bengali'), she and her colleagues faced abuse on social media, were called 'Muslim lovers' by the *Ma Ba Tha* and their supporters (Davis, 2021, p. 113). Despite this kind of inflammatory phrasing and its

material consequences – vicious physical atrocities, rape, arson, homelessness, land-grabbing and other anti-Rohingya violence – Facebook chose not to act. Our analysis suggests that this decision may have been based on its estimation that Myanmar was too insignificant in terms of international geopolitics, and too profitable as a home for Facebook's Free Basics programme and other potential investments.

Between 2017 and 2018 the volume and content of anti-Rohingya, anti-Muslim, disinformation on Facebook in Myanmar rose. Repeated hostile misrepresentations served to legitimise extreme forms of violence and atrocity by the armed forces and Buddhist-controlled paramilitaries, which were, in turn, represented online as efforts to 'safe-guard' the nation. Military personnel stationed in mixed Buddhist and Muslim villages would segregate Buddhist villagers from their Muslim neighbours, saying they had come to 'protect them'. Western governments remained determinedly silent until the grim reports of murder and atrocity began to emerge publicly through investigative journalists in Rakhine:[27] Ten men shot and buried in one village; boys and men burnt to death in torched houses in another; women and girls gang-raped and denied medical treatment. Estimates from the UN and Human Rights Watch in 2018–19 suggest that over 7000 were murdered and more than 10,000 raped.

One of our key informants, HS, (anonymised for his safety) discussed the context of the racist socialisation that intersects with hateful content online:

> We have been trained since we were children to be racists. I mean, I attended High School in Yangon. Although our teachers and headmasters already understood the [racist] context, there were some forms of discrimination for example . . . our school [played and still plays] a patriotic song every day at the end of the day: 'The Pride of Birth', in the lyrics there are some words like 'The country for martyrish Buddhists . . .' We had been indoctrinated by such words since we were young. . . . Once, I did not take notes on history in a writing book, I just wrote a question and answer on the textbook. *So some teachers told my family that I had a Muslim friend and got spoiled*. In different ways, such as singing, routine activities and informing parents by teachers, we were trained to hate others [especially Muslims] . . . If a Buddhist student is poor in study, he or she is not blamed, but if students from other religions are poor in education, it is because of their religion.

The Myanmari diaspora are also affected by the abject failure of Facebook to act upon years of online dehumanisation and violent incitement. Rohingya refugees in India face persistent dehumanisation and threats online,

with disinformation about them tied to anti-Muslim pogroms and changes in laws. Camps have been burnt down in India. Even individuals living in the West who have spoken out against the February 2021 military coup fear for their relatives' lives. Hate speech in Myanmar is thus contemporary as well as historical and remains consonant with discrimination and racism in changing circumstances.

Rohingya under constant 'threat of genocide': Social media companies flailing

In the *Tatmadaw's* genocidal project against the Rohingyas which has unfolded over decades, social media has added fuel to inflammable rhetoric and policies. In response to disquiet amongst platform owners about the 2021 coup, the *Tatmadaw* opted for cruder forms of infrastructural control such as Internet shutdowns, intranets, surveillance through biometric technology, verification cards, facial recognition technologies and whitelisting applications that delete the vast majority of online content except on authorised applications.[28] Meanwhile, the response of social media companies continues to be disjointed. Even as Facebook had close to seven million users in Myanmar, a Reuters report[29] revealed that Facebook in 2014 had only one employee who spoke Burmese (based in Dublin) and a year later, only four employees (based in Manila and Dublin). Currently, Facebook has outsourced its content moderation in Myanmar to the global business process outsourcing firm Accenture, based in Kuala Lumpur, through a project called Honey Badger. Approximately 60 employees who speak Burmese are moderating content posted on Facebook related to Myanmar. Since connectivity came to Myanmar relatively late, it affected the development of a Burmese Unicode script and online users in Myanmar had developed their own script called Zawgyi. Oblivious to these technical issues, Facebook created a Burmese to English interface unable to decipher Zawgyi. As a result, their content moderators cannot even read and understand content accurately, let alone moderate it.

For example, an inciting phrase posted on Facebook '*Kill all the K***rs that you see in Myanmar; none of them should be left alive*' was translated by Facebook interface as '*I shouldn't have a rainbow in Myanmar*'. Kalar is a derogatory term that is commonly used to refer to Rohingya and other minoritised groups as foreigners and outsiders. In an attempt to show the world that it is taking firm action, Facebook banned the word, filtering out its Burmese characters, which they later discovered literally translated as 'from the west'. However, due to its reliance on automatic filters rather than contextual moderation, Facebook has unwittingly been removing any word that has kalar in it such as *kalar pae* (chickpeas) or *kalarkaar* (curtains).

More recently, in 2019 and 2020, Facebook removed accounts linked to the Tatmadaw, explicitly admitting that Facebook has played a role in the genocide against Rohingyas. However, in an attempt at supposed fair-play and balance, Facebook also deleted[30] several Facebook pages of what are commonly known as Ethnic Armed Organisations (EAOs). While some EAOs do discuss armed resistance and recruit on social media, many more use social media to highlight human rights abuses by the *Tatmadaw*, which otherwise have no chance of publication on mainstream media. Not all EAOs have been affected but only those which are anti-regime, thus positioning Facebook – possibly inadvertently but more likely by choice – in favour of the *Tatmadaw*, and entrenching Facebook in the entangled politics of ethnic supremacism in Myanmar (Sablosky, 2021).

Since the coup in February 2021, minority groups including political organisations mobilising in favour of minority rights, activists, academics, journalists and political dissenters have to change strategy. A young activist from Myanmar who has been speaking out against the regime since 2017 and faces death and rape threats, shares her experience[31]:

> People can't get real information . . . They restored the Internet but not the television. On Facebook there is no legitimate news verification, there's no accountability, so now misinformation is growing dangerously. There is a lot of hate speech, a lot of fake accounts. I relocated, but I don't feel safe at all, anything can happen to us. I feel a mounting threat. . .

Media literacy and sensitisation programmes for young social media users are being offered by organisations such as Myanmar Institute of Theology and Myanmar ICT for Development Organisation[32] while UNICEF runs youth media literacy and voice campaigns in Cox's Bazaar amongst the Rohingya refugees. However, since state and non-state actors, militias and Buddhist vigilante mobs are bent upon eliminating the Rohingya population, such efforts while worthy and sensible, are drops in an ocean of hate.

Given the discussion of multiple policies in the first half of this chapter and the limited but notable achievements that international pressure and boycott threats had in leading to acknowledgment of anti-Black racism on Facebook, it is important to note that many platforms continue to be used by the military and by Buddhist supremacist groups in Myanmar and therefore need to be scrutinised and held to account by international pressure. The need for such pressure will become all the more evident in the case of Brazil, which we delve into in the next chapter, and where social media disinformation and hate is implicated in far right vigilantism and authoritarian-populist regime change.

Notes

1 See www.apc.org/en/pubs/statement-facebooks-internal-guidelines-content-moderation
2 See www.facebook.com/communitystandards/hate_speech
3 See https://help.twitter.com/en/rules-and-policies/hateful-conduct-policy
4 See https://support.google.com/youtube/answer/2801939?hl=en-GB
5 See www.dazeddigital.com/artsandculture/article/24279/1/pavel-durov
6 See https://telegram.org/tos
7 See www.nytimes.com/2021/01/13/technology/telegram-signal-apps-big-tech.html
8 See www.whatsapp.com/legal/terms-of-service-eea
9 See https://faq.whatsapp.com/general/safety-in-india/?lang=en
10 See www.bbc.co.uk/news/technology-56135817
11 See www.buzzfeednews.com/article/craigsilverman/facebook-ignore-political-manipulation-whistleblower-memo
12 See www.wsj.com/articles/facebook-hate-speech-india-politics-muslim-hindu-modi-zuckerberg-11597423346
13 See https://time.com/5883993/india-facebook-hate-speech-bjp/
14 See www.statista.com/statistics/268136/top-15-countries-based-on-number-of-facebook-users/
15 See www.statista.com/statistics/289778/countries-with-the-most-facebook-users/
16 See www.theguardian.com/technology/2020/jul/02/whatsapp-groups-conspiracy-theories-disinformation-democracy
17 See www.theguardian.com/technology/2020/jun/29/how-hate-speech-campaigners-found-facebooks-weak-spot
18 Seewww.statista.com/statistics/267031/facebooks-annual-revenue-by-segment/
19 See www.vanityfair.com/news/2020/11/facebooks-election-tweaks-curb-misinformation
20 As Chouliaraki and Banet-Weiser (2021, p. 4) argue in their introduction to a special issue on the logic of victimhood 'that victimhood today operates as a 'master' signifier, a dominant communicative logic that relies on auxiliary vocabularies – of injured white masculinity, celebrated survivorship or heroic sacrifice – to reclaim power for the powerful and retrench existing hegemonic arrangements in liberal polities. Whether it is authoritarian populists, networked misogynists or imperialist state actors, this logic of weaponized victimhood is, we contend, a crucial site of political struggle and, as such, a scholarly terrain of urgent interrogation'.
21 See www.naacp.org/latest/statement-stop-hate-profit-meeting-facebook/
22 See www.washingtonpost.com/technology/2020/12/03/facebook-hate-speech/
23 Arjun Appadurai. (2018, May 22). Across the world, genocidal states are attacking Muslims: Is Islam really their target? *Scroll*. https://scroll.in/article/879591/from-israel-to-myanmar-genocidal-projects-are-less-about-religion-and-more-about-predatory-states
24 Readers familiar with recent legislation in India, especially the Citizenship Amendment Act and the National Citizenship Register, will find similarities between how the Rohingyas have been excluded and dehumanised in Myanmar and how the Muslims in India are being targeted by the BJP government in India.
25 E/CN.4/1993/62. https://documents-dds-ny.un.org/doc/UNDOC/GEN/G93/101/09/PDF/G9310109.pdf

26 Michael Caster. (2015, August 26). The truth about Myanmar's new discriminatory laws. *The Diplomat*. https://thediplomat.com/2015/08/the-truth-about-myanmars-new-discriminatory-laws/
27 www.reuters.com/investigates/special-report/myanmar-rakhine-events/
28 Crisis Group. (2021, May 18). Myanmar's military struggles to control the virtual battlefield. *Crisis Group*. www.crisisgroup.org/asia/south-east-asia/myanmar/314-myanmars-military-struggles-control-virtual-battlefield
29 Steve Stecklow. (2018, August 15). Why Facebook is losing the war on hate speech in Myanmar. *Reuters*. www.reuters.com/investigates/special-report/myanmar-facebook-hate/
30 https://about.fb.com/news/2021/02/an-update-on-myanmar/
31 Peter Guest. (2021, February 2). How misinformation fueled a coup in Myanmar. *Rest of World*. https://restofworld.org/2021/how-misinformation-fueled-a-coup-in-myanmar/
32 Samantha Stanley. (2017, May 16). Misinformation and hate speech in Myanmar. *First Draft*. https://firstdraftnews.org/articles/misinformation-myanmar/

References

Amnesty International. (2017). *Caged without a roof: Apartheid in Myanmar's Rakhine state*. Amnesty International.

Anwary, A. (2021). Sexual violence against women as a weapon of Rohingya genocide in Myanmar. *The International Journal of Human Rights*, 1–20.

Arvidsson, A., & Colleoni, E. (2012). Value in informational capitalism and on the internet. *The Information Society*, *28*(3).

Aung-Thwin, M. (1985). *Pagan: The origins of modern Burma*. University of Hawaii Press.

Beer, D. (2009). Power through the algorithm? Participatory web cultures and the technological unconscious. *New Media & Society*, *11*(6), 985–1002.

Bucher, T. (2012). Want to be on the top? Algorithmic power and the threat of invisibility on Facebook. *New Media & Society*, *14*(7), 1164–1180.

Cady, J. (1965). *A history of modern Burma*. Cornell University Press.

Callahan, M. (2005). *Making enemies: War and state-building in Burma*. Cornell University Press.

Cheesman, N. (2017). How in Myanmar "national races" came to surpass citizenship and exclude Rohingya. *Journal of Contemporary Asia*, *47*(3).

Cheng Guan, A. (2007). Political legitimacy in Myanmar: The ethnic minority dimension. *Asian Security*, *3*(2), 121–140.

Clarke, S. L., Myint, S. A. S., & Siwa, Z. Y. (2019). *Re-examining ethnic identity in Myanmar*. CPCS Publication.

Dahlberg, L. (2015). Expanding digital divides research: A critical political economy of social media. *The Communication Review*, *18*(4), 271–293.

Darusman, M., Coomaraswamy, R., & Sidoti, C. (2019). *Detailed findings of the independent fact-finding mission on Myanmar.* https://www.ohchr.org/EN/HRBodies/HRC/MyanmarFFM/Pages/Index.aspx

Davis, A. (2021). Hate speech in Myanmar: The perfect storm. In S. Jayakumar, B. Ang, & N. D. Anwar (Eds.), *Disinformation and fake news* (pp. 103–116). Palgrave Macmillan.

Edelman, B. G., & Luca, M. (2014). *Digital discrimination: The case of Airbnb. com*. SSRN Electronic Journal. Elsevier BV.

Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police and punish the poor*. New York: Picador.

Faraj, S., & Azad, B. (2012). The materiality of technology: An affordance perspective. In P. Leonardi, B. Nardi, & J. Kallinikos (Eds.), *Materiality and organizing: Social interaction in a technological world* (pp. 238–258). Oxford University Press.

Farrelly, N., & Win, C. (2016). Inside Myanmar's turbulent transformation. *Asia & the Pacific Policy Studies*, *3*(1), 38–47.

Ferguson, J. (2015). Who's counting? Ethnicity, belonging and the national census in Myanmar. *Journal of the Humanities and Social Sciences of Southeast Asia and Oceania*, *171*(1), 1–28.

Fink, C. (2018). Dangerous speech, anti-Muslim violence, and Facebook in Myanmar. *Journal of International Affairs*, *71*(15), 43–52.

Fuchs, C. (2010). Labor in informational capitalism and on the internet. *The Information Society*, *26*(3).

Gebru, T., Krause, J., Wang, Y., Chen, D., Deng, J., Erez Lieberman Aiden & Li Fei-Fei. (2017). *Demography with deep learning and street view Proceedings of the National Academy of Sciences, 114*(50), 13108–13113.

Gibson, J. (1982). Notes on affordances. In E. Reed & R. Jones (Eds.), *Reasons for realism: Selected essays of James J. Gibson* (pp. 401–418). Lawrence Erlbaum Associates.

Guyot, J. (1989). Burma in 1998: Perestroika with a military face. *Southeast Asian Affairs*, 108–133.

Hall, S. (1996). Who needs identity? In S. Hall & P. du Gay (Eds.), *Questions of cultural identity* (pp. 1–17). Sage.

Hanafi, S. (2009). Spacio-cide: Colonial politics, invisibility, and rezoning in Palestinian territory. *Contemporary Arab Affairs*, *2*(1), 106–121.

Houtman, G. (1999). *Mental culture in Burmese crisis politics: Aung San Suu Kyi and the national league for democracy*. Tokyo University of Foreign Studies.

Hutchby, I. (2003). Affordances and the analysis of technologically mediated interaction: A response to Brian Rappert. *Sociology*, *37*(3), 581–589.

Jones, L. (2014). The political economy of Myanmar's transition. *Journal of Contemporary Asia*, *44*(1).

Kipgen, N. (2019). The Rohingya crisis: The centrality of identity and citizenship. *Journal of Muslim Minority Affairs*, *39*(1).

Kyaw, N. N. (2015). Alienation, discrimination and securitization: Legal personhood and cultural personhood of Muslims in Myanmar. *Review of Faith and International Affairs*, *13*(4), 50–59.

Lee, R. (2014). A politician, not an icon: Aung San Suu Kyi's silence on Myanmar's Muslim Rohingya. *Islam and Christian – Muslim Relations*, *25*(3).

Lee, R. (2019). Extreme speech in Myanmar: The role of state media in the Rohingya forced migration crisis. *International Journal of Communication*, *13*, 3203–3224.

Lewis, R. (2018). *Alternative influence: Broadcasting the reactionary right on YouTube*. Data & Society Research Institute.

MacLean, K. (2019). The Rohingya crisis and the practices of erasure. *Journal of Genocide Research*, *21*(1).

Mahmood, S. S., Wroe, E., Fuller, A., & Leaning, J. (2017). The Rohingya people of Myanmar: Health, human rights, and identity. *The Lancet*, *389*(10081).

McCarthy, G., & Menager, J. (2017). Gendered rumours and the Muslim scapegoat in Myanmar's transition. *Journal of Contemporary Asia*, *47*(3), 396–412.

Meehan, P. (2011). Drugs, insurgency and state-building in Myanmar: Why the drugs trade is central to Myanmar's changing political order. *Journal of Southeast Asian Studies*, *42*(3), 376–404.

Noble, S. (2018). *Algorithms of oppression*. New York University Press.

Renshaw, C. S. (2013). Democratic transformation and regional institutions: The case of Myanmar and ASEAN. *Journal of Current Southeast Asian Affairs*, *32*(1), 29–54.

Rieder, B., Matamoros-Fernández, A., & Coromina, Ò. (2018). From ranking algorithms to "ranking cultures". *Convergence: The International Journal of Research into New Media Technologies*, *24*(1), 50–68.

Rigi, J., & Prey, R. (2015). Value, rent, and the political economy of social media. *The Information Society*, *31*(5). https://doi.org/10.1080/01972243.2015.1069769

Sablosky, J. (2021). Dangerous organizations: Facebook's content moderation decisions and ethnic visibility in Myanmar. *Media, Culture & Society*, 1–26.

Sandvig, C., Hamilton, K., & Karahalios, K. (2014). Auditing algorithms: Research methods for detecting discrimination on internet platforms. In *Data and discrimination: Converting critical concerns into productive inquiry*. New America Foundation.

Selth, A. (1986). Race and resistance in Burma, 1942–1945. *Modern Asian Studies*, *20*(3), 483–507.

Smith, M. (1991). *Burma: Insurgency and the politics of ethnicity*. Zed Books.

Taylor, R. (1987). *The State in Burma*. C. Hurst.

Than, T. M. (2005). Dreams and nightmares: State building and ethnic conflict in Myanmar (Burma). In K. Snitwongse & S. W. Thompson (Eds.), *Ethnic conflicts in Southeast Asia* (pp. 65–108). ISEAS-Yusof Ishak Institute.

Ullah, A. A. (2011). Rohingya refugees to Bangladesh: Historical exclusions and contemporary marginalization. *Journal of Immigrant & Refugee Studies*, *9*(2).

Ware, A., & Laoutides, C. (2019). Myanmar's "Rohingya" conflict: Misconceptions and complexity. *Asian Affairs*, *50*(1). https://doi.org/10.1080/03068374.2019.1567102

Whitten-Woodring, J., Kleinberg, M. S., Thawnghmung, A., & Thitsar, M. T. (2020). Poison if you don't know how to use it: Facebook, democracy, and human rights in Myanmar. *The International Journal of Press/Politics*, *25*(3), 407–425.

Yiftachel, O. (2005). Neither two states, nor one: The disengagement and "creeping apartheid" in Israel/Palestine. *The Arab World Geographer*, *8*(3), 125–129.

Zarni, M., & Cowley, A. (2014). The slow-burning genocide of Myanmar's Rohingya. *Pacific Rim Law and Policy*, *23*(3), 681–752.

3 Brazil colonisation, violent 'othering' and contemporary online hate

Introduction

This chapter presents a case study of Brazil where modest political reform led by previous regimes – for instance, with regard to food security under the Worker's Party or LGBTQIA+ rights – have been followed by a virulent rightwing backlash. This backlash is intimately connected to the rising popularity of the Brazilian far right whose authoritarian politics are most notably personified by Jair Bolsonaro. Bolsonaro's rise to power and his current grip on the political imagination of at least a third of the voting public, was supported and accompanied by the systematic production and reproduction of hateful discourse and action. This systematic hate was deployed against specific groups – including the working classes, Indigenous populations, Afro-Brazilians, the left, LGBTQIA+ groups – always exacerbating existing lines of racism and ethnic difference. From the *creation of difference* to a process of consolidation of political and economic power, the production of targets for populist rage and violence has been an extended socio-political process. This process includes but is not confined to the co-optation and misuse of state institutions and actors specific to Brazilian history, culture and politics, and the rapid demonisation and delegitimisation of groups and individuals who publicly stand for peace, equity and social justice through reference to traits of identity, character or politics which act as excuses for violence.

Within this milieu, social media use including Facebook, YouTube, Twitter, WhatsApp, Instagram and TikTok for the circulation of hate and misinformation has further intensified violence through the seeding of disinformation that deliberately deceives specific groups about their fellow Brazilians. The intimacy afforded by social media spaces and the illusion of privacy on some cross-platform apps has led to the emergence of new forms of political trolling where the identity traits of those championing social justice and human rights (such as religion, gender, race, sexuality or disability) are 'discovered'

DOI: 10.4324/9781003083078-3

by organised far right political actors and used to build narratives of deceit and contamination, whose solution appears to be the annihilation of the voices and bodies of the 'other' from the public sphere. For instance

> The story that was spread via Bolsonaro's social media channels was that Haddad had created a "gay kit" which he planned to introduce in primary schools so that children from the age of six "would be encouraged to become gay". In fact, Haddad, as Brazil's former minister for education – and alongside other politicians – had promoted an educational programme for primary school students to understand sexual diversity and combat homophobia.[1]

This hateful and mendacious but politically effective campaign culminated in deep-fake videos of Fernando Haddad implicating him and his supporters in 'deviant' sexual practices and, in particular, paedophilia, that were circulated in a targeted manner towards sections of the Catholic and Evangelical population, and in particular towards middle-aged women, swinging them away from the Workers Party at the last minute.

Despite many brave, reputable and insightful journalists, the mainstream media in Brazil, meanwhile, have played a deeply problematic role in maintaining various forms of colonial thinking, white supremacism and heterosexual privilege. Built on a model of private investment and weak regulation since the early 1990s and in a manner startlingly similar to India, even the few supposedly educational media outlets are run by quasi-religious or oligarchic interests: 'Brazil has always had a weak public media sector which has been composed mainly of the respected but funded-starved TV Cultura in SP and its counter-part TVE in Rio, as well as other regional outlets controlled by local politicians and by sectors of the evangelical Church' (Matos, 2011, p. 7). While TVE has since ceased to broadcast and the primary public channel now is TV Brasil, the broad point still holds good. In some regions a single family with affiliation to or entanglement with conservative politicians controls all of the media outlets, and therefore the political messages. As Alfonso de Albuquerque argues powerfully with regard to Brazil,

> elites and their media portray themselves as a westernized minority endowed with a civilizing mission regarding their societies as a whole, and manipulate the Fourth Estate discourse toward their own benefit, as a means for securing and legitimizing their own privilege.
>
> (2017, p. 906)

Associating the western democratic tradition with capitalism in an unregulated form, these elites use mainstream news media to spin their attempts

to suborn and subvert the institutions of accountability as 'pro-democratic' when their targets are Indigenous populations or leftist politicians. Sponholz and Christofoletti's informative analysis traces the ways in which co-optations and rightwing ownership models have led to a situation whereby 'the media system has been enabling public figures to use hate speech to enhance their media prominence' (2019, p. 67). The toppling of the left government and the ascendence of rightwing forces who utilise hate speech is thus, in their view, a combination of social and media capital that pre-dates the embedding of social media as a popular tool.

As we discussed in Chapters 1 and 2, the hate, disinformation and misinformation circulated on social media are not isolated phenomena. Nor do they just appear out of the blue in the online world as by-products of the affordances of new and emerging technologies. Rather, social media, hate and disinformation spring from complex and intersecting histories and genealogies. They are embedded in and fuelled by structural and material violence, racism, prejudice, discrimination and inequality. Accompanying these adverse processes and behaviours, we begin to see more recently the entry of a techno-communicative development model, untethered from an ethics of rights and solidarity. It is a model that thrives because of social media's value in advancing neoliberal capital's agendas and in empowering surveillant, carceral states. In this chapter, we ground our analysis of the phenomena of hate-filled trolling, disinformation and misinformation in the powerful narratives of a cross-section of Brazilian citizens whose work, identities or politics have 'resulted' in their experiences of extreme trolling, racism, misogyny, homophobia, death threats and/ or physical attacks and violence. Before diving into these cases, however, we provide our readers with some basic historical background.

Histories of violence: Colonialism to the present

Prior to the 16th century Brazil was inhabited and controlled by Indigenous peoples, comprising over ten million and distributed between two thousand tribes, some constituted into discrete nations and occupying specific parts of the continent, others semi-nomadic, and moving between the coast and the interior. Struggles for power were common between and amongst the largest of these groups – for instance, Tupis and Tapuias – who tended to settle along the coastal regions and in the interior respectively. Many of the groups contained numerous subgroups grappling for internal dominance.

Following a 1494 treaty between Spain and Portugal which arbitrarily allocated all territories east of a certain line and west of that line in the landmass of South America to Portugal or Spain respectively, from 1500 onwards Portuguese explorers and traders led by Pedro Alvares Cabral opened a pathway to the decimation and pillage of Indigenous territories. In the subsequent

centuries, millions of Indigenous people were subjected to colonial genocide. This happened both directly – through massacres, the sacking of their tribal homelands, and violent take-over of their land – and indirectly, through the transmission of deadly illnesses that swept across their populations, the hunger that ensued, and the suppression of their Indigenous spiritual practices by swathes of Jesuits and other European Christian priests.

The white Portuguese invaders and colonisers treated Indigenous South American populations with extreme contempt, deploying mass suppression, erasure, violence and multiple forms of dehumanisation common across Belgian, British, Dutch, French, German, Italian, Spanish, and other colonial systems. Local populations responded by fighting back fiercely, going underground, doing deals with opposing colonial powers in internecine struggles and, on occasion, intermarrying with or joining forces with escaped parties of African rebel slaves transported to South America by the slave trading colonisers.

Despite early lack of interest in their new colony which seemed simply too unyielding and distant, the Portuguese soon discovered that sugarcane – a then almost priceless commodity – could be grown in abundance. Thus began several centuries of slave-trading and plantation slavery to enhance Portugal's monopoly in the global sugar trade. We will not dwell here on the various European and colonial wars for ascendancy that occurred between the Portuguese, French, Dutch and British over parts of the Brazilian territory, and over the slave, sugar, gold, diamond and coffee trades, as these are all detailed with incredible nuance in histories such as those by Braudel (1981) and Puntoni (2019). Highly relevant to our volume, however, are the hidden groups of escaped African slaves, who began to form communities of *quilombos* (Anderson, 1996; Ferretti, 2019), gathering strength from the increasing numbers of slaves transported to Brazil during until the middle of the 19th century (Brazil was one of the very last countries to officially abolish slavery in 1888). Although there continue to be contests over the histories of quilombo descendants in modern Brazil, up to the mid-19th century these subaltern communities frequently banded together in insurrectionary movements fighting off the colonisers, sometimes with help from surviving Indigenous populations.

The rebels were often brutally suppressed but also never surrendered and were thus increasingly successful, particularly once the abolitionist movement took hold amongst the European middle classes. Intermarriage and sex (both coerced and consenting) between African, Indigenous and Caucasian European populations was common. New creole languages developed which mirrored the mixed communities living outside the purview of the Portuguese crown. Although there are alternate histories which point to multiple complicities with the white state, many of these mixed communities continued to operate in the 20th century as focal points for socialist

mobilisation, particularly in urban areas, and for African and Indigenous spiritual practices that contested the stranglehold of the Catholic church.

In line with Colin Snider (2018), our work takes 'an expanded approach . . . that explores the ways in which memories, discourses, and policies from military regimes continue to shape politics, society, and discourse decades after militaries left power' (p55). Therefore, as we highlight historical events and processes primarily in order to elucidate the *complex contexts and theoretical underpinnings of hate* in contemporary Brazil, we skip ahead now to the 1960s, which saw the beginning of two decades of military dictatorship. Supported by Johnson's administration in the United States with its virulent anti-communist propaganda, Brazil's democratically elected left-wing government was brought down in the coup of 1964 and a spate of relentless state killings, disappearances and torture began crushing dissent (Chirio, 2018; Schneider, 2011). Unquestionably propped up by imperialist military support and intervention from the United States and by the highest ranks of the Catholic church (which shed its earlier institutional connections to socialist groups), the military dictatorship lasted until 1988. Democracy only returned to Brazil as a result of almost unthinkable suffering and courageous protest on the part of multiple civil society groups. Unfortunately, there was no concomitant widespread change to social values and attitudes to race, gender and violence which had been shaped under colonial and then dictatorial rule. The wounds and scars from these years run deep, and influence much of contemporary politics and social life.

Founded in 1980, the Workers Party is the largest socialist or social democratic party in South America. During the 1980s, when unemployment and inflation were at an all-time high and violence was rife in the urban factories and favelas, many grassroots union organisers were threatened, beaten and arrested for organising protests and strikes. Amongst these community organisers was Luiz Inácio Lula da Silva, popularly known as Lula, who became the leader of the Workers' Party and the 35th President of Brazil from 2003 until his former chief of staff, economist Dilma Rousseff took over in 2011. Lula's and Dilma's regimes oversaw some of the most sweeping pro-people socio-economic reforms in Brazil's history. These were fiercely opposed by US corporations and the rightwing politicians allied to US interests.

Ironically, corrupt prosecutions over apparent corruption saw Dilma ousted in 2016, at which point disagreements over the agenda of the Workers' party and its opponents spilled over into cyberspace. When the mainstream media focused on the highly politicised corruption scandals — foregrounded to undermine social democratic reforms — and rumours spread on social media, in quick succession Lula was imprisoned followed by the impeachment of Dilma and the electoral hobbling of the Workers party. A former member of the military during the dictatorship, famed for his homophobic, racist and

misogynist views, Jair Bolsonaro came to power in 2016 on a tide of violent bigotry.[2] While his military service and nostalgia for the dictatorship era are a matter of public knowledge, his connections to Steve Bannon and US far right networks are less well known. The misogynist and white supremacist legacy which brought Bolsonaro to power also includes the vigilante murders of multiple Indigenous land defenders, women politicians and Black trans folk, more than half a million Covid deaths, multiple complex sets of disinformation emanating from the dreaded "Office of Hate"[3] and a flood of homophobic deep-fakes against Bolsonaro's opponents.

Brazil always has been and remains a ferociously racist society, with stratification between Black and Indigenous Brazilians and their white counterparts now at an all-time high, and overlaps between misogyny, homophobia and racism making certain communities and people even more vulnerable. While no modern social history of Brazil can ignore the issue of race, the unfounded romanticist myth of "racial democracy" (the idea that the three racial groups in Brazil – Indigenous peoples, African and European descendants – enjoy equal rights and that racism *per se* does not exist) has dominated many accounts since the work of Gilberto Freyre in *Masters and Slaves*. Now challenged both by accounts of Afro-Brazilian lived experience (Keisha-Khan, 2004; Sheriff, 2001) and by historical and theoretical analysis of racial hierarchies and practices (Anderson et al., 2019; de Vasconcelos, 2019), this myth continues to serve as a salve in outward-facing international discussions of race relations, but is no longer believed with conviction amongst affected populations. As Keisha-Khan points out,

> Resistance to urban renewal plans in Salvador demonstrates how struggles for urban land rights are a crucial part of engaging in the broader national and international politics of race. In Black communities in Brazil and throughout the African diaspora, urban land and territorial rights are the local idioms of Black resistance.
>
> (2004, p. 811)

Each of these historical threads informs our analysis of the data in coming sections.

Will to power as the root, hatred as the tree

In addition to being professionals who found themselves unexpectedly in the eye of a political storm (journalists, doctors, fact-checkers) or spiritual leaders, academics and social activists whose work for social justice made them targets, our interviewees in Brazil often had *intersecting lived experience* of being members of groups against whom hate in Brazil has reached

new peaks – Indigenous, LGBTQIA+, Black, women leaders, political dissidents. Gradually as their stories and words flowed across our screens, the picture which took shape became clearer and clearer: The legitimisation of discrimination, dehumanisation and violence by powerful authorities and government figures is linked to a steep rise in hate crimes and hate speech online in Brazil. Beatriz Buarque, a journalist, researcher and founder of NGO *Words Heal the World* consistently described how the views, values and behaviours of powerful leaders in government are linked to social media hate amongst the populace:

> [M]ost hate crimes [in Brazil today] are generated by racial prejudice. Gender comes in second place. In third place, we find homophobic crimes. On social media, we see a higher incidence of racist messages, homophobic messages and political-ideological messages, especially after Bolsonaro's election. We have a president who reinforces and legitimises this kind of narrative. Bolsonaro and his ministers are very active on Twitter and they use this platform to post hate messages against the left, as if any person who diverged from the government was an enemy. This is their narrative. Besides the ideological side, there's also the racist one – the government is composed mostly of white people. They don't talk about diversity. They are extremely homophobic. During Carnival, the president himself reinforced stereotypes that exist in relation to homosexuals. When authorities legitimise hate speech, we expect that some people will start reproducing this behaviour. We are seeing far-right groups proliferating; groups that actively hate homosexuals. And they don't even use subtle language. On Facebook you will find posts defending 'killing gays and lesbians in the name of a clean society, in the name of a Christian society' . . . African or African-Brazilian religions are frequently attacked. This is closely linked to racism too. These religions have been historically delegitimised, they are often said to be linked to the devil . . . This narrative feeds religious fanatics who are against diversity. So Brazil has an interconnected society that uses social media to express every desire and hate, without any fear of punishment. We have the president and other authorities legitimising these stereotypes and discourses.

Our interviewees and other key informants were uniform in their description of the communities and individuals most targeted by hate. Sonja Guajajara who is finishing her second term as Executive Coordinator of a group of Brazil's Indigenous People (APIB) also confirmed that hate is directed largely at 'Left-wing groups, Indigenous people, Black people, LGBTQIA+ and women, with Indigenous people and Black people leading this ranking... people tell me I'm incompetent because I'm indigenous, that I'm fat, they tell me "go back to

the jungle because that is your place"'. Although they gave different accounts of the role of platforms and technology in facilitating this, there was agreement about the root causes and perpetrators (see Table 1.1 in Chapter 1). As in India, the situation in which discrimination, hate speech, dehumanisation and online incitement to violence appear to carry no social stigma for perpetrators let alone any criminal proceedings, and are sanctioned by members of the ruling party, has led to ever increasing attacks.

Gilberto Scofield, a journalist who has worked for the major news organisations such as Globo and now manages a well-established fact-checking agency told us that the attacks against him were deeply personal. The attackers use every aspect of his private life to delegitimise his public role and political stances or interventions. The *less visible* he tries to be as an individual in order to assist in the work of his organisation, the more he is made the personal focus, with accusations of being an LGBTQIA+ activist, with details of his private life publicised in malicious and destructive ways by organised trolls. His attitude now is not to hide anything in his online presence. Nor does he publicise anything about his identity online. He tries to exist and retain his integrity, regardless of the vitriol:

> I'm gay. I've been out of the closet my whole life. These hate posts always say I'm an LGBT activist. I don't consider myself an LGBT activist, I consider myself a normal person who talks about his personal life. People talk about their lives, their husbands, their kids and their career. Why can't I talk about these subjects too? We have problems as well as virtues. We're normal human beings. For most of my career, I covered business/economic issues – a sector that is extremely closed to diversity and is very uptight. Journalists who are gay are usually covering fashion, culture, music. It's a kind of closet for gay journalists. I met [my husband] 17 years ago. We got married and we adopted two kids.

However, while the strategy of being open but keeping a lower profile online might maintain the careful balance of privacy for his children, he cannot prevent them from experiencing anti-Black racism because of the colour of their skin, their features. The accounts he gave of their experiences tally with accounts from the United States (Bailey, 2021; Sethi, 2018) and UK (in this book), and show the extent of everyday discrimination, microaggression and prejudice both in the association of Blackness with criminality and in the association of Blackness with poverty and ghettos.

> Racism is a major issue in Brazil. Once, my son was coming back from school and, when he tried to enter our building, a woman was going out and she told him he didn't live there. My husband arrived and asked what was happening and the lady told him 'this boy was trying to get in

> the building, but he doesn't live here'. So my husband told her that he was our son. She felt embarrassed and tried to justify herself. It turned out she was the one who didn't live in the building, she was just visiting someone. Even not living there, she assumed a Black child would never live in that building. *Things like that happen all the time*. Once we were swimming in a pool and I heard a [white] teenager saying 'this place is going downhill, there's a Black kid in the pool'.

Gilberto and his fellow fact-checkers are also targeted precisely when they are most successful at managing to get hate speech taken down or called out online:

> When you go on Facebook, there are three dots next to every post where you can click and report that post. One option is to report it as fake news. Every time someone chooses that option, Facebook sends that post to us so we can fact-check it. We analyse them based on two guidelines: How relevant is the person who is posting and how many times the post was shared. If it's fake, we tell this to Facebook. They don't delete the post, but every time a person sees that content, there will be a pop-up saying that our agency has fact-checked it and found out it is fake. Also, they change the algorithm so fewer people can see the post. Hence, people who want that content to be seen and shared get very angry; they say it is censorship, they say it's an attack to their freedom of expression. If you think hate speech is freedom of expression, I'm sorry, it's not. There are ethical and moral limits for everything in life.

Indeed, Gilberto's children's safety is often threatened simply because of his work. He has begged them not to go on social media in order to protect them from further exposure to the most vile abuse and also to keep their images from being circulated and doctored by fascist trolls.

> The right uses disinformation as a strategy. The groups that support the government using bots are very active, they are professionals. Some don't necessarily share the ideology, they are being paid to do that. So we avoid posting photos showing where we live, we avoid posting family photos. We are careful not to say where our kids go to school. Attacking our loved ones is part of their strategy. Some people send emails saying 'I know where your kid goes to school' – this is more common than you imagine, it happens all the time . . . There's another fact-checking agency that live fact-checked Bolsonaro's speech at the UN and they are still being attacked for that.

Another well-known journalist and fact-checker anonymised here (at her request) as Antonia P, has been forced into exile after being targeted by Bolsonaro's regime. Asked what kinds of messages she and her female colleagues received, Antonia went on to name others who had faced such attacks and to associate the attacks with any attempt on the part of women to take a role in Brazil's public sphere either professionally or through socialist politics:

> They [the far right] call you a hooker and worse. It is always an attack against our bodies, our existence, our mental capability. It's always linked to sex, like what happened to Patrícia Campos Mello,[4] Miriam Leitão, Vera Magalhães . . . It doesn't happen only to journalists. Maria do Rosário, Manuela D'Ávila,[5] the list of women who've been through this is endless. . . . There was an account with a statue in the profile picture that sent me an inbox message every day, saying 'I'll shoot you in the face, you whore'. It's always a statue, a comic strip, but it's not a bot because they react to reality, they talk.

Antonia's verdict: '"Techno-populism" is highly responsible for online hate' did not differ much from the narratives we were hearing in other countries, except in its insistence on the culpability of the left too in the growth of misinformation. Another interviewee, Mariana (a pseudonym, as it is not safe for her or her family if we use her real name), reiterated that misogyny becomes a blunt weapon of the far right in inciting hate and violence towards liberal or leftist women who comment on politics. When the integrity and courage of her work – which involved critical engagement with the public sphere and evaluations of the accuracy or mendacity of political communication – initially triggered an avalanche of hate speech and incitement against her online, much of that hate took the form of misogyny:

> They cursed me – the usual social media behaviour . . . things started escalating. I received messages mentioning my son; telling me that, if I wanted my son to be safe, I should leave the country. I reported it to the police. A person called my mobile saying I was going to get punched in the face; my schedule was shared in WhatsApp groups as an encouragement for people to confront me face-to-face. Fake news about me was also being spread; people were sharing a doctored photo . . . People recognised me in the streets . . . We were afraid someone would do something physical, so [my work] hired me a security guard. . . . People started spreading horrible memes of me, indescribable memes; they put my face on a naked woman with her legs open, with the most disgusting and pornographic subtitles; Another one was a naked woman with

> my face and a line of men waiting their turn. I also received messages saying I deserved to be raped. It was so overwhelming.

Mariana's ability to survive a routine dose of trolling which she describes as 'the usual social media behaviour' suggests that we need to distinguish between different forms of trolling. It is the multiple forms and sites of aggression, the threats, the formal and substantive connections between politicians and vigilante publics (Banaji, 2018), that makes the fascist social media sphere of hate quite distinct from some of the early troll behaviour from individuals and groups on message boards in the 1990s and early 2000s, and connects it to older and more organised forms of Nazi, Fascist and white supremacist action and propaganda from pre-Internet days.

Another interviewee, philosopher and anti-racist activist, Djamila Ribeiro has authored three books on racism and Black feminism, and writes a weekly newspaper column. Djamila is very active on Facebook as well as on Instagram, where she has a million followers. Since she posts frequently about violence against women, abortion and racism, she told us that she is used to getting hate comments and messages, and has come to expect them, getting attacked primarily by the right but even on occasion by those ostensibly belonging to the left. This year, she was attacked by a leftist woman who accused her of being against the working classes after she participated in a sponsored post for a taxi company. She was attacked on Twitter (even though she doesn't have an account) and on other platforms. Someone sent messages to her 15 year-old daughter's phone saying: 'Your mother is a disgrace! Aren't you ashamed of being her daughter? We know where your mother lives! There is no way out!', at which point Djamila reported the case to the police but to no avail.

Along with other activist organisations, she also brought a lawsuit against Twitter arguing that the platform benefits economically from racist and misogynist attacks. While she was clear with us that the structure and protocols of spaces like Twitter and Facebook make them 'toxic environments' for Black women in particular, she also emphasised how the takeover of politics and mainstream media by the right and far right were fuelling overlapping forms of hate and violence:

> Since the coup against president Dilma, the polarisation has intensified. It was partly encouraged by the hegemonic media that has definitely contributed to the polarised discourse. But I think it can be traced back to the criminalisation process imposed against the Labour Party.[6] Now, the current government encourages hate against other parties' members, against people who don't share their ideology and

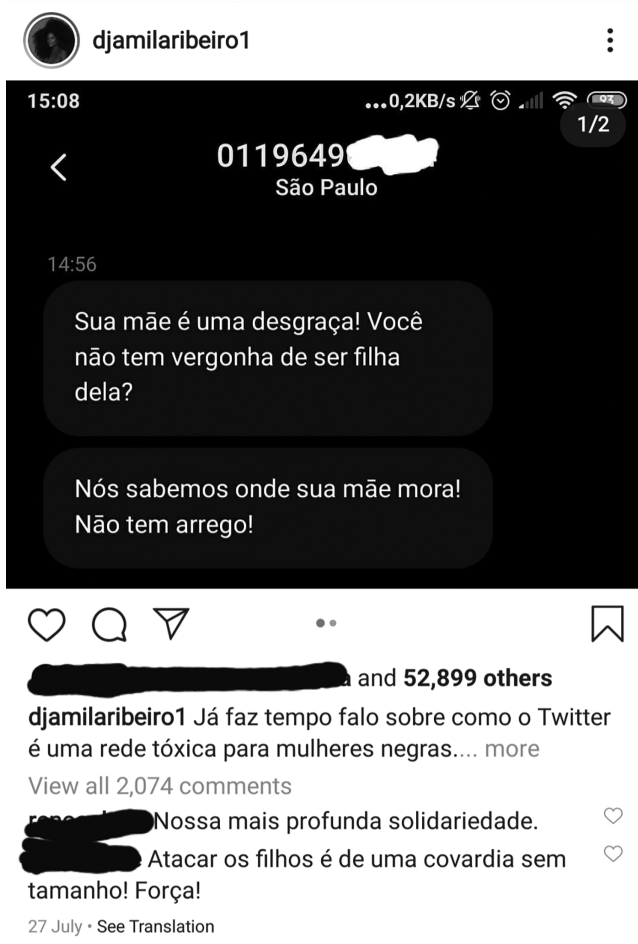

Figure 3.1 Sample of hateful material received online by Djamila Ribeiro. Credit: Djamila Ribeiro.

against intellectuals. There is a clear anti-intellectualism. So there is this political side of it, attacks against left-wing politicians. Manuela D'Ávila,[7] for example, has been victim of several attacks . . . These spaces are not detached, since these are structural discriminations; therefore, they structure all social relations. Social media platforms are also spaces where hate speech is spread against historically discriminated groups.

A repeated theme in our interviews both in Brazil and elsewhere was that historically marginalised and discriminated populations, who have faced vigilante violence and atrocity from both state and non-state actors over the decades, also bear the brunt of hate online. De Vasconcelos summarises the multiple overlaps of erasure, violence, exclusion, inequality and discrimination when it comes to Brazil's African descendant population, and connects current circumstances to the history of slavery:

> According to Brazil's National Statistics Institute (IBGE), 53% of the population identify as Black or mixed. According to the Atlas of Violence 2018, the Black homicide rate was more than double that of non-Blacks (40.2% versus 16.0%). A Black person is murdered in Brazil every 23 minutes. According to Oxfam International (2017) . . . white Brazilians earn twice as much as Black Brazilians. Between 2003 and 2013, the number of Black women murdered rose by 54%, while the white femicide rate fell by 10% . . . . Only 10% of Brazilian books published between 1965 and 2014 were written by Black authors. *The Face of National Cinema* . . . (UFRJ), revealed that of all Brazilian film directors, only 2% are Black men, and none are Black women. The racial gap also exists among writers, where only 4% are Black. Of all the films and novels analysed, only 31% had Black actors in the cast, in which they commonly play roles associated with poverty and crime . . . In Brazil, slavery persisted for a longer period of time and over a larger geographical area than anywhere else in the world. Forty percent of enslaved Africans sold between the 16th and 19th century landed there. The presence of Black people influenced customs, language, religion and culture. For almost five centuries of Brazilian history, the Black community has been marginalised.
>
> (2019, pp. 1–2)

Whatever efforts have been made by progressive governments to redistribute resources, their overall impact has not been the kind of decolonial approach that would radically alter such statistics. And, although the Black Lives Matter movement has seen a surge of support in Brazil, community organisers and social justice activists are also in contention with one of the most racist regimes in living history. The tentacles of this racism extend into every area of life, including pedagogy and religion.

A war between narratives: Colonial dehumanisation versus liberation

Afro-Brazilian religions face daily racism and intolerance in Brazil. Catholic priests persist in trying to delegitimise them, linking them to witchcraft

and the devil. Bolsonaro-aligned evangelical churches openly preach against them, often urging evangelical congregations to act violently against them. Some studies show the connection between evangelical churches and drug militias,[8] which makes it even more dangerous for Afro-Brazilian religious centres to remain open. Further emphasising the connections between violence against African descendant individuals and communities in Brazil with religious discrimination, our interviewee MER was a religious practitioner of the Afro-Brazilian religions Candomblé and Umbanda which gained popularity amongst autonomous communities in the 19th century. Leader of a religious house in São Paulo and holding a PhD in Theology with a focus on Afro-Brazilian religions, MER explained to us her position and activism as a white ally in the fight against religious racism. Evangelical Christian leaders are increasingly occupying political space; they represent one of Bolsonaro's main support pillars. Their official sanction by the government makes it easier for them to spread hate against Afro-Brazilian religions and cultures. Sparked by her postings on these topics on social media, MER has received hate messages calling her a whore and a charlatan. She reported only one of the posts to the police.

MER was forced to postpone her interview with us when someone painted a swastika on the wall of her residence. Other Afro-Brazilian religious centres have been attacked and even burned to the ground. MER's explanation of her work sheds further light on the divides in Brazilian society:

> I'm both Umbanda and Candomblé. Modern society is used to thinking based on written tradition; our world conception is rational and systemic – there is a beginning and an ending. Umbanda and Candomblé are not based on written tradition and, therefore, they don't follow that linear mindset. They also don't have a central power, they are polycentric. Thus, Afro-Brazilian religions are very diverse. However, there is some common ground, there is a shared skeleton. So they have a generational transmission and they are based on oral tradition, which doesn't mean that we don't have a secondary written tradition. We do have books, but they are not the main axis. One of their central characteristics is religious trance. We value our ancestral memory, so it is really important to talk about ancestry and to experience it during our rituals. We are based on a circular time, a mythical time. It's not like Christianity that you can say it was invented 2,000 years ago. We acknowledge two worlds that coexist: A human world and a supernatural world that intervenes directly in our lives. The divine is not outside people, we incorporate the divine, and we go into a trance with the divine. Simplistically, this is the best way to explain Umbanda and Candomblé . . . . I don't post my rituals, but since 2018 I tend to post the opening speeches I do before the rituals. These are 10, 15 minutes

> videos in which I talk about spirituality. Every time I talked about Exu, I got a lot of criticism but also a lot of sympathy . . . People tend to link Exu to the devil. Along with the colonisation process, there was also colonisation of religious values. So Exu was linked to the devil, this is still present in people's imaginary. When the pandemic came and I started posting more often, there were people calling me macumbeira[9] trying to offend me, but I would just joke and say 'I love being a macumbeira!'. . . . I avoid using key words because of the bots, I never say 'Bolsonaro', 'hate'. When we talk about Black people, poor people, we get a lot of attacks. But I still talk about that. Instead of saying 'hate', I talk about peace, culture and otherness, empathy. When you say 'Afro', 'Umbanda', 'Candomblé', 'Exu', you trigger these people.

Many of MER's postings are popular and receive praise and solidarity. She has a following across the globe. So, for her, it's imperative not to get frightened off, but to continue engaging and posting content, despite the risks. Commenting on the vicious animosity towards Afro-Brazilian religious beliefs and communities, MER explained:

> The main reason is eugenics. Black people were objectified in order to be enslaved. So their culture didn't have any value because they couldn't be seen as human beings with values. The idea that everything that comes from either Black or Indigenous culture is negative is still very present in Brazil. . . . When we talk about evangelicals, we are talking about a process of white supremacy including a high number of Black people who have converted to evangelical religions. Studies show that there's a mentality that Black people feel less Black when they are part of an evangelical church. . . . There is the idea that we have to evolve, I can't stand that. I tell these people that this is our choice, I didn't choose this religion because I didn't have other options. This idea is completely connected to religious racism: The idea that Black people's culture is primitive. Evangelicals are mostly fundamentalists and they want to silence our voices, so they attack us. We are also attacked because of political reasons when I talk about eugenics, homophobia, gender and racial issues.

So, alongside the new populist rhetorics of the far right, the complicity of the church and the criminalisation of dissent, it is clear that the majority of those who face the worst violence for their social justice and pro-democracy activism share overlapping histories of marginalisation or oppression, and intersecting demographic characteristics being either LGBTQIA+ and/or Black and/or Indigenous in terms of their presentation and cultural or

spiritual identifications. Like MER, Djamila had also initially gone online to connect with communities with whose experiences she felt solidarity and with whom she shared common praxis. She explained the trajectory of her own online practice – from blogging to social media – and the daily barrage of hatred that she faces as a consequence:

> I started writing for a blog for Black women, in 2013. It was a very cool website that promoted meetings with Black women and encouraged us to write about them. In 2014, I started writing for Carta Capital magazine, so I started posting my columns on social media. At first, I was scared when I saw the kind of reactions some of my columns generated on social media. . . . If the post is about violence against women, I get many aggressive comments, mostly from men. At the same time, these posts also engage people who are more aware and on our side. . . . when I talk about sexual violence or rape, men feel more comfortable to attack me. I think this is a curious aspect. They feel more comfortable to say that it is a lie, that women overreact. About racial issues, there is the myth that Brazilians are not racists, so they are more ashamed to do that. Racist attacks happen, but they are more restrained.

In Djamila's view the complexities of people's overt reactions to her feminist and anti-racist postings are reflective also of their beliefs about Brazil as a society, of their socialisation into spaces of violent male dominance, of Bolsonaro's encouragement of anti-intellectualism and of the demonisation of Brazil's left movements. One particularly pernicious aspect of being harassed and violated on social media remains that those spaces for thought and debate amongst marginalised communities which were welcomed and often saved or still save lives, giving people the strength to resist and to organise in ways that mainstream media has long failed to do, have now become sites of further trauma and anxiety.

Another interviewee, Wari'u Tseremey'wa, who is Xavante from Mato Grosso in the Central-West region of Brazil, uses emerging technology and graphic design to challenge and rebalance misrepresentations of Indigenous cultures. He uses his personal knowledge and experience of Indigenous politics to correct the preconceptions of those he comes into contact with. In 2018, while still a teenager, he started the YouTube channel 'Wariu' where he posts videos explaining what it means to be Indigenous in today's Brazil. Patiently and in a calm, soothing tone of voice, he demystifies preconceived ideas that non-Indigenous Brazilians hold about indigenous people[10]:

> I am from the Parabubure Indigenous territory, near Campinápolis, in the state of Mato Grosso. . . . I am Xavante mixed with Guaraní, but

> the Xavante culture doesn't allow any mixture. . . [This] means that, if the Dad is Xavante, the descendant will only be Xavante. My Mum is Guaraní, but my name is Xavante and the cultural traditions I follow are Xavante. It's the kind of protection we have . . . to avoid cultural loss. We do have relationships with other Indigenous people, even blood relationships, but we keep our culture intact. This is not a general Indigenous rule; the Xavantes were always 'war people', so these strategies exist to protect us, our culture, our people. I belong to one of the most traditional Xavante families, my ancestors were important in Xavante history; even today, my father and I still make history, in a way. My father is the president of the Federation of Indigenous People and Organisations of Mato Grosso (Fepoimt) . . . . As an Indigenous communicator, I currently do communications for the Federation of Indigenous People and Organisations of Mato Grosso, where I organise events and contribute to the national Indigenous movement. We are always in contact with Apib (Brazil's Indigenous People Articulation). That is an Indigenous social and political movement that is often ignored.

Many of Wari'u's postings on YouTube have drawn positive comments and interactions, with some Brazilians 'surprised' to see a young Indigenous person explaining their culture so calmly and clearly, and others joining him in critique of Portuguese colonial mindsets. There has also been hate, particularly on Twitter, which is a platform that he rarely uses because of its lack of enclosed community. Wari'u's response to hate received online is tempered by his constant experiences of dehumanisation and racist stereotyping since his earliest years:

> I have 25,000 followers [on YouTube] and almost 300,000 views. On Instagram, I have 14,000 followers. . . . I don't usually feel so bad with online comments because *I've already heard terrible things being said to my face*. Sometimes people send comments, but I just think 'this person is such a coward hiding behind a screen' . . . . Once, a teacher told me to "go back to the forest". She had a joking tone of voice, she said it was only a joke. . . . Sometimes online hate is not very clear for me because I've faced hate in person. People were racist in front of me, shamelessly. On Twitter, I've already received this kind of comment, but I see myself in control of my content. In order to explain my behavior, I need to tell you a bit about my culture. Xavantes are taught to resist since they are little. Our rituals test us. In a rite of passage, you spend one month in cold water, having water thrown on your ears. After this, they pierce your ears with a jaguar's bone. In another ritual, they

> give us a borduna[11] and we need to take care of it as if it was our baby. So the rituals test us physically and psychologically. . . . When I think it is hard now, I remember it was harder back when I was a child. . . . The country's president defends this old colonial narrative. . . . On one of the videos I posted on Instagram, I poked the Portuguese a little bit. Brazilians joined us Indigenous people to criticise the Portuguese. It was like 'post a moment when everything went wrong' and I showed an image of Portuguese boats arriving in Brazil. The Portuguese have a perverted idea of Brazil's colonisation. They learn in school that they were the heroes, it's shocking. So Portuguese people commented 'if we hadn't colonised you, you wouldn't be where you are today' . . . . This colonial narrative has always imposed itself, so when we bring our Indigenous narrative, it bothers some people. And they attack us to protect their narrative, they have their own interests. All information has an underlying bias. This is the Indigenous fight now – to show our side of the narrative. The other narrative justifies people taking our lands, justifies a development that can't really be called development, it's a destructive development.

In this fascinating account, a new form of resistance to colonial and racist dominance is outlined in the way of life of Xavante and of this young representative of the community. The mental tenacity cultivated in order to survive, to struggle against ongoing attempts at erasure and dehumanisation enables Wari'u to deal placidly with hate speech or belittling and stereotyping online. While complex intersections of Indigenous, Afro-Brazilian spiritual, Black, LGBTQIA+ communities and women are longstanding targets – experiencing hate and harassment in venues both real and virtual regardless of their professions – some individuals find themselves in entirely uncharted waters.

Marcus Lacerda, a doctor and researcher specialising in infectious diseases, based in Manaus, capital of the state of Amazonas, in the North of Brazil, led a clinical research trial with Covid-19 patients. His team's first finding was that high dosages of chloroquine were dangerous for patients.[12]

> When the pandemic came, I started posting videos explaining aspects of the disease and people liked it, they thought it was important at that moment. So I got many followers on Instagram. . . . Our study ended up being the first randomised clinical trial with chloroquine in the world. Its visibility increased because it was the first one. The first publication was a preliminary one; we had 81 patients at that time and it was already clear that the highest dosage we tested had toxicity problems. So we thought it was necessary to publish quickly and tell the world

> that a high dosage isn't safe. Our idea was to block the studies that were being conducted with high dosages; there were doctors all over the world using these high dosages. So we published this preliminary report and continued the study with lower dosages, which had proven to be at least safe . . . We had to interrupt most of these studies when the Ministry of Health started recommending the use of chloroquine. . . . Our conclusion was that the use in high dosages had an important toxicity.

At the outset the team had hoped to save lives. Instead, they suddenly found themselves objects of national suspicion and hatred. The day after they published their first findings, Eduardo Bolsonaro, the president's son, tweeted calling the researchers murderers and saying they were from the Worker's Party (Lula's party). Marcus immediately received hate messages on Instagram and had to acquire armed security. His team was threatened and as the controversy grew many researchers became depressed and were unable to work. They concluded the study sooner than planned but suffered the consequences of having been framed by Eduardo Bolsonaro and government supporters as anti-national:

> Most of the profiles that attacked us were not fake, they belong to real people. [I received messages like]: 'are you the doctor who killed those people?', 'we are watching you, we are watching your family', 'we are going to make you and your family suffer like you did the patients' family suffer'. It was really hard because we didn't know if this would lead to physical aggression. I talked to the police and to some lawyers . . . I was escorted by a policeman hired by the Amazonas government for two weeks. . . . my wife panicked because I have three small kids, so we never know if something will happen, if this hate will become tangible. Up until a month ago, when the media talked about this subject again, some people from other states sent me messages saying I was going to burn in hell, calling me a 'communist son of a bitch', saying 'you're going to end up like Marielle[13]' – *these people have no idea how to do research, they don't know me, they don't know anything*. We reported it to the police department that investigates cybercrime and to the state's Public Prosecutor's Office, but *nothing has been done so far*. [Speaker's emphasis]

Psychological damage from defamation, threat of violence and legal harassment is not easy to quantify or overcome. Several interviewees detailed the efforts that they go to in order to reassure friends and families, even while suffering panic attacks and depression themselves. Other repercussions included the destruction of the reputations of team members and repeated, targeted harassment by law enforcement and judiciary loyal to Bolsonaro.

> Pro-Bolsonaro prosecutor Bento Gonçalves[14] started investigating us; they published it on social media before I was even given the inquiry. They did it publicly in order to harm us. They summoned the researchers, we had to hire a lawyer and answer many questions. . . . Bolsonaro said that more people were dying in Manaus because we *had a protocol to use high dosages of chloroquine*. He wanted to convince everyone that the high dosages were not used only in the 40 patients we were studying, but in every patient in Manaus, that the deaths were caused by our study. He said it and hoped it would stick. He also said 'according to what I was told, the doctor is affiliated to PT, but I'm not going to comment on that'. It wasn't only his son's tweet, Bolsonaro said it. . . . The president's supporters started to report us everywhere; it is a kind of repression that is extremely organised. The prosecutors from Bento Gonçalves published their inquiry online, so another prosecutor from Goiania copied the same text and opened another investigation. Then a Congressman from Amazonas reported us to the Federal Council of Medicine. These investigations are endless.

The conjunction of political and legal harassment and online hate – the use of the state apparatus to support the misinformation and hate-speech against the teams of doctors and medical researchers attempting to win the fight against Covid-19 – was potent both as a warning to other scientists and in silencing critique of the government's laissez faire policies and reliance on conspiracy theories. While this case is somewhat unique, similar patterns of harassment are visible in India in the case of doctors such as Kafeel Khan and journalists or clinicians who question establishment lies or blow the whistle in the medical field.

What is to be done?

As in India and the UK to which we turn in Chapters 4 and 5 respectively, Brazilians' strategies for dealing with the affective burdens placed on individuals and families targeted by hate varied significantly from trying to respond with love to hate, trying to respond reasonably to hate, getting one's supporters to report or respond to hateful comments, trying to ignore hateful comments, systematic troll blocking, complaining to platforms, leaving platforms, complaining to the police, leaving jobs, quitting blogs, and disengaging from platforms because they are unsafe. In some cases our respondents who could manage to organise and afford this had been forced into exile to protect themselves and their families. Almost everyone we interviewed in Brazil who could afford counselling and therapy had been forced to avail of it in the aftermath of particularly heinous incidents

of online harassment and incitement against them, their families and their communities. While the online vitriol and abuse appeared to be a trigger for increasing levels of anxiety and depression, most expressed the view that their group identities or personal identities were almost always under threat or stress due to ongoing political circumstances regardless of their online presence. Going offline or leaving a job were simply stop-gap measures aimed at shielding their loved ones or selves from immediate violence.

All of these strategies are also clearly mediated by people's theories about why hate happens, its linkage to histories of violent discrimination and the clear and present danger to our interviewees and their families as well as the responses of platforms and law enforcement to their predicament. Acknowledging that in the medium and long-term there is a need for profound social change akin to a revolution in attitudes to gender, sexuality, race, religion, caste, class, disability and other protected characteristics, we engaged deeply with all our interviewees about ways forward in the short-term. Hate costs health, mental health, livelihoods and lives. In response we heard time and again about the frustrations incumbent on being good citizens and following policies and procedures to report hate:

> I reported a tweet for using my photo three months ago. One month ago, Twitter replied asking for my ID to prove who I was. I called Twitter's press office to ask if this was serious or if it was phishing. It was serious. Yesterday, they blocked only that one tweet. The person who posted was angry with my organisation because they fact-checked something and he thought I was responsible for it, so he posted a doctored photo of me . . . that took Twitter three months to act. It's necessary to improve the system, being very careful not to be partial, since it's not always easy to define what hate is. This tweet wasn't blocked for hate, but because they used my photo. Social media platforms have to be more agile in acting on hate speech, they need to hire more people to work with this goal, they need better definitions and criteria. In second place, this subject needs to be discussed in schools. I am in my 40s, my generation is gone! My daughter is 11, we need to start talking about hate speech in the classroom.

Education, and media education in particular, was brought up several times as a medium-term solution. However, it was also seen to have its limits, and to be curtailed by existing regimes of discrimination and domination which require quite different types of action:

> We should educate people, but it's important to have a system that will punish perpetrators. If you type 'Hitler' on Facebook, you will find

> many pages worshiping Hitler and genocide. It's not subtle. This is a crime, these people need to be punished. Education won't solve the problem concerning those people who are already using social media to spread hate speech. . . . We have a project to build a Stereotype Guide: the students are writing this guide to send to journalists in order to encourage them to stop reproducing stereotypes. We have a group that is finishing the 2019 Hate Crime Report. In the UK, we're going to work with students . . . on a campaign against fascism, explaining to people fascism's characteristics, and also a campaign to stop spreading hate. In Latin America, we are working on a report about femicide.

Several of our interviewees expressed deep frustration, hopelessness and overwhelm. Mariana was suspicious of new laws to curb online hate speech, arguing that they would be utilised in corrupt ways to curb the free speech of human rights defenders and critics of authoritarian behaviour. Weary as she was of being forced into hiding her identity, being threatened with violence, having images of her sexuality and body constantly used in demeaning and dehumanising ways, and having her child threatened, she was sceptical of change while the ones wielding the power over courts and laws remained unchanged.

Yet others whom we interviewed had only recently faced hate in this form, since their lives had otherwise been protected by unexamined privileges of race, gender and class.

> Before, I didn't believe in hate speech and in the existence of an "Office of Hate".[15] I always thought it was a bit of left-wing fantasy. But it is in fact really organised, I experienced it. The impact is huge, people still haven't realised it. They understand what this means when they become the victims. Disinformation makes it even harder for people to understand scientific research. . . . I'm very pessimistic about the future. Hate speech is here to stay. I can't see a way to neutralise it. . . . On WhatsApp, there is an excess of freedom, of sharing, nobody controls it. When you share content, you don't know exactly what it is and where it is going to go. You're a cog in a much larger process, someone is overlooking it. I'm going to be honest, I'm afraid hate speech will transform our society in a bad way, make it worse.

Others told us that their solidarity overrode potential threats:

> My children – blood related and of saint – are really afraid something will happen to me, especially now. They are really afraid. . . . We reported it to Decrin,[16] in Brasília. We don't have a special police department dedicated to these crimes where I live; one of my projects is

> to implement policies that protect the *terreiros* in this area . . . we don't know if our case will have any result. We're thinking about suing this person in a civil procedure as well. The attack was public, it was made in a comment in one of my posts on Facebook and everybody saw it. Yesterday, someone drew a swastika on my wall. [But. . .] I can't stop. I won't be silenced, I can't be indifferent to what is happening. My indignation is stronger than my fear.

Conclusion

Throughout this chapter we noted the insistent relationship between the past and the present, between seething prejudice and structural discrimination, between sanctioned disinformation and the rise in hateful trolling. Histories of colonial repression, imperialist aggression, media manipulation and religious alignment with European and US-based churches all converge in the accounts given by our key informants and interviewees of racist, anti-Indigenous, anti-leftist, homophobic and misogynist bullying, harassment, trolling and aggression. On the streets and online, in grocery stores, churches, sidewalks, newsrooms, chatrooms, message boards, messaging apps and on platforms, dehumanising language and imagery, and symbolic representations of lynching and death, run in parallel with each other. Common denominators mentioned by everyone were the legitimacy lent to hate by powerful politicians and religious leaders, the complicity or co-optation of mainstream media, and the profit-seeking complicity and inadequate response of corporate platforms. The common political theme of this stark spectacle of hate was a push to silence and make invisible people from Black and Indigenous, LGBTQIA+, feminist and leftist communities, in order to continue to dominate and oppress them, and to profit from their suppression. Their refusal to be silenced, their determination to continue to stand for justice, and to be open about their identities in the face of structural injustice, discrimination, pain, anxiety, family pressure, job loss and other dangers was the single most inspiring finding of our research. As will be seen in Chapters 4 and 5 on India and the UK respectively, this courage and determination is something that sustains not just community organisers and activists but also entire communities in the face of hate.

Notes

1 https://theconversation.com/how-jair-bolsonaro-used-fake-news-to-win-power-109343
2 www.theguardian.com/commentisfree/2018/oct/06/homophobic-mismogynist-racist-brazil-jair-bolsonaro

3 https://brazilian.report/power/2020/10/24/against-vaccine-bolsonaro-son-reactivates-office-of-hate/ and www.zdnet.com/article/fake-news-probe-in-brazil-exposes-office-of-hate-within-government/
4 https://advox.globalvoices.org/2018/10/28/brazilian-journalists-face-hacking-doxxing-and-other-threats-as-election-draws-near/
5 Left-wing politicians. Manuela D'Ávila was the vice president candidate on Fernando Haddad's ticket.
6 Worker's Party – Partido dos Trabalhadores (PT).
7 She was the vice presidential candidate alongside Fernando Haddad and a member of the Brazilian Communist Party who also ran to be Mayor of Porto Alegre.
8 https://theconversation.com/in-brazil-religious-gang-leaders-say-theyre-waging-a-holy-war-86097
9 A pejorative way of referring to Afro-Brazilian religious followers. It's similar to calling someone a 'witch'.
10 https://observers.france24.com/en/20190211-brazil-indigenous-youtube-combat-racism-12
11 An indigenous weapon.
12 www.sciencemag.org/news/2020/06/it-s-nightmare-how-brazilian-scientists-became-ensnared-chloroquine-politics
13 Marielle Franco, Rio's city councillor who was murdered in 2018.
14 A town in the state of Rio Grande do Sul, in the South of Brazil, more than 4,000 kilometres away from Manaus.
15 'This toxic environment has been fomented by what Brazilians call the "office of hate," an operation run by advisers to the president, who support a network of pro-Bolsonaro blogs and social media accounts that spread fake news and attack journalists, politicians, artists and media outlets that are critical of the president. The office of hate does not have an official title or budget – but its work is subsidized with taxpayer money.' – www.nytimes.com/2020/08/04/opinion/bolsonaro-office-of-hate-brazil.html
16 Special police department that investigates crimes motivated by discrimination against race, religion, sexual orientation, and against elderly or disabled people.

References

Anderson, R. (1996). The Quilombo of Palmares: A new overview of a Maroon state in seventeenth-century Brazil. *Journal of Latin American Studies*, *28*(3), 545–566.

Anderson, W., Roque, R., & Ventura Santos, Ricardo. (2019). *Luso-tropicalism and its discontents: The making and unmaking of racial exceptionalism*. Berghahn Books.

Bailey, M. (2021). *Misogynoir transformed: Black women's digital resistance*. New York University Press.

Banaji, S. (2018). Vigilante Publics: Orientalism, Modernity and Hindutva Fascism in India, *Javnost - The Public, 25*(4), 333–350.

Braudel, F. (1981). *Civilization and capitalism*, *15th–18th century*. Fontana Press.

Chirio, M. (Trans.). (2018). *Politics in uniform: Military officers and dictatorship in Brazil, 1960–80*. University of Pittsburgh Press.

de Albuquerque, A. (2017). Protecting democracy or conspiring against it? Media and politics in Latin America: A glimpse from Brazil. *Journalism*, *20*(7), 906–923.

de Vasconcelos, A. V. S. (2019). *The persistence of racism and its impact on the Afro-Brazilian culture*. Researchgate.

Ferretti, F. (2019). Decolonizing the Northeast: Brazilian subalterns, Non-European heritages, and radical geography in Pernambuco. *Annals of the American Association of Geographers*, *109*(5), 1632–1650.

Keisha-Khan, Y. P. (2004). The roots of black resistance: Race, gender and the struggle for urban land rights in Salvador, Bahia, Brazil. *Social Identities*, *10*(6), 811–831.

Matos, C. (2011). Media and democracy in Brazil. *Westminster Papers in Communication and Culture*, *8*(1), 178–196.

Oxfam. (2017). Brazil: extreme inequality in numbers. [online] Oxfam International. Available at: https://www.oxfam.org/en/even-it-brazil/brazil-extreme-inequality-numbers.

Puntoni, P. (2019). "The Barbarians war": Colonization and indigenous resistance in Brazil (1650–1720). In N. Domingos, M. Jerónimo, & R. Roque (Eds.), *Resistance and colonialism: Cambridge imperial and post-colonial studies series*. Palgrave Macmillan.

Schneider, N. (2011). Breaking the "silence" of the military regime: New politics of memory in Brazil. *Bulletin of Latin American Research*, *30*(2), 198–212.

Sethi, A. (Ed.). (2018). *American hate: Survivors speak out*. The New Press.

Sheriff, R. E. (2001). *Dreaming equality: Color, race, and racism in urban Brazil*. Rutgers University Press.

Snider, C. M. (2018). The perfection of democracy cannot dispense with dealing with the past: Dictatorship, memory, and the politics of the present in Brazil. *The Latin Americanist*, *62*(1), 55–79.

Sponholz, L., & Christofoletti, R. (2019). From preachers to comedians: Ideal types of hate speakers in Brazil. *Global Media and Communication*, *15*(1), 67–84.

4 Social media, violence and hierarchies of hate in India

Introduction: The politics of the digital sphere

As in Brazil, many of our informants in India evinced a shared cynicism about the chances of imminent political or policy change that would alter the volume, types and reasons for targeted hate and violence. The subset of Indian social media users whom we interviewed theorise online hate and discrimination in an impassioned manner. Like us, they too link it to the histories of postcolonial religious pogroms, poverty, caste discrimination and gender-based violence. Addressing online hate was, for them, just one facet of a wider need for action towards economic and social justice. Since 2014, under the aegis of the Bhratiya Janata Party (BJP), a vast swathe of the Hindu population have weaponised their religious identity (intersecting caste and gender) against overlapping minoritised groups – Muslims, Dalits, Adivasis, Christians, queer and trans groups – with different forms of slurs, atrocity and violence directed at men and women in these communities. This weaponisation manifests offline and online in both physical and symbolic forms.

In India, where mass digitisation has taken place through a series of forced bureaucratic measures (Bhat, 2020; Ferguson & Gibson, 2015), these interconnected domains act on each other in increasingly complex ways, some of which are addressed in this chapter. By dint of subverting democratic processes in all but name, the BJP has managed to achieve electoral and ideological hegemony unlike their political predecessors. The BJP's politics is comprised of strands familiar to those studying authoritarian populist governments across the world – vicious privatisation and neoliberal reform with capital flowing freely between top corporate and political interests while the populace becomes ever more destitute. This is accompanied by the discursive championing of a faux-emasculated nation-state as the moral horizon set up to justify and underpin all political action, cultural values, public space and so on – a paradoxical formation of, as Tambiah (1986) puts

DOI: 10.4324/9781003083078-4

it, 'a majority with a minority complex'; the concomitant diffusion of hate against minoritised groups into the minutiae of daily life; widespread suspicion and paranoia about the Other (as the cause for all real and imagined ills of society); and deep-seated patriarchy and misogynist control of women's bodies and youth relationships at peril of death.

An aspect of the BJP's dominance is its effective capture of public institutions, including universities, public health bodies, the electoral commission, banking and financial institutions, legal and regulatory institutions and media. Captured through political appointments, co-optation, bribery, threat, steamrolling legislation through parliament without discussion and other strategies, these are used to discriminate against, surveille and harass opponents and dissenters. There is also subservience to and complicity in the BJP's politics from the private sector – particularly from the largest corporations. Increasingly, global social media companies have recently acquired stakes in or are working closely with dominant domestic multinationals such as the Ambani group, the Adani group, the Tata group and so on.

A further aspect of the BJP's dominance is the personal 'brand' of Narendra Modi. During his previous stint as BJP chief minister of Gujarat, he was accused of complicity in the 2002 Gujarat pogrom in which more than 2000 Muslims were killed by Hindu vigilante groups (Ayyub, 2016). Once banned throughout Europe and North America and refused visas, with the help of global image management firms, he has steadily consolidated his political image through a mix of neoliberal politics (acquiring agricultural land and providing it to large industrial investors along with heavy subsidies) and a hardline Hindu nationalism that has constantly vilified Muslims (Jaffrelot, 2007). After his entry into national politics, Modi rebranded himself as a champion of national security, development and technologised, corruption-free governance (Chakravartty & Roy, 2015). Simultaneously, his party members unabashedly used disinformation to vilify Muslims. They promoted a culture of Hindu vigilantism that led to the public lynching of Muslim boys and men, and the rape and murder of Dalit and Muslim women and girls, the beating and murder of Dalit men and multiple hate crimes. Modi was an early adopter and is an active user of social media. His supporters and his party use social media heavily for propaganda and to shut down criticism. They employ trolling, abusive speech, doxing, hacking and other strategies that we will expand on as we elaborate our interviewee's experiences.

Technological networks are embedded in socio-political, cultural and economic contexts, but also, to a large extent, reproduce and strengthen the abiding tendences of the contexts into which they are embedded. In the late 2000s, a series of transformative changes precipitated the widespread use of

social media platforms and cross-platform applications, primarily accessed via smartphone. Since 2015, Indians have enthusiastically taken to Google (YouTube) and Facebook (Facebook, WhatsApp and Instagram), and, to a lesser extent, to Twitter, Instagram and Tik-Tok (now banned in India). The precise number of users is uncertain since financial motivations can inflate or exaggerate the extent of social media and internet usage. While social media usage has particularities (for instance, its limitation to at most 40% of the population, and concentration amongst males and in urban areas) that help us distinguish it from other aspects of daily life, we argue that this distinction rests on a thin and blurred line between the online and the offline.

Founded in 1980, the BJP's ideological roots are in the Rashtriya Swayamsevak Sangh (RSS) which was founded in 1925. In less than a hundred years, the RSS has gone from a small cultural chauvinist initiative beginning with fewer than 20 members into a mass fascist organisation with a membership above five million. It is no coincidence that the core ideas of the RSS draw deeply from the ideologies of fascism and Nazism that were flourishing in 1930s Europe. The RSS has also diversified its organisational structure to target different communities. For instance, there is a separate wing for girls and women, a separate organisation for students, for youth members and so on. Collectively these organisations are called the Sangh Parivar or the Sangh Family (Andersen & Damle, 1987; Hansen, 1999). The BJP is the political face of the Sangh Parivar. Narendra Modi was cultivated and promoted by the RSS and most of the BJP's senior leaders come from the RSS. This background is important to keep in mind when we refer to the BJP or to Modi.

A long history of inequality

For more than 2000 years, much of what is now South Asia has been subject to the caste system which, far from remaining static, has changed in multiple ways in response to historical events and movements. The caste system comprises four broad categories of castes, arranged in vertical hierarchy, with Brahmins (priests) at the top, followed by Kshatriyas (warriors), Vaishyas (traders) and Shudras (manual and menial labour). These *savarna* groups imagine themselves in distinction to the fifth and outside caste *avarna* group, commonly known as Dalits. Located at the bottom and outside the structure, Dalits alongside Adivasis (Indigenous peoples) have suffered continual discrimination, exclusion and violence. Online too, in what Shanmugavelan[1] has called 'caste-hate speech', Dalits face extensive abuse and discrimination.

Some aspects of the caste system invite comparison with oppressive systems in other parts of the world. The most common comparison is race in

the United States (c.f. Wilkerson, 2020). While we have provided an abstract and condensed picture of the caste system, in practice it is far more complex. Each caste category is in turn comprised of various sub-castes or *jaatis*, that are also constantly competing for superiority over various other *jaatis* in their caste group. Conflict in the caste system is therefore not strictly vertical but also to some extent, horizontal (Manor, 2010). The graded inequality of the system also distributes power in a peculiar way – since a group is incentivised to ally with those above it, oppress those below it, and negotiate with those who are on par, even if only to maintain the status quo (Gorringe and Rafanell, 2007). The distinctive aspect of the caste system is that it, unlike other oppressive systems, links occupations to birth and is considered permanent. In other words, there is technically *no control over an individual's entry into a caste* and there is *little way out of the caste you are born into* – apart from the gradual change introduced through *jaati* politics, inter-caste marriage and religious conversion. The concept of inter-caste marriage is especially significant since the caste system is essentially a system of 'exogamy superimposed on endogamy' maintained through a control over women's sexuality (Ambedkar, 1968[1917]). The caste system is therefore inherently a patriarchal system and patriarchy profoundly informs how gender identities are constructed, and refracted, through caste. Much of the current online discrimination, bullying and hate cannot be separated from misogyny and the inability to accept any deviation from a dominant caste masculinity that pervades Indian and wider South-Asian online culture. A caste perspective arguably enriches the complex intersectionality that underpins these contemporary practices of gender struggle and identification.

This does not mean that all conflict, discrimination and violence in India can be analysed exclusively through a caste perspective. Caste should be seen as what Bourdieu calls a 'structuring structure' (1984, p. 170), a generative space within which we act but do not always or necessarily experience as a restriction, but rather develop a 'feel for the game' (Bourdieu, 1990, p. 66). For example, we do not maintain that the identities of all Adivasi groups, or Indian Muslims are reducible to caste. Some Muslim sects and groups – such as Sunnis, Shias, Ahmadis, Boras, Sufis and so on – impose their own sets of divisions and rules. However, the caste perspective provides a way of deepening our understanding of the distribution of power when it comes to how Muslims in India are treated. Those who are lynched or face physical attacks, such as the Delhi pogrom of late February 2020, are more likely to be Pasmanda Muslims; whereas Muslim political dissenters who are targeted through draconian legislation are more likely to be from dominant castes. At the same time, the constant reproduction of anti-Muslimness through governmental practices, academic research, media discourse, landmark events (such as the demolition of Babri Masjid or the Shah Bano

case) produces a clear sense of Muslim identity that can sometimes conceal the intersection of Muslim-ness with caste (c.f. Ansari, 2009). Working with a caste perspective necessitates an intersectional approach (c.f. Arya & Rathore, 2020) which can help us situate contemporary problems.

Socio-political contexts and the emergence of online and social media usage

As in Latin America, the mid-1980s to late-1990s in India were years where neoliberal structural adjustment policies were forced on successive Indian Governments by Bretton Woods institutions in order to mitigate global debt crises sparked by western policies. One of the key objectives of these policies was to 'open up' public sectors to the 'free market' and western direct investment. Governed under the colonial 1885 India Telegraph Act, media and communications had been tightly controlled by the central government. After the 1990s, one of the first industries opened to private investment was telecommunications. Telecommunications was seen as a luxury service (compared to more 'basic needs') by previous governments. Other socio-political shifts took place in tandem with foreign direct investment.

After the assassination of Congress Party leader Indira Gandhi in 1984, her son and successor Rajiv Gandhi essayed a series of symbolic actions pandering to the sentiments of chauvinist Hindus (and in particular, dominant castes) in order to consolidate the fallout from his mother's regime. A crucial symbolic act was the regular broadcast of the Hindu epics *Ramayana* and *Mahabharatha* by the sole public broadcaster in India – Doordarshan (Mankekar, 1999; Rajagopal, 2001). In 1990 however, the coalition government headed by prime minister V.P. Singh announced that it intended to implement a 27% reservation of central government and public sector jobs for Other Backward Classes (OBCs), the largest population bloc (approximately 52–54%). This announcement about affirmative action signalled a big boost for parts of the population that had hitherto been systematically excluded from access to education and employment and other benefits deriving from these fundamental aspects of citizenship. This move resulted in a major recalibration of electoral politics at the national level. The BJP, led by their then leader L.K. Advani, mobilised rightwing support through a deployment of the so-called forward or upper castes in a vicious and spectacularly orchestrated mosque demolition and temple building project that left a trail of violence and death in its wake (Teltumbde, 2005). In this journey, which utilised multiple forms of media to communicate its message of exclusion and provocation against Muslims, we can see the antecedents of Modi's contemporary Hindutva fascist government (Sarkar, 1993; Banaji, 2018).

The BJP's strategy for the unification of a caste-fragmented Hindu population (wherein upper castes had no desire to give up privileges and inherited networks of position), was through the othering of Muslims. In a stunning feat of modern propaganda and disinformation, the BJP, aided by the RSS's fast-growing infrastructure of fascist training schools and programmes, reaped the benefits of this fictional construction of Hindu identity. Capitalising on imaginary wounds inflicted against upper caste Hindu masculinity by past (and present) Muslim 'others' and by affirmative action caste policies (known as reservations), the BJP went from two seats in 1984, to 120 seats in 1991. By 1999, the BJP led a rightwing coalition (National Democratic Alliance). In the early 2000s, the RSS leader Narendra Modi was elected Chief Minister of Gujarat and accused of complicity in the 2002 pogrom where more than 2000 Muslims and several Muslim-aligned Hindus were killed in a horrific manner by Hindu fascist mobs (Sarkar, 2002; Ghassem-Fachandi, 2012). Since then, Modi has consolidated his political career as a figurehead of masculinist Hindutva. His neoliberal subsidies for major corporations and destruction of welfare schemes endeared him to elites, who in turn helped him rebrand as a messiah of development.

Alongside Modi's rise to power, there was a decisive shift in the telecommunications industry. The 1990s saw poorly regulated privatisation where the government struggled to arbitrate disputes between private players or between private players and the government's own public telecommunications service providers (BSNL and MTNL). Poorly managed privatisation led to organisational change. For example, a fixed license fee regime shifted to a revenue sharing agreement between private telecommunications service providers and government (Athreya, 1996; Chowdary, 1999; Roy, 2009; Sridhar, 2012); meanwhile a rising middle class created the material conditions for a booming telecommunications market. Governments too would stand to earn substantial income from spectrum charges, tax on value added services and so on. The 2000s saw a crowding of the market with more than ten (domestic and international) telecommunications service providers. Heavily skewed towards urban areas, most service providers were in long-term debt to license spectrum at high costs – either expecting consistent long-term growth or expecting to re-sell assets to competitors at a profit.

Voted out in 2004, the far right BJP came to power again in 2014 after an unprecedented use of social media during the campaign with Modi at the helm. By this point, the telecommunications market in India was crumbling as the online ecosystem tilted towards usage related to products from companies such as Google, Amazon, Facebook and Microsoft. Supported by fascist cadres, the Modi regime has seen a marked increase in a Jim-crow style politics involving frenzied mobs who lynched, tortured, harassed and intimidated religious minorities (Banaji, 2018). Although several tracker websites and

initiatives have been attempted in order to provide an aggregated picture of the atrocities and violence under Modi, the government has repeatedly shut these down.[2] We estimate that there are hundreds of such instances of mob violence connected to 'defending' or advocating Hindu nationalism (Hindutva).

Social media platforms and cross-platform applications have played a role since 2014 in enabling the formation of vigilante mobs and allowing the fascists to upload and spread their propaganda. Aside from these incidents, the BJP has constructed a fictional trope of India as an ascendant Hindu nation state vulnerable to the conspiracy of Islamic and communist-backed terrorism. This fascist rhetoric has been amplified by a subservient mainstream media and used as a cover to target both real dissidents and imagined opponents – including Dalit and Adivasi activists, religious minorities, university students, human rights activists, academics, artists, journalists and many others. Many have been jailed on trumped up charges.

In 2016, one of the richest men in the world, Mukesh Ambani owner of Reliance Industries Group, launched Reliance Jio Infocomm Ltd. From its inception, the group enjoyed a competitive advantage as it was allowed by TRAI (Telecom Regulatory Authority of India) to offer free subscriptions, free internet usage and free mobile phone devices for nearly a year.[3] Such corrupt subsidies enabled Jio to 'steal' millions of customers from competitors who could not afford equally low tariffs. In 2020, Reliance Jio Infocomm became a subsidiary of Jio Platforms, which raised more than 20 billion US dollars by divesting a nearly 33% stake. Companies with stakes in Jio Platforms include Facebook (9.9% stake for 6 billion US dollars) and Google (7.7% stake for 4.7 billion US dollars).[4] According to a report by researchers at Azim Premji University,[5] Mukesh Ambani, a key beneficiary of the Modi government in multiple sectors from energy to telecom, and a few others such as the Adani group, have increased their wealth many times over, while the Covid-19 pandemic has pushed 230 million Indians into poverty. This brief account of the macrolevel structural matrices within which social media companies operate in the Indian economy sets the scene for an understanding of the role of social media in circulating and promoting hateful content.

Modi and the media

The distinctive relationship between Modi's regime and the media is the absolute control that Modi appears to wield over his own image. Unlike all previous Prime Ministers, Modi controls his media appearances by eliminating any possibilities for spontaneous interaction with journalists (there are no unscheduled interviews or press conferences). He uses the public broadcaster to voice his rhetoric, speaks only to journalists who will report

favourably, and actively uses social media platforms to get his versions of events across to millions of followers. Amit Shah, the Home Minister and Modi's key aide since his days as Chief Minister of Gujarat, has also made effective use of Facebook in the 2014 and WhatsApp in the 2019 elections to mobilise BJP voters. Multiple IT cells and hundreds of thousands of BJP-linked social media groups circulate content to keep the BJP's ideological vision in supporters' minds (Thakurta & Sam, 2019). Many (often, although not exclusively, men and from dominant castes) have become vocal supporters of Modi on the internet and brook no criticism of him or his government. These online Modi supporters are notorious for bullying, abusive speech, trolling, doxing and spreading disinformation. There is sufficient evidence to indicate that social media platforms and apps are regularly 'gamed' by the BJP IT cell to make topics trend or go viral, manipulating opinion through coordinated behaviour.[6] Legacy media (newspapers and television news) then report the 'buzz' uncritically, selectively favouring the BJP. Bollywood, the Hindi film industry, with a massive viewership and disproportionate cultural and symbolic capital at its disposal, has also promoted Modi and Hindutva[7] both directly through Islamophobic and anti-communist propaganda films and via Hindutva-supporting actors and social media 'influencers'.

In line with other tactics of suppression, the Modi government is trying to control social media use and the use of other digital media and communication networks (such as streaming platforms). The government has introduced new rules for intermediaries that establish procedures for addressing complaints from users.[8] However, the government has claimed for itself the power to take final decisions on complaints that cannot be resolved through internal mechanisms. Further, the central government has pressured companies to break the encryption of peer-to-peer messaging applications such as WhatsApp, allegedly to track the originators of disinformation, but in actuality to surveille and incarcerate opponents of the regime. Authoritarian restrictions on NGOs and digital news and current affairs initiatives have had a 'chilling effect', while enabling government supporters to dominate social media. In 2021, social media companies, especially Facebook and Twitter, have been under increasing pressure from the Modi regime to shut down even mildly critical accounts, delete posts, and stop tagging government supporters' disinformation posts as manipulated and so on. As we write, it appears likely that the courts will ask social media companies to comply with government demands, thus further endangering social media users who have identities targeted by Hindutva fascist mobs, are critical of the regime or do social justice work. The situation, as our key informants explain, is already one of deep anxiety and existential threat for many communities.

Difference and discrimination

Meena, a queer activist who has been targeted by the state for their political views, explained their experiences with social media in terms of how difference of identity and social position plays out, both with them and other progressive people they follow on social media:

> A lot of it was targeting my appearance, just saying that I'm ambiguously gendered, essentially. Sort of saying that I don't have a sense of morals or ethics because I went out and created problems for people. It was mostly attacking my general queerness and the fact that I'm assigned female at birth. . . . Something recently happened to an artist I follow on Instagram, where she parodied another Instagram account and the person who she parodied got really upset. That person (an Upper Caste Hindu), told her followers that she's being bullied for being Hindu but she wasn't being bullied . . . Her entire follower(base) turned into this troll army and would send the girl who posted the parody so many hate messages. . . . They went after her for being fat, having darker skin, being Dalit, being queer – every aspect of her identity, and said absolutely horrible things. It was really painful to watch.

Preeti, an activist working in a rural part of a North Indian state, shares how quickly an ideological argument can shift to caste-based abuse:

> A few days back there was news that the Indian government is selling a part of LIC. I had posted on Facebook saying that LIC is the backbone of many people in India and if government sells even that it will be a loss of people. Now this is not against the government as such. This is about ideology. There were some Modi Bhakts who jumped on it. First a lady posted saying 'all this is fine but Modi is doing good work'. I didn't respond, this is her opinion. Then a few more people responded saying, 'you hate Modi, what is your status (*aapki aukat kya hai*?), you should remain where you have come from. Caste system should exist, people like you should be suppressed (*pair ke neeche daba ke rakhna chahiye*)'.

These experiences reveal the extent to which the online environment in India is dictated by dominant caste, heterosexist and pro-regime users. A particular set of techniques – trolling, hacking, abuse, doxing and so on – ensure that those who are marked as different online by virtue of skin colour, gender, sex, sexuality, caste, class and so on are silenced and humiliated to the point where they do not venture online again. These experiences also

reveal that much like the offline world, an intersection of caste, gender and sexuality or class and gender heightens probabilities of harassment and discrimination. Indigeneity, Dalitness, Muslimness, queerness, darker skin, and leftist and/or feminist political opinions compound each other in terms of the consequences that individuals and groups face.

Apart from the 'pure' hostility and abuse towards specific identities (LGBTQIA+, Muslims, Dalits), there is also an immediate torrent of abuse that users – particularly from a vulnerable social identity – face for ideological dissent. In the excerpt above, critiquing the privatisation of a public sector company was the catalyst that unleased caste-hate speech. The ease and extent to which online 'political' discourse (for example, about privatisation of public sector industries) transforms into casteism should challenge illusions about the Internet as a Habermasian public sphere where rational-critical dialogue can be put to use towards deliberative democracy (Anderson & Jaffrelot, 2018; Bürger et al., 1992; Khorana et al., 2014). Historically formed asymmetrical hierarchies dominate the media sphere in India while historical discrimination and violence manifest themselves repeatedly on social media. A particularly violent maintenance of difference can be seen in the abuse targeted against Muslims. Hatred against Muslims has political support from prominent BJP leaders who have not hesitated to use the full power of the state apparatuses at their disposal against Muslims.

Anti-Muslim hate online

One of the most prominent forms of social media discrimination in India is Islamophobia – discrimination, dehumanisation and incitement to violence against Muslims. Several interviewees spoke about this. Syed, a journalist at a mainstream media organisation, and Shruti, who leads an independent digital news organisation explain:

> I get abuse just because of my name and my religion. . . . People try to make it a Hindu-Muslim issue – you can check my page; very few people write against radical Islam the way I do. But at the same time, I write about other things as well. I won't forget how Palestinians are being robbed and killed and how Israel has been treating them for years and it has the backing of the so called developed and modern world . . . Just a few months ago I wrote something about cricket, someone started abusing my name and just the usual abuse they used on social media which I can't even tell you. They say 'go to Pakistan, you are a traitor, you are a mullah, and you are a motherfucker', these types of things. Things that are so abusive that I can't even tell you. . . [Syed]

> I've done reports on 'love jihad' and seen the response of people like . . . I don't even know how to say it . . . I've seen people say, 'the first thing you need to do when you meet a Muslim man is really make sure he's Muslim by taking off his pants' . . . . And just really gross things like 'you should just behead them, hang them by their balls'. . . . I really don't know what to do about it. I think when the Rohingya thing came up and I was reporting on it, I just saw the kind of things people said about them, like calling them 'terrorists', 'dump them in the sea' – *who says things like that*? And they said that 'these guys are terrorists', but there is not one FIR against [the Rohingya] so far like major FIR all the complaints against them have been for pickpocketing or small theft but nothing, nothing to do with national security. . . . I have Muslim friends who are journalists. The minute they write something they are called 'jihadis', all of us are called 'jihadis' and 'terrorists' – if I'm called that twice, that person will have been called that 100 times. I know a lot of Muslim friends who are otherwise very brave journalists, who refrain from writing on the Internet or talking about religion, because it is just so explosive. There was this example of a Muslim journalist from [redacted], things got to such an extent that there were blogs saying that he was a mujahideen member, that he's a terrorist – with absolutely no evidence, just because he's a Muslim you can write things, and there's a percentage of people who will actually believe it simply because of the fact that you're a Muslim. I think the kind of hatred Muslim journalists face is something that I cannot even comprehend. [Shruti]

Any pretensions that social media platforms and cross-platform applications may have had to being an 'alternative public sphere' ring hollow given this context. Silencing comes through Islamophobic abuse directly aimed speakers' identities ('traitor', 'mullah', 'go to Pakistan', 'he's a mujahideen member', 'he's a terrorist' and so on). In most instances, it is not just the intensity of abuse but also the magnitude that can be overwhelming ('called that 100 times') making it impossible for many Muslims to post about anything, even about ostensibly non-political domains such as cricket. Muslimness as an internally heterogenous and complex social identity is discursively reduced to a homogenous religious stereotype. Although these discursive strategies against Muslims may resemble anti-Muslim discourses in other parts of the world, noticeably in the UK and Myanmar, the hatred against Muslims as the 'Other' has a distinctly Indian lineage that has been strengthened by post-9/11 Islamophobia across the globe.

Historically, a positioning of Muslims as being on the 'outside' and 'enemies' has been how the RSS and the BJP have constructed 'Hindu' identity.

This spurious religious identity overwhelms other potential ways of engaging fellow citizens on or off social media, especially when these are critical of the casteism, chauvinist nationalism and neoliberalism that the BJP has cultivated. The ontological equation of Muslimness with threats to national security or Hindu womanhood is now so ingrained that many caste Hindus barely give it a second thought. While such racist imaginaries depend on a binary of good Muslim/bad Muslim so crucial to post-9/11 imperialist geopolitics (Mamdani, 2004), the Indian history of discrimination against Muslims constantly references Mughal rule, a trope through which Hindu majoritarians construct a false history of persecution for which they are seeking redress (Anderson & Jaffrelot, 2018; Truschke, 2020). Discrimination against contemporary Muslims is especially intense when it comes to Muslim women.

Islamophobia and misogyny

The intersection between caste and religious identity is exacerbated by the misogyny and hate targeted at Muslim women which has risen steeply since the early 2000s. In 2018, Amnesty International published a report about misogyny on Twitter, detailing the ways in which women were targeted in western countries such as the UK, USA, Denmark, Italy, New Zealand, Sweden, Spain and Poland. The problem is even more ubiquitous and intense in India. A particularly pernicious instance of doxing (a process where a person's private documents are revealed without that person's consent, in order to silence or harass) targeting Muslim women came to light in mid-2021 when an app and website 'Desi Sulli Deals' was revealed. 'Desi' is a common term for South-Asian and 'Sulli' is a derogatory term used by far-right Hindutva groups to refer to Muslim women. Developed through the open-source software website Github, 'Sulli Deals profiled more than 80 Muslim women, including several outspoken, high profile and dominant caste Muslim women, using images stollen from their social media and public spaces and conducted faux auctions.[9] These 'auctions' provided opportunities for Hindutva trolls and sock-puppet accounts to post vicious misogynist and anti-Muslim comments objectifying the women, reducing them to their body parts and threatening rape. Although rights activists forced the Delhi police cybercell to file a complaint, the issue was under-reported and political leaders kept silent. It was only after several Muslim women complained vocally that Github and various social media companies including Twitter and YouTube shut down the application, the website and the accounts promoting it.

Other violence goes beyond the virtual silencing of outspoken women. Women in semi-urban and rural areas told us that their daughters' profile pictures are circulated accompanied by lewd and abusive comments, with

'upskirt videos' of young teenagers and women frequently circulated without their knowledge in closed all-male WhatsApp groups. In mixed gender WhatsApp groups, women sometimes encountered hardcore pornography shared by men, supposedly 'by mistake'. In other places, for example in the state of Uttar Pradesh, incidents of rape and gang rape are regularly filmed by perpetrators and then sold for a nominal cost at local mobile phone shops, transferred to clients' SD cards and further shared.[10] During our research we were told of at least one public health worker who committed suicide. A sexual health activist, Preeti, who works with young women in rural North India, and is active on social media explained:

> . . . if you are online at 1am in the night, suddenly you will get a call and obscene language is used. They say very obscene things like- 'you are online at this hour, are you sexually not satisfied?' Some ask, 'why are you still single? If you need some, just call me, I will come there.' If you posted something about gender equality on Facebook, people write so obscenely, they start commenting on your private parts like vagina, breasts, on your weight . . . There are about 4–5 girls with me who said they don't want to do this work as people send obscene content on [Facebook] messenger. Those girls went into depression, we had to put them on medication. Some children come from small families, and they struggle and reach here. For them these kinds of things are very new. They are not able to handle it.
>
> Recently, we were working on the Hathras case, there was a committee made. In this regard we had posted on Facebook – in that case boys from Jat community were the perpetrators – we posted about the case. The first comment that came was, 'we will rape you, we will do what was done to Manisha'. These comments were not even in the inbox, they would write it directly on our Facebook post. They would write, 'we will rape you, will throw acid on you, you try and step out of your house alone!' It is a lot of mental harassment. People who make these comments, they all have fake accounts. Nobody used their personal account to write all this. They will be in my friend list from their personal account and will read what I have posted but will post comments from a fake account. They keep themselves safe. Though from their style of writing one can understand who it is actually.

Feminist non-profit spaces, which might have reported physical instances of intimidation and abuse through legal protocols, struggle to deal with anonymously delivered abuse. Apps such as Facebook Messenger lead women to experience this abuse as a deeply invasive phenomenon that must be suffered in private, and internalised individually, resulting in fear, depression and loss of

motivation. This is not restricted to women in rural areas. A prominent Kashmiri ex-student leader from Jawaharlal Nehru University in Delhi, Shehla Rashid, spoke about the massive levels of abuse she has faced from men in all communities, in response to her work as an outspoken woman, Muslim, Kashmiri, and progressive activist from university with a history of leftwing activism.

> First of all these are not very innovative . . . rape threats, and you know, 'whore' and 'prostitute'. . . . So you have degrees in it. You can have rape threats, you can grade it by prostitute or a worse word. The second kind is misogynistic content. Now that's more difficult to identify. Through keywords you can look at rape, whore, slut, etc., you can still do a keyword search. How do you do a keyword search for misogyny? Now only qualitative analysis can tell you that. It is very difficult to pinpoint that but its very very widespread and across the political spectrum. So is abuse, don't get me wrong. . . . If I were to comment on an issue which is sensitive to Muslims in general, despite me being a Muslim, despite me having a claim to speaking for Muslims myself, or as a Muslim, I will get abused by Muslim men and in pretty much the same terms.

As Rashid says, it is difficult to use quantitative methods to ascertain instances of misogyny which go beyond expected keywords. Misogyny in the Indian social media space, like casteist communication and Islamophobia, can either be open and easily recognisable or highly contextual, culturally and linguistically nuanced, presenting a challenge for social media teams that emphasise Artificial Intelligence approaches to identify and take down discriminatory content.

Ever more disturbing physical and symbolic violence against women circulates on WhatsApp. Padmini, a postgraduate student from a well-known academic institution in Maharashtra recounts her experience:

> . . . I also talked about how the army is masculinist and assists in state sponsored crimes. These are words that we use in academia very frequently. Someone said that a WhatsApp group wasn't an appropriate place for this . . . But screenshots [of my posts] were circulating all around campus. And then those were posted on Twitter by a student who very proudly claims to be an office bearer of the ABVP [the student wing of the BJP]. He later pulled down those posts. He blurred everyone's numbers but mine. This post was getting shared by a BJP MLA and another BJP spokesperson. It was flooding all over twitter. Then I started getting calls that I didn't answer. Later I started getting messages on Facebook, so I deactivated my account. But before that

> someone had got a photo of me and my mother and circulated that on Twitter. My number and my email id, my Facebook, Twitter account was made public. The institute was also getting a lot of political pressure from outside to act, so they set up a committee which conducted a hearing. They asked me to give an apology and I didn't want to make this into an issue, so I gave a written apology. The committee gave me notice saying that I must vacate the hostel and that they are letting me finish my PhD but if something like this happens again, I would be expelled. They also said that they won't consider me for further employment or study in the institute.

The most startling thing about this account to those who have not followed the rise and spread of fascism in India may well be the university administration's abject refusal to support their victimised student and their positioning of themselves alongside her oppressors and doxers. Her comments about the abuse meted out by the Indian army and the 'masculinist' nature of the army space have ample evidence in human rights reports over the years. However, since the BJP's rise to power, critiquing nationalism is dangerous, and any critique of the armed forces is deemed unacceptable. Voicing such opinions as a woman multiplies the danger. The excerpt above clearly reveals the ways in which the BJP, through its various licit and illicit networks, create offline consequences for online utterances.

Complex psychic and social consequences

While justice and rights activists have been targeted through overt and devious means, public opinion against them is stirred precisely through the legitimisation and amplification of discourses that vilify them (as anti-nationals, for instance). An activist, Father Stan Swamy, jailed under the deeply anti-democratic UAPA law for fighting for Adivasi Indigenous rights in central India, died in custody.[11] A cruel irony is that many activists, such as Surendra Gadling, Sudha Bharadwaj, Rona Wilson and Father Stan Swamy, who have been jailed for years on false charges of being 'Naxalites', were fighting precisely for the rights of Dalit and Adivasi political prisoners. Apart from direct state sponsored harassment and imprisonment, public disinformation and vilification also has other effects that are internalised and have long term health consequences.

> If someone hits you, you obviously feel pain, and if someone hits you with their words you feel pain. I don't think there's much you can do about it. You have to become like some cold-hearted monster where nothing affects you anymore; but that's not a desirable thing to be.

> So yeah, it does affect you very badly, it lowers your self-esteem . . . I mean the government and ministers have called us "anti-nationals". But social media never lets you forget that. . . . Even today they are using the word in official communications. [Shehla]
>
> I was already in a pretty bad headspace, and seeing the media coverage of me was really difficult. Part of it is that they were constantly misgendering me and referring to me as a girl, but I think the part that really hurt is that I felt like my whole character was being put on trial. . . . And then I read the comments, and it made me feel scared for myself and for my own physical safety. . . [Meena]
>
> . . . the first thing I did is make my Instagram private and for a while I was still posting things on my Instagram story and I had a close friends list, so it would only go to them, but now I don't use social media at all . . . I only read things on Twitter to keep up with the news. [Meena]
>
> I used to get badly affected, I used to have nightmares and I couldn't sleep. I used to think that some of them are going to come to my house. I remember one election where I was really scared of going out because someone did recognise me and said 'you were the one who tweeted against. . .' Initially I used to be very closed about this feeling of fear, I wouldn't even tell my husband, who actually works with me. I never used to share with anyone that I'm feeling very disturbed. I used to act very brave and say it doesn't affect me but inside it was eating me up, it was impacting my work and sleep. [Now. . .] I don't give my opinions on social media anymore, I am very cagey about saying stuff – I self-censor a lot. [Shruti]

The twin constructions of nationalism and the figure of the 'anti-national' or 'urban Naxal' has been a longstanding authoritarian political project, with the previous Congress and coalition regimes complicit in previous decades, particularly targeting Adivasis working for labour and land rights, and critics in Kashmir and the Northeast. Under the BJP, targeting critics and dissenters across the country has become a vast government-sponsored strategy, with systematic discursive support from mainstream media and social media trolls. The 'drip-drip' of hate that our interviewees receive seeps into and acts on how they construct their identities and live their lives.

Much of the policy debate around taking down problematic content hinges on the extent to which content may incite violence or poses a clear and present danger conceived as *physical* danger. However, as explained in Chapter 1, we argue, 'violence' and 'danger' are defined far too narrowly in an attempt to protect an abstract and often disingenuous idea of free speech. The actual experiences of our subjects demonstrate that it is precisely a destruction of their free speech which recurs in response to subtler forms of threat

and symbolic violence. This may not always fall under the ambit of the law, but must be addressed by social media companies to protect multitudes of socially and politically at-risk users. It is imperative that affective forms of violence (which cause pain, fear, lack of sleep, and silencing) should be considered hate crimes. Since such violence has historically been targeted towards excluded and discriminated groups, there is a greater probability that those groups face different and tailored modes of violence online as well. Our analysis has found this pattern repeated across all the cases and countries.

What is to be done?

As with users from other countries, our research subjects in India spoke about a wide range of strategies and recommendations which we believe should be taken up seriously by social media companies and policymakers, as well as by international human rights organisations that work on discrimination and violence.

In India, there are (very few) institutional resources available for recipients of hate who need support, and most of these are created by feminist, Dalit and anti-Fascist activists themselves, in a personal or collective capacity:

> But I do report, I tell my groups of friends to report . . . For me, what has helped me to cope is common support groups, understanding what my rights are and recognising that it is a problem. [Shruti]
>
> I've become very reclusive and if you knew me before any of this happened that's very much *not* how I am. It made me quiet down a lot. [Meena]
>
> One shouldn't get scared. When one moves ahead without fear – that kills their power. Their aim is to scare you. But when you continue to do your work while scared then they get scared that we are doing so many things but she is not getting scared. So we should keep our focus on work, you should not ignore other things but should deal with it. . . . All the girls in our circle who are usually trolled, we all have made a group on WhatsApp and Facebook. So if a girl is trolled then we all get down to trolling them back. We get number of that boy then we take his case. We record the call and out it on Facebook. . . . 3–4 boys deleted their Facebook account. We are about 68 members in the group. At times there should 'behengiri' [aggressive sisterhood] also. [Preeti]

Inspiring as these accounts are, they demonstrate the immense labour that individual social media users in India have to undertake to grapple with discrimination and violence levelled at them, often demonstrating remarkable patience and resilience in the face of consistent threat, hate and hostility. There are signs of more experienced social media users self-organising

and fighting back, not just around specific incidents, but also in a more systematic way.

> There's an organisation in Karnataka called hate speech beda [we don't want hate speech]. They brought out a report last year on how Karnataka media is spreading hatred. It's documented for posterity but what is the action that we are aspiring for on the ground? The government is not reacting, nor is the police. Our documentation is all that we're doing now, *are we creating any impact?* Are we going to judge only by election results whether this kind of polarisation works? I'm hoping that people will see that this is not right, bigotry is not right. Once that realisation happens, or the other way is, if you have ten people who are behaving in this way, like what's happening in America, the voices of these people are getting drowned and these people are getting stronger [depicts a model with hands], it does not mean that these people don't exist, their racism and bigotry are still there. The other voices have become more powerful. That's the only way out for India. The so-called fence-sitters have to choose their sides. [Shruti]

Concluding on a slightly more optimistic note, the *Hate Speech Beda* group's efforts in Karnataka finally paid off. Following their campaign, on 23 June 2021, the self-regulatory broadcasting body National Broadcasting Standards Authority (NBSA) ordered a TV news channel, News18 – part of the powerful Network 18 group – to air an apology and pay a fine.[12] They also censured Suvarna News and Times Now for provocative and hateful coverage of Tablighi Jamat Muslims in 2020.

Conclusion

Several of our key informants, interviewees and focus group participants expressed disappointment and resignation with regard to social media companies taking prompt action against the hateful and discriminatory communication to which they find themselves subjected. Social media companies must recognise the current conjuncture in India as a crisis for democracy and for mental health that needs to be comprehensively and courageously addressed, beyond profit, algorithms and AI. Rather, if the social media companies are serious about democratic mandates and free speech, they must listen carefully to human and civil rights activist and users from historically marginalised groups who have been consistently targeted both online and offline, and act to protect them.

Groups that have been physically harassed and targeted historically based on their caste, gender, sexual orientation, class and religious identities, and/or intersections of these, continue to be targeted on social media. Although the latter mode (of social media abuse, incitement and threat) uses different techniques, it has no less serious consequences particularly in the area of free speech and mental health, leading in extreme cases to suicide, rape, lynching and/or targeted pogroms in which entire communities are expropriated and their homes burnt to the ground or razed. Given the enabling environment for such discrimination and hate against these groups at the highest levels of the ruling party and Modi government and from most state institutions including the police and local magistrates, it is disingenuous to tell affected individuals and groups to seek legal protection, report online abuse to the police and so on. Many of them as we have seen, have individually and collectively come up with strategies to fortify their own well-being. At the risk of imprisonment or other serious forms of censure and detriment, they continue speaking against authoritarian and fascist developments in their own local and national contexts. That our collective global democratic well-being rests on the shoulders of some of the most vulnerable and discriminated persons should be a matter of deep shame to technologists, scholars and policymakers alike.

Notes

1 Dr. Murali Shanmugavelan. (2021, March). Caste-hate speech: Addressing hate speech based on work and descent. *International Dalit Solidarity Network*. https://idsn.org/wp-content/uploads/2021/03/Caste-hate-speech-report-IDSN-2021.pdf
2 Scroll Staff. (2019, September 12). FactChecker pulls down hate crime database, IndiaSpend editor Samar Halarnkar resigns. *Scroll*. https://scroll.in/latest/937076/factchecker-pulls-down-hate-crime-watch-database-sister-websites-editor-resigns
3 Daniel Block. (2019, February 1). Data plans: How government decisions are helping Reliance Jio monopolise the government sector. *The Caravan*. https://caravanmagazine.in/reportage/government-helping-reliance-jio-monopolise-telecom
4 Scroll Staff. (2020, July 16). Why Google and Facebook are investing in Mukesh Ambani's Reliance Jio. *Scroll*. https://scroll.in/article/967618/explainer-why-google-and-facebook-are-investing-in-mukesh-ambanis-reliance-jio
5 Azim Premji University. (2021). *State of working India 2021: One year of Covid-19*. Centre for Sustainable Employment, Azim Premji University. https://cse.azimpremjiuniversity.edu.in/wp-content/uploads/2021/05/State_of_Working_India_2021-One_year_of_Covid-19.pdf
6 Buddhadeb Halder. (2020, December 17). How the BJP tried to manipulate public opinion on social media in favour of the CAA. *The Wire*. https://thewire.in/politics/how-bjp-tried-manipulate-public-opinion-social-media-favour-caa

7 Kamayani Sharma. (2019, April 1). Supporting role: How Bollywood acted under the Modi government. *The Caravan*. https://caravanmagazine.in/perspective/how-bollywood-acted-under-modi-government
8 Internet Freedom Foundation. *How the intermediary rules are anti-democratic and unconstitutional*. https://internetfreedom.in/intermediaries-rules-2021/
9 Geeta Pandey. (2021, July 10). Sulli deals: The Indian Muslim women "up for sale" on an app. *BBC*. www.bbc.com/news/world-asia-india-57764271
10 Asad Ashraf. (2016, October 31). A dark trade: Rape videos for sale in India. *Al-Jazeera*. www.aljazeera.com/features/2016/10/31/a-dark-trade-rape-videos-for-sale-in-india
11 Mrityunjay Bose. (2021, July 5). Tribal activist Fr Stan Swamy passes away hours before the hearing of his bail plea. *Deccan Herald*. www.deccanherald.com/national/tribal-rights-activist-fr-stan-swamy-passes-away-hours-before-the-hearing-of-his-bail-plea-1005065.html
12 Campaign Against Hate Speech. (2021, June 29). News 18 Kannada: Why fighting hate speech in news is crucial. *The Quint*. www.thequint.com/voices/opinion/network-18-apology-why-fighting-hate-speech-in-news-is-crucial-for-democracy

References

Ambedkar, B. R. (1968). *Annihilation of caste with a reply to Mahatma Gandhi; and castes in India: Their mechanism, genesis, and development*. Bheem Patrika Publications.

Andersen, W. K., & Damle, S. D. (1987). *The brotherhood in saffron: The Rashtriya Swayamsevak Sangh and Hindu revivalism*. Westview.

Anderson, E., & Jaffrelot, C. (2018). Contemporary South Asia Hindu nationalism and the "saffronisation of the public sphere": An interview with Christophe Jaffrelot. *Contemporary South Asia*, *26*(4), 468–482.

Ansari, K. (2009). Rethinking the Pasmanda movement. *Economic and Political Weekly*, *44*(13), 8–10.

Arya, S., & Rathore, A. S. (2020). *Dalit feminist theory: A reader* (S. Arya & A. S. Rather, Eds.). Routledge.

Athreya, M. (1996). India's telecommunications policy: A paradigm shift. *Telecommunications Policy*, *20*(1), 11–22.

Ayyub, R. (2016). *Gujarat files: Anatomy of a cover up*. Rana Ayyub.

Banaji, S. (2018). Vigilante publics: Orientalism, modernity and Hindutva fascism in India. *Javnost – the Public*, 333–350.

Bhat, R. (2020). The politics of internet infrastructure: Communication policy, governmentality and subjectivation in Chhattisgarh, India. PhD thesis, *London School of Economics and Political Science.* Etheses. http://etheses.lse.ac.uk/4175/

Bourdieu, P. (1984). *Distinction: A social critique of the judgment of taste*. Harvard Universtiy Press.

Bourdieu, P. (1990). *The logic of practice*. Polity.

Bürger, T., Lawrence, F., & Habermas, J. (1992). *The structural transformation of the public sphere: An inquiry into a category of bourgeois society*. Polity.

Chakravartty, P., & Roy, S. (2015). Mr. Modi goes to Delhi: Mediated populism and the 2014 Indian elections. *Television & New Media*, *16*(4), 311–322.

Chowdary, T. H. (1999). Telecom demonopolization: Why did India get it so wrong? *Info*, *1*(3), 218–224.

Ferguson, J., & Gibson, T. (2015). *Give a man a fish: reflections on the new politics of distribution.* Duke University Press.

Ghassem-Fachandi, P. (2012). *Pogrom in Gujarat: Hindu nationalism and anti-Muslim violence in India*. Princeton University Press.

Gorringe, H., & Rafanell, I. (2007). The embodiment of caste: Oppression, protest and change. *Sociology*, *41*(1), 97–114.

Hansen, T. B. (1999). *The saffron wave: Democracy and hindu nationalism in modern India*. Princeton University Press.

Jaffrelot, C. (2007). *Hindu nationalism: A reader*. Princeton University Press.

Khorana, S., Parthasarathi, V., & Thomas, P. N. (2014). Public spheres and the media in India. *Media International Australia*, *152*(1), 75–76.

Mamdani, M. (2004). *Good Muslim, bad Muslim: America, the cold war, and the roots of terror*. Pantheon Books.

Mankekar, P. (1999). *Screening culture, viewing politics: An ethnography of television, womanhood, and nation in postcolonial India*. Duke University Press.

Manor, J. (2010). Prologue: Caste and politics in recent times. In R. Kothari (Ed.), *Caste in Indian politics* (pp. xi–lxi). Orient Blackswan.

Rajagopal, A. (2001). *Politics after television: Religious nationalism and the reshaping of the Indian public*. Cambridge University Press.

Roy, H. (2009). *Telecom growth trajectory in India*. SSRN.

Sarkar, S. (1993). The fascism of the Sangh Parivar. *Economic and Politicaln Weekly*, *28*(5).

Sarkar, T. (2002). Semiotics of terror: Muslim children and women in Hindu Rashtra. *Economic and Political Weekly*, *37*(28), 2872–2876.

Sridhar, V. (2012). *The telecom revolution in India: Technology, regulation, and policy*. Oxford University Press.

Tambiah, S. (1986). *Sri Lanka: Ethnic fratricide and the dismantling of democracy*. University of Chicago Press.

Teltumbde, A. (2005). *Hindutva and Dalits: Perspectives for understanding communal praxis*. Samya.

Thakurta, P., & Sam, C. (2019). *The real face of Facebook in India: How social media have become a propaganda weapon and disseminator of disinformation and falsehood*. Oxford University Press.

Truschke, A. (2020). Hindutva's dangerous rewriting of history. *South Asia Multidisciplinary Academic Journal*, *24–25*.

Wilkerson, I. (2020). *Caste: The origins of our discontents*. Random House.

5 White male rage online

Intersecting geneologies of hate in the UK

This chapter is based on the testimonies of multiple informants ranging in age from their early twenties to their seventies. They have experienced trolling, doxing, hacking, stalking, targeted harassment, death-threats and violence on and offline in the UK. Contrary to the scholarly and media narrative which sees this kind of hate originating in 'Russian troll farms' or 'abroad', most of this hate is homegrown. We find ourselves at a frightening historical juncture which is repeatedly misunderstood and downplayed by being called 'post-truth' or 'populist'. This neglects the elaborate infrastructures of the political right and far right that now permeate everyday consciousness. Ruptures have reached breaking point between those who actually (rather than rhetorically) uphold freedom of speech, equality before the law, transparency, democracy or human rights, and those white and white-adjacent supremacist ideologues who mobilise around retrograde ideologies masquerading as pride in nation, rights of the individual and freedom of speech. Although we know the UK intimately, the painful and grotesque testimonies of our UK interviewees in the context of ongoing public struggles over definitions of antisemitism, anti-Zionism, feminism, sexual and gender identity, race, racism and resistance make this country strange in multiple ways.

As elsewhere, we insist on the overwhelming importance of history in understanding different forms of violent incitement, dehumanisation, discrimination and hateful content. We want to eschew presentist claims, even while the present envelops us in ever more bizarre circumstances. Writing in 2021, degraded public services are groaning under the onslaught of government cuts, mismanagement and graft. Children are going hungry and being abused while also being exhorted to learn online, to fulfil their potential and to achieve. While maintaining its own very British fiction of freedom and fairness, a largely unregulated media-sphere has moved significantly towards the expression of authoritarian and majoritarian values (now with the launch of a new alt-right news channel to mimic Fox news).[1] To these

DOI: 10.4324/9781003083078-5

processes, one might add a series of intertwined media representations and Government actions that are fuelling vicious collective prejudices. Curbs on migration and a suspicion of refugees and asylum seekers have led to *de facto* murder, with the Home Office criminalising the saviours of drowning refugees. Definitions of antisemitism, such as that by the IHRA, which encompasses those who uphold the humanity of Palestinians, have made it ever more challenging to point out the horrific humanitarian repercussions of settler colonialism and European exceptionalism in Palestine. The rise of celebrity Islamophobes and anti-Black racists whose visibility is linked to the anti-democratic opinions they spew in tandem with a deliberate fuelling of suspicion against anti-racism and anti-racists has had direct repercussions. This was clearly demonstrated when three young Black players for the England football team missed penalties in Euro2020 in July 2021 and found themselves on the receiving end of racist hate and disapproval.

Debilitating divides between those who voted for and against Brexit remain, while the worst economic repercussions of leaving the European Union are only beginning to be felt. Resurgent transphobia amongst parts of the British intelligentsia has mobilised diverse publics against an already dangerously marginalised group and exacerbated homophobia at the same time. All of this has been and continues to take place against the backdrop of the second year of the global Covid-19 pandemic which has caused a hundred and forty four thousand officially counted deaths in the UK as we go to print, left several million in poverty and jobless and provided an opportunity to an endlessly corrupt Conservative party and Government to line the pockets of their friends and donors. In the pandemic's wake, we have barely had time to count, name and mourn the dead, here and in diasporic 'home' countries. Meanwhile millions of people find themselves diverted by anti-vaccine conspiracy theories or caught up in the incapacity and unwillingness of capitalism and its favoured private systems to adjust to the shared burden and responsibility for material and psychic survival. In the UK these events and processes are filtered through relative and respective class privileges or burdens as educators, cleaners, bus drivers, retail, care and healthcare workers attempt to bear the unbearable risks of maintaining – and changing – this failing system.

Meanwhile, migration from former colonies in Asia, Africa and the Caribbean in line with government policies of the 1950s, 60s and 70s has been caught up in the post 9/11 demonisation of Muslims in the UK. British participation in wars against Afghanistan, Iraq and Libya, as well as the supply of weapons to Yemen and Syria has provoked an influx of newer, educated migrants and refugees from these countries. The drastic destruction of social welfare by a decade of ideologically motivated privatisations, cuts and closures covered up and excused by a media steeped in neoliberal dissimulation around economic mismanagement has ensured a constant

sense of resentment on the part of the mainly white citizens who view people of colour and migrant communities as competing with them for scarce resources which they think of as their birth-right.

As we analyse overlapping accounts of political abuse, racism, homophobia, transphobia, misogyny, trolling and physical violence in this chapter, we find ourselves returning time and again to these events and processes to explain, and counter, the simplistic and vicious imaginaries revealed in racist tweets and direct messages, dehumanising Facebook and Instagram groups, posts, racist viral videos on TikTok and YouTube and newspaper comments' sections. We also find ourselves repeatedly grateful to scholarship which theorises intersectionality (hooks, 2014; Crenshaw, 1989) and interrogates race and caste critically with attention to sociolegal frameworks and atrocity (Bell, 1995; Delgado & Stefancic, 2013; Davis, 1981; Teltumbde, 2010), since almost all accounts detail how overlapping identities and positionality within constructions of race, class, gender and sexuality exacerbate the harmful, vitriolic and violent content of social media hate.

Doing God's work: Faith-based support for LGBTQIA+ rights around the world

The intersection between organised religion and sexuality has historically been a troubled one, not only for the Church of England and Catholicism but across other religions too. LGBTQIA+ people of faith have experienced suppression, exclusion and active persecution both in the global north and in the global south. Worse still, much violent homophobic and transphobic persecution has been enshrined in law and is often first experienced in childhood. From the trauma of conversion therapy (Adamson et al., 2019), homelessness and playground bullying to that of being spat at and assaulted on public streets (Tyler & Schmitz, 2018), many LGBTQIA+ teenagers in the UK remain closeted for survival, with religious rhetoric often providing a convenient disguise for parental and community lack of support. Being a migrant or person of colour almost always adds a further layer to the hate levelled at LGBTQIA+ people of faith. While discourse in the UK has changed somewhat, largely due to the constant struggles of activist groups and individuals, much of the above still holds true in British faith communities.

We interviewed British-Nigerian Reverend Jide Macaulay only weeks after he had become a priest in the Church of England. We discussed the 14 years since he had founded House of Rainbow, a faith-based campaign group and space for 'sexual minorities, lesbian, gay, bisexual, transgender,

intersex, and queer people, and more particularly, Black Africans, Black Caribbeans and others'. House of Rainbow has extensive collaborations – from work in over 22 countries including Jamaica, Guyana, Lesotho, Malawi, the United States and Colombia. These collaborations mean that they have influence beyond Nigeria and the UK where much of the day-to-day work takes place. They hold workshops both for support and education, connect Buddhists, Muslims, Sikhs and other individuals of faith to Christians, have prayer, music, sermons, and food-based events to build community solidarity. They maintain safe spaces for improving mental health and well-being, with regular training events and webinars on combatting misinformation about what the Bible teaches about homosexuality. This can be seen collectively as a form of 'sanctuary'.

With a long history in African Indigenous and independent churches, including Pentecostalism and other charismatic ministries, Reverend Jide is a critical insider. His stance and that of the organisation is both personal and political:

> [House of Rainbow] started 14 years ago out of the need that also affected me as a gay man who had come out about his sexuality and was looking for a safe space where I can express my Christian beliefs and my understanding of God . . . the early missionaries and even some modern day missionaries, largely within the evangelical or extreme right-wing conservative Christians are making this forceful claim that God hates gay people and that homosexuality is an abomination. And of course, for the millions of sexual minorities around the world, it is inconceivable to have them brush on this canopy of hate and misinterpretation of the religious texts, regardless of the faith, regardless of the religion, whether it's Christianity or Islam or Buddhism or any other religion. I think that the essence of queer people is very important . . . House of Rainbow is the work of creating spaces for queer people of faith, and also bringing the message of affirmation to the people contrary to many conservative beliefs that 'God condemns homosexuality'. We believe God did not, nor did Christ, condemn same sex relationships. . . . Homosexuality is a taboo. HIV is a taboo. . . . So that is why we use social media, so that we can connect with people and people can connect with us. The traffic of people connecting with us is higher than what we can handle as an organisation. . . . We are *faith based* because we do not discriminate against people of other religions outside of Christianity. . . . We have colleagues who are gay imams or lesbian imams who take on the responsibility of teaching and nurturing and providing pastoral care for queer Muslims.

Reverend Jide and the other organisers of House of Rainbow face multiple forms of hate – from discrimination and legal threat to abuse and violence.

> We are considered an *abomination*. We are considered *against nature*. They believe that we came from the bottomless pit of hell. And so here we are, saying that god is loving, god is inclusive, god is liberating, god is a freedom and god is queer because this is all of who we are. . . . House of Rainbow started in the atmosphere where the Nigerian government had introduced the anti-gay bill to parliament . . . 14 years imprisonment with hard labour for anyone convicted of homosexuality. There was, about 5 to 10 years imprisonment for organising a gay group or assembly of gay people [including a religious one] . . . . From day one I received hate messages on social media. . . . [Under the blog we wrote] *the hatred was just unbelievable*. And just two days ago, I came across an article and I was reading it and then I read the comments about me and my ministry. *A few days ago, people were asking for me to be killed and executed* . . . . I report them to the police as soon as I get them. People have taken out a petition page on change.org against me as a "fake pastor". People have set up an alternative Twitter account in order to create a massive following of people that will hate me and cuss me. Yeah, so I have a designated police officer in London who I will just send all these things to. [Emphasis added]

One thing that struck us about the way in which Reverend Jide has been targeted and his account of the intense anxiety, grief, anger and fear that he has felt was the intersectional nature of the abuse, discrimination and violence he faces. From being stalked and threatened on the streets of Nigeria for being gay and openly Christian to having his profession mocked and ridiculed and being treated with contempt in the UK because of the colour of his skin on the street or because of his accented English when he went to an acting trial for a Shakespeare play in Glasgow, all the different facets of his life come under pressure and scrutiny from people's prejudices. This intersectional experience of hate was repeated time and again during our interviews.

The intersections of race and online abuse: 'If you're Black or Brown your life is political'

Grace Blenkinsop, a young BLM activist, told us that her passion for the movement was about a concern for the Black lives that often are left out of debates: 'Queer black lives, working class black lives, women of colour'. Grace uses Twitter, Instagram and TikTok to comment on current events

such as the death and 'memeification' of Breonna Taylor, with a high level of engagement, but also high levels of targeted hate messages:

> I've seen a lot of hate towards Black women, women of colour predominantly . . . Tik-ToK is such a *horrible, horrible site* for breeding hate, it's really nasty. . . . I have a lot of people comment on my videos, who don't follow me, and say that I'm always on their 'for you' page. I guess if the wrong person can see your videos, there's just a lot of racist hate, a lot of sexist, misogynous comments, comments on girls weight, appearance, especially with people that have links to the BLM movement, just racist, derogatory terms surrounding that. [Speaker's emphasis]

The overlap between different forms of identity such as race and gender or race and transness were key issues drawing the largest numbers of dehumanising and abusive comments online. Neither trying to ignore the comments knowing that there are a lot of sock-puppets (fake identities) and bots involved, nor putting on filters, has reduced the shock and distress of encountering hate in the intimate spaces of inboxes and comment sections:

> [D]uring the start of the year, when the whole BLM movement resurfaced, I was active making videos about protests and calling systems out, like the police and the government. . . . But people took it very personally. It was actually really troubling. I had a lot of comments on my videos – people saying that I should *die* or *get hit by something*, or *really graphically explaining how they want me to die or my family to die*. That was on Tik-Tok. Then, because my Instagram was linked to my Tik-Tok, after I'd block those people or delete the comments, it came on Instagram. I was just so confused as to why I was trying to help educate people and then suddenly I was getting told to *get hit by a bus*. It wasn't just comments directly about me, it was comments about the movements. People would DM me on Instagram with jokes about George Floyd's death. . . . I made a video about police officers being racist and I had loads of people message me like 'my dad's a police officer, f*** you'. Some people were really angry to the point where I would read their message and not reply because I wouldn't want to spur it on and they'd keep messaging me like, 'f*** you, N-word', and all of these racial slurs. [Speaker's emphasis]

Reading through the transcripts of our interviews for this book in the context of the torrent of racial abuse endured by Black British footballers Raheem Sterling, Marcus Rashford, Bukayo Saka and Jadon Sancho in July 2021

after the Euro2020 cup final penalty defeat, that reaction by a large section of the British public to these successful young stars (which shocked many white British commentators and engaged a response of disavowal) was utterly predictable. It reveals only the tip of the iceberg when it comes to everyday life for Black British citizens and many other people of colour in the UK.

Comedian Sandy A. (a pseudonym, at her request) explained forcefully that intersections of gender and race provoke much hateful content she receives and underpin experiences of discrimination. Given her occupational visibility and the fact that she's British Caribbean, she cannot hide. She told us passionately, 'If you're Black or Brown your life is political, isn't it? Unless you're stupid. It has to be, we're politicised as kids' and later 'Being an opinionated Black or Brown person in this country is going to bring you grief from everywhere'. In Sandy's experience, racism is a driver of much social media hate:

> Racism would be the biggest one. Sexism. And in racism, I include Islamophobia. That's another one that gets people going. Transgender issues get people going – that generates *a lot of hate speech*. Black people in adverts seems to upset people! Anything to do with identity, I think. Anything outside of white identity, and it used to be white male, but I think it's white female as well now because I think white women are just as bad . . . I think white supremacy adapts, doesn't it, and changes itself to survive. . . . So, the kinds of hateful stereotypes I see online are 'transgender women are really rapists in disguise – they're just trying to get at women'; 'Muslim people are extremely violent and they just want to Islamify Britain and steal Christmas'; 'Black people are stabbing everybody and lazy'; 'Jewish people are greedy', but that one's not so much anymore. . . . The really nasty ones will be talking about *the death of my daughter* . . . .
>
> I wrote an article about Islamophobia and then I had to write a follow up article because the response was so bad. It ranged from, 'go and live in Iran, or why don't you go and live in Iran' – I had a dress on in my profile picture at the time – to 'go and live in Iran *in that dress*'. I was like, 'why would I? It's a winter dress?' So, that would be the mild. Then it's like, 'take Muslims into your house, I hope they rape you'. Then, 'get out the country', then accusations of being a secret Muslim. . . . I had EDL [English Defence League] members saying they were going to come to my [workplace], and they were posting pictures of a noose and saying they were going to lynch me, and 'Hope you die like your daughter, you fucking Golliwog'. You'll report it and

> [the Platform] goes, '*It hasn't breached our community standards*'. [Speaker's emphasis]

Failed by the inadequate and biased implementation of rules on hate speech and harassment to apps and platforms, Reverend Jide, Grace and Sandy, like many other key informants, took issue with being told to just 'ignore' the racist, misogynist or homophobic threats and slurs, 'don't feed the trolls'; or to just 'have fun with the trolls' by trolling them back. As Sandy summed up: 'We *can't* have fun with them, because they *literally want us dead*'.

The refusal and/or inability to recognise racism in language or action, denial, identification with racists and systemic racism are all named as issues preventing moderation teams and the police from taking adequate action against those posting hateful content and misinformation. Reverend Jide and Sandy in particular have faced such serious threats that they now report regularly to cybercrime units and pursue platform moderation teams. They are sceptical about the ability and commitment of these institutions and intermediaries to change the overall context in which hate is manufactured, circulated and does its divisive political work.

Clayton Wildwoode, a working-class, white queer student who co-founded 'All Black Lives, Bristol'[2] at the age of 19, insists that racism, classism, misogyny and transphobia intersect, underly rising social tensions in the UK, and have been fuelled strategically by the 'Tory' elites with Government connections. As someone who counts himself an anti-racist 'ally', Clayton was clear that he was still learning about the profound effects of racism, misogyny and transmisogyny on Black communities and Black women, in particular. He explained how *All Black Lives, Bristol* often gets trolled on Facebook and Twitter by white people 'who write essays under our posts explaining that we're wrong because racism doesn't exist, or that we're being racist to white people'. In his experience dehumanisation and misrepresentation are incredibly varied, covering body shaming within the gay community, comments on appearance and gender, as well as suspicion of bisexuality and Biphobia. All of these, he noted, are exacerbated at the intersection of race and class. As such it is worth ending this section with Clayton's reflexive commentary, which emphasised that there is often an assumption that all minoritised groups support each other. There is, however, sexism amongst Black cis and gay men and racism and misogyny within the white gay community that comes out in online spaces as well as in discriminatory actions, attitudes and violence:

> Everyone that's not a straight white cis man thinks that, oh, 'if I'm queer or if I'm a different race, it automatically means I'm woke' and

> that you are accepting of everyone, but it's not true. There's so much misogyny in the gay community and then there might be different types of racism from different races. So just because you're not the stereotypical racist or sexist person, I think you kind of shrug it off as, 'oh I'm gay and there's prejudice against me so *I'm obviously not prejudiced against other people*', but that's not true.

We were deeply moved by and learnt a lot from the ways in which Clayton and our other interviewees used the space of the interviews to reflect on their relationships to different facets of identity in their chosen and birth communities.

'We're not allowed to be Jews because we're not the sort of Jews that they want us to be': Political censorship and the destruction of democracy

Clayton's and Sandy's commentaries were pointed with regard to UK institutional politics. In their opinion, stereotypes about Black people, Muslims and Jews could be found amongst commentators on the right and the left, and amongst majority and minoritised groups. Sandy was, however, very clear that this was most evident amongst members of the Conservative Party, and in factions on the right of the Labour Party who ousted Jeremy Corbyn (leader of the Labour Party from 2015–2019) after an insidious campaign of disinformation. To get a closer insight into the disingenuous and destructive tactics utilised by the Labour Right against their own members who support the Palestinian people's struggle or call out Zionist racism, we interviewed two long-time left-wing activists. Their experience is in trade unions and the Labour Party. Mike Cushman, a founder member and Membership Secretary of Jewish Voice for Labour (JVL) who frequently represents the group, explained how complex the tensions have become around trolling, anti-Zionism and antisemitism in the UK. He explained that although there are some 'lone rangers' who engage in trolling or sock-puppetry, much of it is coordinated by groups of hard-line Zionists with access to funding, media and social media:

> The fact that we are dissident Jews offends them greatly because they really don't know what to do with us. So, a lot of it is trying to deny that we're Jewish. We're not allowed to be Jews because we're not the sort of Jews that they want us to be. . . . [The trolls have] a vastly inflated idea of our impact and reach. They complain about conspiracy theories about Jews. Not incorrectly, because these can be very dangerous theories. But they then seem to engage in the same conspiracy theory about [members of JVL], ascribing to us all sorts of malevolent power and

> influence. . . . So things like we are 'pulling Jeremy Corbyn's strings', 'we had all this privileged access to Jeremy' (In fact, Jeremy kept us at a very great distance.) That we are 'a nest of anti-Semites', 'holocaust deniers'. The phrases go on and on and reverberate and reverberate. Their favourite label is 'cranks'. As soon as you say you're not a crank you're proving your crankishness. It's a double bind. . . . There was a whole Twitter feed – it was recently taken off – called "JVL Watch" which was people who just continually trolled us and harassed us . . . We're Jewish, and it's very hurtful. I know of at least one person who attempted suicide [because of this trolling].

In a similar vein, Labour councillor Jo Bird is on the left of the Labour party, and recounted the impact on her life when online and offline harassment against her intensified, including an unjustified suspension from the party in February 2019 over her support for Palestine. She has found herself with doors slammed in her face on doorsteps because she was canvassing for the Labour Party (which at the time had endless disinformation published against it in the mainstream media). She also had non-Jewish people telling her that she doesn't understand antisemitism. She's experienced bullying and harassment from those on the rightwing of the Labour party[3] amplified by the media because, as an anti-Zionist Jewish socialist, she did not 'fit' the Labour right and Zionist narrative about institutional antisemitism in Jeremy Corbyn's Labour Party.

> From my lived experience as a Jewish person in the Labour Party it's been a very welcoming place. I very quickly became a councillor. I've not, in day to day, face to face interactions seen antisemitism and I would notice [if it were there], *because I do know what it looks like*. I come from a generation of a family that know what anti-Jewish racism looks like. That's why my family is in the UK, not in Prussia, in Poland where it started off. So, me saying that doesn't fit the narrative that is being pursued by other organisations and the [Zionists inside and outside the Labour party] don't like that I continue to speak. . . . In October, maybe November 2019, [the hate campaign] effectively stopped me becoming a candidate to be an MP. So, the Jewish Leadership Council and Merseyside Jewish representative committee put out a joint Tweet statement saying how awful it would be [if] I was selected as an MP candidate . . . and how it would exemplify the antisemitism in the Labour Party – they didn't mention that I was Jewish. *The Jewish Chronicle* ran a story on it amplifying those voices – didn't mention I was Jewish again, erased that protected characteristic and the key part of who I am and what I do. And I wasn't shortlisted, wasn't invited for interview. . . . and back to

> my [suspension] in March, they tried to stop me from speaking. They were basically saying that the allegations that had been made against me were enough that I shouldn't be heard even to defend myself against those allegations. Three independent councillors walked out when I, a Jewish councillor, was talking about the Holocaust [in the council chamber]. And that is hatred. That is anti-Jewish hatred.

Similar to the Brazilian activists discussed in Chapter 3, Jo has endured an excruciating cocktail of hateful misinformation, with 'dossiers' of false or sensationalist stories denigrating her work and character spread by right-wing commentators and amplified by mainstream media. While much of the hate was aimed at the left of the Labour party, the toll on Jo's life has been severe, and in a recent statement on her unfair expulsion from the Labour Party, she evinces relief rather than disappointment. The political implications in terms of democracy and voice for those who challenge political authoritarianism and settler colonialism are equally acute. As Jo notes, the ultimate goal is to silence all dissent.

> The attacks on people like me are constant; it's just like these never-ending picking up on what you've said, what you've written, trolling, going through Facebook pages. Hundreds of people have been suspended or investigated for things that are just not at all a breach of the rules, for political disagreements. And *the combined effect of all of that is, is silencing*. So, people, including myself, we use Facebook less. We use Twitter less. We don't talk about Israel and Palestine as much. And we don't talk about antisemitism as much, and people don't respond as much either. It's no longer seen as a legitimate topic for political debate. It's seen as something that if you say anything about it, you could be suspended. You could have your reputation trashed in the press. You could lose your job. You can lose your position, elected position or your candidacy. There's fairly severe penalties in this society for that. . . . there was factions within the Labour Party that wanted the party to lose the last two general elections. . . . And the evidence is there in the public domain. And antisemitism, that constant smear of 'the leadership is racist', was an integral part of that strategy.

Jo's and Mike's accounts of political manipulation of public discourse in order to further particular political causes, prevent the ascension to power of pro-Palestinian activists, and ensure that any criticism of Israel's political violence or discrimination, and of Zionism is classified as a form of antisemitic hate speech, have been confirmed multiple times by scholars of this period and this politics from the late David Graeber to the poet Michael Rosen and the actress and activist Miriam Margolyes. These

Jewish critics and commentators have neither denied the existence of antisemitism within British society and in the Labour party more broadly, nor suggested that being Jewish means that someone is incapable of antisemitism; their positions are subtle, well thought-out and rights-based. They have all supported mechanisms for ensuring that racism of all kinds is called out and that Jewish individuals and communities in the UK are protected. Their well-evidenced stance, however, is that the exceptionalist narrative which conflates the actions of a state (in this case Israel) with a protected group across the globe should be challenged. In challenging antisemitism, anti-Palestinian and Islamophobic racism at the same time, like Jo and Mike in this chapter, all have followed a long history of Jewish socialism and internationalism. Losing this voice from the history taught to young people and from the media that circulates to make sense of contemporary events increases rather than lowers the visceral dangers of racism faced by both Jewish and Muslim communities in the UK.

The narrowing of the school curriculum[4] and focus to a Zionist ideology that pushes to exclude Palestinian voices within the British history curriculum is, to these commentators, another example of the selective and hypocritical championing of freedom of speech by the now mainstream alt-right. Fuelled by rightwing politicians in the UK with links to alt-right think tanks in the US, funded by corporate and political interests, and circulated on both social and mainstream media, the idea that anyone who stands up for Black or Palestinian rights is a 'racist' or is 'antisemitic', that those who challenge capitalism are vicious authoritarians harking back to Stalinism, has taken a strong hold of the imaginations of a significant section of the British public. Untethered from ethics, and targeted at particular demographics through Facebook, Twitter and Instagram, these ideas have gained currency in ways which the more complex counter-narratives have been unable to. Like the duplicitous slogans of the 'Vote Leave' campaign around the Brexit Referendum, ideas about 'mad Marxists' bankrupting the treasury and destroying British culture by pandering to minorities have a long history in far right propaganda campaigns from Nazi Germany and Fascist Italy to Hindutva India (Koonz, 2003; Caprotti, 2005; Banaji, 2018).

Mike turned to history to illuminate why we find ourselves in this situation of exacerbated public acceptance of mistrust and hatred against antiracists and socialists:

> Hate speech has always existed, but when you say it to five people, in a pub at 11 o'clock at night, it isn't the same as saying it on Twitter or Facebook . . . so the technologies are significant. . . . Add to that, the collapse of a socialist ideal has all sorts of effects. . . . We get this bizarre statement from the Department of Education that schoolteachers are

> not allowed to use anti-capitalist texts. *You defend freedom by abolishing it!* [Speaker's emphasis]

Jean Clemens (a pseudonym, at the request of our informant) is an academic at a well-respected British university who researches far right populism, racism, and the mainstreaming of the far right through an examination of elite discourse. Jean, who has a significant following on Twitter and also receives a range of vicious trolling there, points to the ways in which both mainstream and social media giving platform to the views of extreme racists has made fascist ideas increasingly acceptable: Through 'the hyping of the far right, whereby far right ideas have been given exaggerated platforms and ways to portray themselves as the voice of the people or as bigger than they really are in many ways':

> The trolling has been mostly on Twitter . . . I've been trolled by the typical trolls, the people who are anonymous, who just latch on. Interestingly, some of the more violent insults that I got were addressed to me *as a woman* because my name, Jean, for some people sounds like it's a woman's name. . . . it's not surprising that people who are racist are usually also sexist. For them, it's a way to justify them not reading my article at all. Just saying, oh look at that woman saying something rubbish. Quite interestingly, as well, I've been attacked sometimes by French people for commenting on French politics as they thought that I was American or British.

In our discussion with Jean it became evident that random posters' hateful commentaries on people's online posts are often based on xenophobic, racist and gendered assumptions about the recipient, so ostensible political disagreement is tethered to other, more visceral phobias and prejudices. Jean has managed to ignore much of the trolling, assuming that he isn't a big enough threat to the powerful for his data to end up on organised troll farm lists or machines. However, he has been targeted through his work and his job, which has caused him extreme stress and anxiety. He became a target of violent incitement as a consequence of signing a letter against hate:

> The final trolling that I got was death threats in the comments of an article that was written by other academics on a right-wing website and the article only mentioned my name, even though it was talking about a letter that had been signed by more than 200 people. And in the comments, there were some very clear death threats. Like some guy was

> talking about his gun in great detail and about how that gun could be the solution to political correctness. I challenged them, the office of the article, online on Twitter and they couldn't care less. . . . I think there's a bit of a difference between saying 'your research is a bit shoddy' and 'here is a gun'.

Jean was candid that putting himself and his research out there was a choice that he as a white male was making and that this distinguished him from others who receive hate due to protected characteristics. Yet he did not minimise the ways in which the trolling affected him. The emotional cost was anxiety and exhaustion. From discussions in Brazil, India and the elsewhere in the UK, one of the key distinctions we draw in regard to online hate and trolling is between the systematic, strategic and politically manipulated groups of trolls and the 'lone warriors' either seeking thrills, performing for their own followers, or airing their resentments and passionate beliefs in rightwing ideological positions. Discussing the kinds of people who engage in violent online trolling, Jean added another group – those who deliberately stoke controversy and prejudice for profit:

> The death threats make me feel scared. But then again, [I'm far safer] than many other people who are at the sharp end of racism, sexism. I've *chosen* to be a public figure, I've *chosen* to fight racism and the far right publicly . . . it's not like me suffering because of my gender, my race, my sexuality, something that is just who I am. But it's extremely tiring. And I'm not even talking about the death threats . . . just the basic trolling gets to you because when you dedicate so much time backing up your work, creating arguments, but also arguments that are really the basic kind of tenets of my work are: we live in societies that are deeply unequal, wouldn't it be a lot better for everyone if they were more equal, right? It's pretty basic. If we were more democratic, if we were more equal, if people had equal say. And then you get people who will fight against that idea not because of ignorance, but because they have something to gain out of it; and they will be disingenuous. They will be well-funded, they will have massive platforms and it's just massively disheartening. . . . For this group, I think the more dangerous ones, it's grift. They know that they can make money. They know that they can find followers using hate and hatred towards other people. Right. They know that racism sells. They know that dodgy understandings of free speech these days sell massively. . . . And they use the anger of certain sections of a population to sell books, to get money from right-wing think tanks. . . . [Speaker's emphasis]

Religious groups and individuals can engage in both individual and organised hate online and, while some is motivated by inegalitarian or conservative belief and fervour, there is also some that falls into Jean's category of 'grift' and 'profit'. As Reverend Jide recalled, this kind of hate can be a multimedia affair starting in a traditional medium such as radio or television and spilling over onto blogs, websites and social media platforms. The continuum of offline and online hate and harm can be even more damaging for those surrounded by communities and families who share their persecutors' hateful beliefs and values. The end result in Reverend Jide's view is a strong possibility of mental health breakdown and suicide attempts or suicidal ideation:

> There was one year where an entire radio station in Nigeria dedicated their entire day program on 'what should the Nigerians do to the homosexual pastor Jide Macaulay?' An entire day. Now, if you are a listener to that radio channel . . . let's not undermine the power of communication online. I mean, we talk about terrorism. We talk about home grown terrorists. Many of them have been impacted by the words that they share online within those communities. So, the homophobia that is driven online is also fueling homophobia in many places. Let's not underplay or downplay how it would then affect people like myself, psychologically and mentally. . . . If you have a Facebook handle with about a million people or Instagram with a million, you have a responsibility when you post something on your page because it will reach . . . 10 to 20% of those people will be picked up. . . . If your message is extremely racist and homophobic, it will ruin lives, there is no doubt about it. . . [W]hat we don't know is what happens to people in various spaces who do not have the power to deal with it. I can only imagine somebody who is in a house where they are homophobic and then their social media is filled with homophobic messages.

In Reverend Jide's view, entities or individuals with large social media followings, like mainstream media, have ethical and legal responsibility to stop the spread of misinformation. Yet time and again, we were told of the misuse of power and lack of action on the part of platforms, when those posting dehumanising and hateful content are savvy enough to disguise their language or couch stereotypes and disinformation in ways that appear to be valid debate. In this penultimate section, we turn to the fraught arena of LGBTQIA+ rights, homophobia and transphobia, where despite decades of systematic campaigning and greater public acceptance,

legally protected characteristics are not guarantors of safety at home, on the streets, or online.

LGBTQIA+ rights, homophobia and transphobia: 'A seed that grew into a hideous tree of hate speech'

In order to understand why it is both pointless and patronising to tell people who face daily abuse and threat online to 'just get off social media', it is important to examine the experiences of people who must perforce use social media regularly as part of their jobs. Ben (a pseudonym, at his request), an LGBTQIA+ charity communications officer, told us that he couldn't carry out his role at work without social media. His work requires the use of different social media for different audiences, and for different purposes:

> My whole life revolves around social media. So, I find that before I even start working, I'm already browsing Twitter and then when I start working, I immediately browse Twitter more. My job is incredibly varied. My job will be everything from running and planning our social media, so everything from writing a tweet to planning a communications strategy [and. . .] events as well. . . . in a work capacity, we use Instagram, but that's a very different kind of platform for us because our following is mainly young people . . . We have LinkedIn for advertising jobs. Facebook is, again, a slightly different audience of general supporters. I'm writing for all these different platforms with slightly different voices. Then, in my own life, I have no interest in Facebook, myself. So, Twitter and Instagram for myself. . . . I've gone through occasional periods of deliberately not looking at Twitter outside of workdays for mental health reasons, purely because there's been moments through the past few years where it's become really emotionally involving to my detriment.

To expand on why social media becomes such a drain on his mental health, Ben explained to us that transphobia and homophobia have grown in the UK, despite some improvements since 2000.

> I started to notice back in 2017 that there was a bit of a push back against the work that we were doing around supporting trans young people. When the government started to look at reforming the Gender Recognition Act at Westminster and in Scottish parliament, there was a

> lot of misinformation flying around *which essentially led to the painting of trans people as a threat to women*. I feel like that was the seed that has then grown into this really hideous tree of quite unveiled hate speech. . . . Back in 2017, I posted something that I thought was vaguely uplifting, something vaguely cheesy about recognising the humanity of everyone involved and being compassionate. And then it received over a hundred replies from people that wholeheartedly disagreed. I think that was the moment when I thought, oh, we haven't come as far as we thought we had . . . there's people out there searching for the word 'trans'. And then, there is a whole bubble of them that will then reply to each other's tweets and to each other's replies. . . [For example] Mumsnet was a site which, essentially, is like an old school message board site, but they also now have big advertising partnerships with companies. It was set up to be a forum for mums to share insights and build a kind of community. But there is a particular channel on there, called "Mumsnet women's rights" . . . Five minutes on there and you'll see that every single post is about trans people. I feel like some of the heat on social media, some of the debate is sparked by Mumsnet users screenshotting other people's tweets, posting them in these forums and that then generating another social media storm. I feel like that particular bubble is almost a bit of a radicalization community at the moment. It's pretty scary. . . . I have older trans friends who are literally struggling to leave their houses now for fear of how people react to them. Personally, as a gay man, I think there's a lot of parallels between what trans people are facing now and what gay men experienced decades ago. It's just like history repeating itself. **[Ben ext. 2, emphasis added]**

Ben recounted how comments targeted at him would say things like 'How disgusting, your mother would be ashamed of you [for supporting trans rights]'. His mother was upset by the comments, rather than by his principled position. Equally worrying, he felt, was the reaction to the educational work that his charity does. When they released a carefully evidenced and compassionate information pack for schools on LGBTQIA+ rights and support, they were deluged with hateful messages from people accusing them of promoting child abuse and trying to get the pack banned. Referring to the Thatcher-era law, Section 28, which saw the teaching of any material which 'promoted' homosexuality banned in UK schools, Ben explained that his own adolescence had been profoundly affected by the lack of information other than the homophobic content in the tabloids at the time.

Like Ben, Raymond Howell (a pseudonym, at his request) is a young, Black gay man who works for 'an LGBT international human rights

charity'. He too described how a large volume of hate encountered during his job and rights advocacy work is against trans people.

> I see a lot of misinformation, and willfully misrepresented information on trans identities: a lot of cherry picking of particular facts around trans people; a lot of trying to paint trans people as inherently antagonistic for existing; trying to paint trans people as potential abusers, when actually the vast majority of trans people that I know and work with are just people who are trying to get on with their lives.
>
> **[RH. Ext. 1]**

One of the most insightful aspects of our discussion with Raymond centred on his understanding of the intersection of racism and dating/sex – what he called 'sexual racism'. His attempt at pedagogic intervention online, however, proved to be a major trigger for all kinds of racism and violent threats against him.

> I've made infographics about sexual racism and racism in dating and in romance and in sex, because I think that's a topic which is really misunderstood and there's really interesting research on it that people don't know about. Another one was talking about racism and queerness and how there's a very strong relationship between homophobia, transphobia and racism. When I posted the infographic on sexual racism, it's a very controversial topic in the sense that people . . . have lots of preconceptions around sexual preferences and who they date, thinking that's personal and you can't touch that. And some of that is true, but there are ways that inherent biases interact with who you do or don't find attractive and it's important to unpack those, but it was something which went a little bit viral. A lot of people who agreed with me and found it interesting shared it and that led to people seeing it and disagreeing with it. Other people who disagreed with it shared as well . . . That was by far the worst instance of online abuse because it was very targeted. So I have some YouTube videos and there were people commenting and putting a thumbs down on every single video. People found my email address through a well-being workshop for queer people of colour and they started like sending me threats through email. Someone contacted my mother on Facebook . . . people sent me death threats, slurs and words that are related to parts of my identity, taking apart my appearance . . . It was crazy. It was really crazy. It had a massive impact on my well-being.

Raymond was accosted on the street by a man who filmed him on his phone and started telling him how much he hates Black people, using the N-word, and threatening him. After that incident and the barrage of homophobic and racist abuse and death threats, Raymond was exhausted, frightened, anxious and dejected and took several months away from social media. In discussion after discussion, we were told of the toll that attempting to engage with a wider audience online around racism, homophobia and transphobia takes on people's confidence, mental health and social interactions. And, while in this instance, much of the abuse was coming at Raymond from whites who could not bear to have their own racism analysed or pointed out, abuse, threat and violence is also directed at Black and Brown LGBTQIA+ people from within Black and Brown communities.

Ferhan Khan, another young LGBTQIA+ rights advocate, was keen to discuss the issue of identity, and to emphasise that the kind of pigeon-holing and typecasting of identities which takes place in online environments by those doling out dehumanisation and hate is far from the fluidity of experienced identity:

> I'm a non-binary male . . . but Muslim as well. At the same time as being Muslim, I'm also an atheist. So I'm a Muslim atheist. I'm a British Pakistani, I'm a Scottish Pakistani. I'm all of these things and that's okay, I don't have to be one thing. I'm also a transportation planner, a project manager, and a political activist, and an LGBTQIA+ activist, and trans right activist . . . I'm a Stonewall BME leader.

Ferhan recounts their early involvement in politics and LGBTQIA+ rights as a participant in a film for the Naz and Matt Foundation,[5] started by Matt after his beloved partner of 12 years, Naz, killed himself when he came out to his family and they rejected him. Ferhan's activism involves making videos with complex and diverse representations so that the media narrative does not pit Muslims (such as the conservatives in England who protest against LGBTQIA+ sexual and relationship education in schools) against the gay community (assumed to be white). However, while there are moments of connection and solidarity, and it brings them closer to other Black and Brown queer people, Ferhan also receives an avalanche of abuse for their social media postings on race and sexuality. Visibility and abuse are directly related.

> I tend to post a lot of stuff on social media, Facebook, Twitter, maybe even Instagram, about mental health, racism, LGBTQIA + topics. I put them out there, and then I just see all the racists come and find me and then scrap at me and say, 'you should just be grateful that you've not been thrown off the top of a building by ISIS'. And I'm like: 'okay

I don't feel like I should be grateful for anything like that, respect me' . . . . I have been getting a lot of hatred in the form of racism, in the form of homophobia from Muslim South-Asians, and racism from white people and white gay people, especially. I get a little bit of transmisogyny as well for my feminine side. . . . I'll get direct messages saying, 'you ought to go to hell. Being gay is wrong. It's against Islam', 'You're going to go to hell' or 'I hope you go to hell, hope you die', 'Hope you get bombed'. And then a racist person will say, 'I hope you got thrown off the top of a building by ISIS' and things like that. Then, there'll be those that do it publicly but framed as a debate, but actually it pretty much feels like racism, feels like homophobia. But, because it's framed in such an almost innocent childlike way, it just

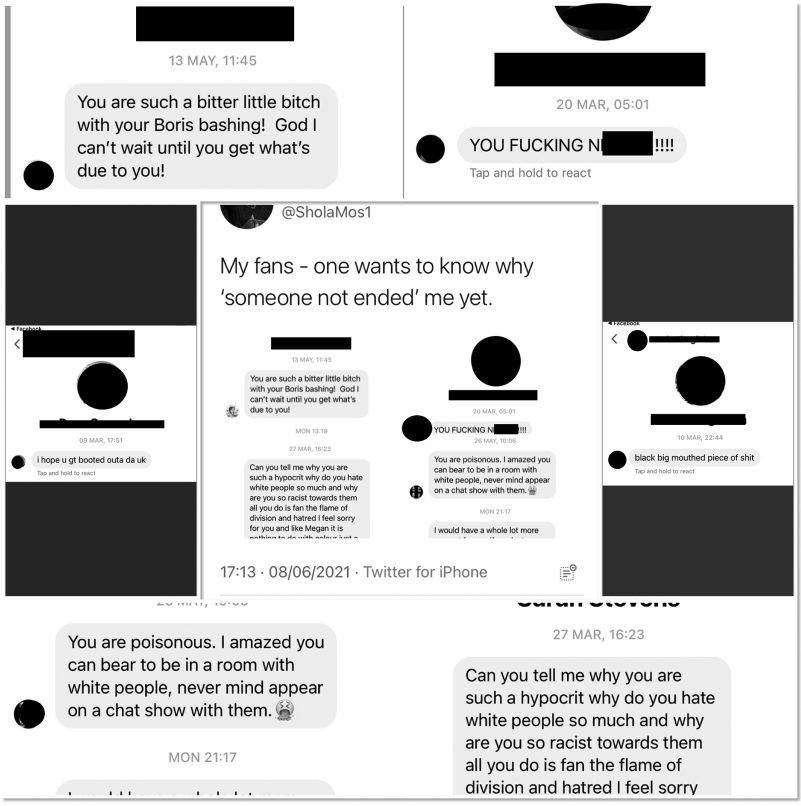

Figure 5.1 A selection of hateful content received on Twitter by Dr Shola Mos-Shogbamimu. Image credit: Dr Shola Mos-Shogbamimu.

> passes under the radar. So there's that kind. I got recognised on the street because I took part in a TV show. I was chased down the street before – that happened three times. That was in [London]. That was a group of South-Asian men.

When Ferhan appears on television, especially on the BBC, there is a spike in hate and threats, some of which spills over into the physical world before slowly lessening as time passes. We've heard this called the *paradox of visibility*, implying that the more visible one is, the more hate one is likely to attract. However, with one or two notable exceptions (for instance Jeremy Corbyn) our research in the UK confirms that the stereotyping, dehumanisation, abuse, incitement to violence and actual physical attacks attached to politically outspoken Black and Brown people of whatever religion, sexuality or gender, are generally far worse than those attached to the same sort of visible politics for white counterparts. This intersection of racism, misogyny, transphobia, Islamophobia and politics equates to a steady and deliberate effort to keep certain voices out of the public sphere.

Dr Shola Mos-Shogbamimu, an outspoken critic of racism and corrupt governance, has given us permission to reproduce some of the hundreds of abusive Twitter messages which she receives on a daily basis (see Figure 5.1).

Shocking as these messages are, they are the tip of the iceberg when it comes to the daily vitriol levelled at outspoken Black, Brown, Indigenous and Muslim women in Brazil, in India and the UK – and as we know, in Pakistan, Colombia, the United States and many other countries too. So, as we conclude this chapter, it is important to dwell on what we might call the 'hierarchy of hate' and its antecedents.

Conclusion

Some of the historical processes and events discussed in this chapter are traceable within the histories of colonial and nationalist fervour invoked by rightwing politicians and latterly by a host of other febrile nationalists: 'Take Back Control' was, after all, the slogan of Brexiteers. Interlinked historical events and processes stand out as the backdrop to the ways in which discrimination, abuse and violence, both on and offline, impinge on and injure the lives and life chances of contemporary communities in the UK. In the language that was reported to us as standard fare of trolling for Black and Brown commentators, activists, students, journalists and academics, we noted constant stark reminders of the British empire and its violent colonial legacy. Chattel slavery and the building of British institutions based on capital accumulated through genocide are frequently whitewashed or disavowed even in the institutions to which we are affiliated and within which we teach. Myths about the second world war — and the misleading rhetoric

British exceptionalism which denies the links between the British far right and Nazism — have led to a cosy relationship between supposedly centrist or conservative politicians, media houses and far right ideas. This has led to a mainstreaming of far right populism even by the supposedly left-leaning press (Brown & Mondon, 2021).

Other processes which we deem relevant to understanding the testimonies of the massive number of people who are at the receiving end of discrimination, prejudice and hate on a daily basis can be traced to deep-rooted social conservatism and the lack of a systematic social revolution in the UK. The ease with which transphobia has swept through the country and the sudden resurgence of biological arguments about gender amongst women who were supposedly committed feminists is deeply troubling. Further, when all is said and done, and despite the Chartists, the Fabians, the Levellers, the Suffragettes and many generations of abolitionists – much white working-class solidarity in overlapping struggles of Black and Brown members of the working class has only ever been fragile, while white middle class solidarity has often atrophied in the face of an assumed competition from Black and Brown fellow citizens. Farcically, but also with fascist undertones, recent Government-sponsored reports have loudly denied the existence of structural racism[6] and have tried to lay the blame for a lack of white working-class progression at the door of anti-racist teaching.[7] In tandem, the concept and foundations of feminism have been derided by an insidious backlash in popular culture and politics, and trade unions have been weakened by decades of rightwing curbs and legislation as well as anti-union propaganda in the mainstream media.

In the UK, which some like to claim is one of the 'oldest and most mature democracies', it is frighteningly easy to recognise parallels with the hate politics and media propaganda of the USA, India and Brazil. Recent eulogies of empire parallel the apparently popular scepticism of social justice work (Social Justice Warrior or 'SJW', 'woke' and 'cancel culture' now being terms of abuse used on message boards of the alt-right but also in the parlance of centrist journalists, politicians and academics). We are alert to the destructive effects of increasingly swift and bitter public shaming and factionalism amongst members of progressive groups online, and aware of instances of misinformation and sensationalism by left-leaning politicians or media. However, those who engage in struggles against racism, classism and sexism find themselves vilified by a majority of the UK's press and public service media which leads to a deep legitimisation of the kinds of vitriolic sentiments that are expressed against them on Twitter, Facebook, Instagram and other social media platforms. We return to this point, made in passing above, because it is the British mainstream media and their servile allegiance to a brutal Conservative regime alongside the vindictive machinations of a now

largely ineffective opposition party, which have forced dissenting voices into marginal spaces or into small but thriving alternative media enclaves and tied them ever more firmly to the use of social media.

Notes

1 www.theguardian.com/media/2020/sep/25/andrew-neil-launches-24-hour-new-channel-to-rival-bbc-and-sky
2 www.vogue.co.uk/arts-and-lifestyle/article/organisers-bristol-black-lives-matter-protest
3 www.independent.co.uk/news/uk/politics/labour-leak-report-corbyn-election-whatsapp-antisemitism-tories-yougov-poll-a9462456.html
4 www.theguardian.com/education/2021/jun/08/uk-history-education-row-israel-palestine-textbooks-pulled
5 www.nazandmattfoundation.org
6 www.runnymedetrust.org/sewell
7 www.forbes.com/sites/nickmorrison/2021/06/22/its-not-white-privilege-thats-to-blame-for-failure-of-poorer-white-students/?sh=58412035f470

References

Adamson, T., Wallach, S., Garner, A., Hanley, M., & Howell, S. (2019). *The global state of conversion therapy – a preliminary report and current evidence brief*. https://osf.io/preprints/socarxiv/9ew78/

Bell, D. (1995). Who's afraid of critical race theory. *University of Illinois Law Review*, *893*. https://heinonline.org/HOL/LandingPage?handle=hein.journals/unilllr1995&div=40&id=&page=

Brown, K., & Mondon, A. (2021). Populism, the media, and the mainstreaming of the far right: *The Guardian*'s coverage of populism as a case study. *Politics*, *41*(3), 279–295.

Caprotti, F. (2005). Information management and fascist identity: Newsreels in fascist Italy. *Media History*, *11*(3), 177–191.

Crenshaw, K. (1989). *Demarginalizing the intersection of race and sex: A black feminist critique of antidiscrimination doctrine, feminist theory and antiracist politics*. University of Chicago Legal Forum.

Davis, A. (1981). *Women, race and class*. Penguin.

Delgado, R., & Stefancic, J. (Eds.). (2013). *Critical race theory: The cutting edge* (3rd ed.). Temple University Press.

hooks, b. (2014[1984]). *Feminist theory: From margin to center* (3rd ed.). Routledge.

Koonz, C. (2003). *The Nazi conscience*. Belknap Press.

Teltumbde, A. (2010). *The persistence of caste: The Khairlanji murders and India's hidden apartheid*. Zed Books.

Tyler, K. A., & Schmitz, R. M. (2018). A comparison of risk factors for various forms of trauma in the lives of lesbian, gay, bisexual and heterosexual homeless youth. *Journal of Trauma & Dissociation*, *19*(4), 431–443.

6 Conclusion

Social media usage and the solidaristic, mutual, collegial and parasocial relationships it enables are an immensely important feature of contemporary society. This holds regardless of the varying numbers of social media users in each country, in our study, and across the globe. Politicians across the ideological spectrum routinely generate and attempt to influence political events via social media. This political discourse is amplified or critiqued and embroidered by broadcast and print media and woven into everyday discourse. Some of this discourse is treated as freedom of speech and expression. Some is not. Many ordinary users as well as activists or well-known journalists we observed and/or interviewed have faced systematic punitive action because of their social media posts related to social and economic injustice. They have been the targets of hateful communication, with dehumanising, discriminatory, threatening and/or abusive messages posted to and about them in public and private messages online. Of the hundreds of messages we evaluated to build our typology (see Table 1.1 in Chapter 1) and to give context to our interviews, and the thousands we collected since beginning our work on social media and disinformation in 2018, most constitute hate speech, some are implicated in hate crimes, and many remained online for extended periods, despite being reported to platforms and/or to the police.

Social media has become an important, sometimes *the dominant*, way for hundreds of millions of individuals to articulate their sense of self, construct their (racial, caste, gender, sexual, religious) identities, and address demands to politicians and corporations. For some social media is a *de facto* 'marketplace of ideas', where opinions are expressed and shaped, affects explored, solidarities built, and information or misinformation exchanged and contested. Although online usage is restricted to a relatively small part of the population in Myanmar and India, and mostly skewed towards dominant castes, males, urban areas, dominant languages and those who can afford the data or coverage, the *perception of its influence* (amongst

DOI: 10.4324/9781003083078-6

different actors) also produces real effects that inform daily practices which spill outside the domain of social media.

The clear-cut distinction often drawn between the online and the offline in terms of hate has been shown to be a distraction. Caste, class, race, disability, gender, sexual and religious inequalities and prejudices from real life are rife on social media, while social media use also informs the ways in which we construct and shore up our racialised, caste-informed, gender, sexual and class identities. Existing power geometries are enhanced through and inform relationships online, pointing to the online and offline, the material and the discursive as being *inseparable*.[1] Infrastructural and political economic approaches which informed our analyses in Chapters 2 and 4 of this book are also helpful in illustrating the unstable distinctions between online and offline, media, state and business, public and private, national and international, local and global.

The advantages of our theoretical framework

Our theoretical framework in this book combined critical race theory, cultural studies, cultural materialist, historical, infrastructural, phenomenological and political-economic approaches to social media hate ecosystems. Market pressures and geopolitical allegiances lead to uneven ways in which hate and violence have been handled in different parts of the world by governments and conglomerates. These different institutional responses are, in turn, implicated in the rise or fall of extreme authoritarian regimes, the undermining or shoring up of democratic practices and the push towards or disavowal of social justice and equity.

On the one hand, drawing on the concepts of intersectionality and othering (and concomitantly invoking the notion of *intersectional othering*), our framework directed attention to the implication of histories and infrastructures of colonialism, casteism, racism and capitalism in hierarchies of hate. From the *creation of difference* through mediated propaganda and disinformation to the consolidation of political and economic power in the hands of elites and majorities, the production of targets for populist rage and violence has been an extended sociopolitical process that now encompasses mainstream and social media.

On the other hand, paying close attention to individual and community memories and discourses of struggle, solidarity, exclusion, othering, dehumanisation, threat and violence, we developed a particular form of deep listening as the core of our methodological and theoretical approach. Applied to original accounts from recipients and survivors of hate, such phenomenological 'listening' foregrounds the contribution of memories and discourses not only to hateful communication and conscientised resistance but also to legal and social media policies on hateful communications, and to the way

in which the balance between freedom of speech and the right to life and dignity are imagined and enacted.

Based on this, and emerging from our literature review and analysis of the data, we offered a definition and typology of social media hate, and its producers, circulators and recipients which will have far-reaching consequences if adopted and understood as a backdrop to social media hate policies, media literacy programmes and in the development of ethical AI to limit hate online and moderate or remove it systematically.

Typologies of hate

As we discussed in Chapter 1, there have been numerous attempts to categorise and classify different forms of hateful communication in different national or community-based contexts. The typology of hate that we delineated in the Introduction (see Table 1.1) emerged from our analysis. Amongst perpetrators, as much as amongst those they target, overlapping aspects of identity are crucial in determining their social and political status and interests in hateful content. There were certain groups – particularly white people, evangelicals and/or upper-caste Hindu males from digitally literate backgrounds – who were most likely to be highlighted as perpetrators of hate on social media in Brazil, India, and the UK. The kinds of hate we examined are targeted both at individuals and more broadly at othered groups. The literature, meanwhile, shows that aggressor groups are some denomination of Muslim in many Muslim-majority countries (Bangladesh, Indonesia, Malaysia, Pakistan, Saudi Arabia), Jewish in Jewish-majority countries (Israel) and Buddhist in some Buddhist majority countries (Myanmar, Sri Lanka). To be precise, it is not just outspokenness, not just frequency of use, and not just the intersection of race or caste with gender/sexuality/disability/age that makes some individuals and groups more likely to be targets: Even within racial and sexual groups there are subgroups who are most frequently and virulently targeted 'just for existing'. So Black women – especially working class, Muslim, gay, and trans Black women, non-binary individuals and trans men face some of the highest levels of violence and abuse on and offline while Dalit women and transpeople, and visibly Muslim women (especially those who veil or those who openly reject the veil) are targets of multiple forms of discriminatory and abusive communicative behaviours. East-Asian women too face a growing amount of racist misogyny online. This is the case even when individuals do not speak out publicly about issues of rights and justice, but simply comment publicly on mundane issues.

These discriminated groups have been shown to be at the sharp end of online hate, closely followed by Black men, Dalit men and Muslim men in all these intersecting categories in all our case study countries (and in

particular, any men in these categories who present as femme and/or advocate for the poor and/or minoritised and oppressed groups and are broadly *on the left*). Notably, hate speech is also aimed from within oppressed groups at other minoritised and oppressed groups. In all of our case study countries and more generally, this is most clearly the case in regard to the disabled and to sexual and gender minorities within racial, ethnic and religious minorities.

We also observed that there are other types of uncivil speech online, often aimed within quasi-progressive groups at those who deviate or dissent from their views or from the accepted or agreed position and line, or at those who have fallen foul of a dominant block within their own communities of choice or practice. Based on our analysis of the literature in Chapter 1, most of these communications, while potentially unpleasant or distressing, and which may cause serious mental health issues and forms of shame to individuals in these circumstances, do not fall into the category of hate speech. Nevertheless, there are times when such discussions make use of the same tactics and prejudiced tropes and/or stereotypes as dehumanising and discriminatory communication does, reinforcing marginalisation, and thus spilling over into toxic or hate speech.

Moving forward: Suggestions for policy and politics

Our analysis of social media and hate incorporates the dialectical relationship between social media, users, production and reproduction of hate (in its material and symbolic forms) within specific national and historical contexts. Precisely because of its dialectical nature, this proves to be a space of flux, trauma and opportunity depending on the context. When groups have been historically vulnerable, social media gives them a public platform but also bombards them with new forms of violence and violation which have far-reaching consequences for their minds and bodies, work, home lives and communities of birth, choice and practice. Our analysis provides the grounds for arguing that international political bodies such as the African Union, the UN and the Council of Europe, and social media corporations such as Alphabet, Facebook (or whatever other name this entity chooses to call itself), ByteDance and Twitter, have a duty to act now, before matters reach extreme stages as they have in several of our case study countries, in taking decisive ethical action. This action should be planned and monitored transparently by well-trained (historically and politically literate) AI, moderation and policy teams, and expert, cross-national, cross-stakeholder groups. Alongside this, both political movements and corporations need to bring pressure to bear on national governments to abide by humanitarian standards and national legal frameworks with regard to the right to life and dignity of those with protected characteristics.

Social media companies must invest greater financial and human resources in countries where there are significant number of users as well

as historical political and religious patterns of harm, such as Brazil and India, and in countries where their platforms and/or applications have been used to exacerbate historical discrimination and violence, such as the UK, Indonesia, Myanmar and so on. This holds true for other social media giants with Chinese and Russian stakeholders operating in China and Russia too. While it may not always be possible to have all personnel located within the country for various reasons, social media companies should have clear and transparent policies to employ staff (moderators, content managers or other kind of editorial resources) proportionate to the number of subscribers in that country, the languages used within the country, and the scale of violence that social media usage enables.

Further, individuals from vulnerable communities need to be employed by mainstream media houses and social media companies as fact-checkers and to train and sensitise staff appropriately so that content moderators and moderation guidelines are compatible with the lingusitic, socio-political and historical contexts that inform much of the discrimination, threat and violence on social media. Our analysis for this book and for other projects over the past five years shows consistently that hate and disinformation circulates transmedially and in intertextual ways: Namely, it is not solely a social media problem – it crosses genres and often appears in fiction and non-fiction contemporaneously. Disinformation that targets minoritised groups and propagates hateful content online feeds off, is linked to and inseparable from content circulating in political speeches, and on mainstream and hyperlocal media systems, as well as within entrenched community imaginaries.

Engaging children, young citizens, the middle-aged and elderly in separate, carefully fact-checked and humane community education with a digital media literacy component may be the bedrock of imparting civic values and human rights as things currently stand. Even these kinds of programmes often don't result in widespread change since they are few and far between, often preach to the converted, exclude the most disadvantaged and allow systemic trolls and ideologically motivated racist and misogynist actors to avoid them altogether. To be even moderately effective, such critical digital literacy programmes would need be ubiquitous, generated and run by combinations of marginalised and oppressed groups and communities, and reflexive about change to demonstrate and illuminate the intimate connections between current and historical conflicts, representations, meaning-making and flows of power.

Meanwhile, technological affordance-based approaches such as AI and algorithms to take down problematic content need to be significantly and fundamentally overhauled after detailed consultations with members of communities who have faced such violence, scholars who have studied it, and ethical technologists. Facebook's actions in engaging with the Black

Lives Matter movement after George Floyd's racist murder, in order to overhaul its algorithms, needs to be applied urgently and widely to other parts of the world too and by other social media platforms and apps including, but not confined to Google, Twitter, YouTube, Instagram, WhatsApp, Telegram, and TikTok owners, ByteDance.

Given individuals' and groups' deeply unsatisfying experiences of reporting hate-speech and overt incitement to moderators on platforms such as Facebook, Twitter, TikTok and YouTube, and their equally depressing experiences of inaction, apathy, contempt and harassment from police and legal entities when they attempt to pursue their attackers in order to gain a modicum of safety, it is clear that following regulations and policies on discrimination, hate speech and incitement to violence does not always protect against it or even prevent ensuing violence. Efforts to reduce social media hate to a technological phenomenon in which algorithms are gamed for popularity or prestige, one that is engaged in by only fringe extremists, one that is populist and used both by far left and far right or one that is about 'fake news' spread by digital illiterates all play into the hands of fascists, rightwing groups and the biggest, most organised spreaders of hateful disinformation as well as unaffiliated individuals who harbour racist and misogynist views.

Likewise, the attribution of all malign online influence to deranged loners, Russian troll farms or western imperialism obfuscates rather than addresses the multiple and powerful sources and beneficiaries of systemic violence, dehumanisation, abuse and prejudice. In this context, a mixture of short and medium-term strategies mentioned by our interviewees – ethical AI, media education and fact-checking, need to be deployed urgently to prevent even worse excesses and crimes. At an individual level, blocking, filtering, taking frequent breaks from social media, leaving platforms entirely, getting therapy when affordable and speaking out about harassment and threat are also important medium-term protections. In the long-term, however, social media hate can only be defeated by international struggles over rights and justice, strikes, boycotts and social movements pushing for profound and far-reaching social and economic change that encompasses all current axes of inequality and injustice.

Discouraged but still hopeful

By keeping the voices of marginalised communities and targeted users at the heart of our book, we have highlighted the long-term individual and collective effects and affective consequences associated with the dramatic increase in the circulation of hate catalysed by social media. While their experiences speak to the immense varieties of hate – including an overlap

of direct and indirect violence, discrimination, threat, incitement, dehumanisation, disinformation and abuse – our key informants and interviewees have also convinced us of the possibilities for a better future. Astoundingly, despite the bleakness of the situation, despite the lack of structural support and the multiple realms of structural violence and discrimination they face, we also noted within our interviewees' narratives something like Camus' wrenchingly beautiful lines:

> In the midst of hate, I found there was, within me, an invincible love. In the midst of tears, I found there was, within me, an invincible smile. In the midst of chaos, I found there was, within me, an invincible calm. . . . In the midst of winter, I found there was, within me, an invincible summer.
>
> (From 'Return to Tipasa')

We hope that this was also your experience of our book. The only thing that we would add is that this resilient and powerful sense of hope, and the courage to act which transcended the fear induced by being the targets of historic and contemporary hate, were not individual phenomena, as in the case of the cited existentialist theorist. Rather, these feelings and the actions they gave rise to were expressed both through personal courage and through solidarity that can best be theorised as a commitment to conscientised and conscientising praxis.

Note

1 This intersection of discursive and material worlds has been acknowledged in the critical media and cultural studies tradition for decades. See work by scholars such as Stuart Hall (1997) and Roger Silverstone (1999) for context.

Reference

Hall, S. (1997). The work of representation. In S. Hall (Ed.), *Representation: Cultural representations and signifying practices* (pp. 13–64). Sage Publications.

Index