The Chilling:
A global study of online violence against women journalists

EDITORS

Julie Posetti and
Nabeelah Shabbir

ICFJ ONLINE VIOLENCE PROJECT LEADERSHIP TEAM:

LEAD RESEARCHER/AUTHOR: Dr. Julie Posetti, Global Director of Research, International Center for Journalists (ICFJ); Senior Researcher, Centre for Freedom of the Media (CFOM), University of Sheffield; Research Associate, Reuters Institute for the Study of Journalism (RISJ), University of Oxford

SENIOR RESEARCHERS: Prof. Kalina Bontcheva (CFOM); Prof. Jackie Harrison (CFOM); Dr. Diana Maynard (CFOM); Nabeelah Shabbir, Senior Research Associate (ICFJ); Dr. Sara Torsner, University of Sheffield (CFOM)

RESEARCH ASSOCIATE: Nermine Aboulez, ICFJ researcher and University of Oregon PhD candidate

REGIONAL RESEARCH TEAMS:

AFRICA: Assoc. Prof. Glenda Daniels (Regional Lead); Fiona Chawana; Dr. Omega Douglas; Dr. Julie Posetti; Nabeelah Shabbir; Alexandra Willis

ARAB STATES: Nermine Aboulez (Regional Lead); Dr. Julie Posetti; Nabeelah Shabbir;

ASIA AND THE PACIFIC: Assoc. Prof. Fiona Martin (Regional Lead); Liana Barcia; Dr. Ayesha Jehangir; Nirasha Piyawadani; Dr. Julie Posetti; Dr. Jenna Price

CENTRAL AND EASTERN EUROPE: Dr. Greta Gober (Regional Lead); Jen Adams; Bojana Kostić; Nabeelah Shabbir

EUROPE AND NORTH AMERICA: Dr. Julie Posetti (Regional Lead); Dr. Greta Gober; Prof. Jackie Harrison; Nabeelah Shabbir; Dr. Sara Torsner; Prof. Silvio Waisbord

LATIN AMERICA: Dr. Luisa Ortiz Pérez and Dr. Yennue Zárate Valderrama (Regional Leads); Dr. Kate Kingsford; Carolina Oms; Dr. Julie Posetti; Nabeelah Shabbir; Kennia Velázquez; and Prof. Silvio Waisbord

SPECIALIST RESEARCHERS: Becky Gardiner and Angelique Lu

UNESCO COORDINATION: Saorla McCabe, Theresa Chorbacher, Guy Berger, and Guilherme Canela

PROJECT SUPPORT: Jen Adams; Fatima Bahja; Heloise Hakimi Le Grand; Mark A. Greenwood; Mora Devi S. Hav; Senka Korać; Juan Mayorga; Cristina Tardáguila; Eunice Remondini; Erin Stock; Joanna Wright; Mengyang Zheng (ICFJ/CFOM); Johann Bihr; Sara Bonyadi; Annina Claesson; Lou Palin; Dana Muresan; Antonia Eser-Ruperti (UNESCO)

ICFJ PROJECT PARTNERS: Centre for Freedom of the Media (CFOM), University of Sheffield; Dart Asia Pacific; Ethical Journalism Network (EJN); International Association of Women in Radio and Television (IAWRT). This project has received financial support from UNESCO's Multi-Donor Programme on Freedom of Expression and Safety of Journalists and the Swedish Postcode Foundation

PEER REVIEW: Damiano Giampaoli, Office of the Director General – Division for Gender Equality, UNESCO; Isabel Tamoj, Section for Global Citizenship and Peace Education, Education Sector, UNESCO; Angela Quintal, Africa Program Coordinator, Committee to Protect Journalists (CPJ); Dr. Abeer Sady, researcher, journalist, and media development expert.

GRAPHIC DESIGN: Luiza Maximo

COVER AND ILLUSTRATIONS: Franziska Barczyk

With the support of

CONTENT WARNING

This document includes graphic content that illustrates the severity of online violence against women journalists, including references to sexual violence and gendered profanities. This content is not included gratuitously. It is essential to enable the analysis of the types, methods and patterns of online violence.

. . . .

Contents

Appendix 1: Country Case Study Summaries

Foreword

Maria Ressa[1]

*Nobel Peace Prize Laureate and UNESCO/Guillermo Cano
World Press Freedom Prize Laureate, 2021*

When I first came under attack online in 2016 after Rodrigo Duterte was elected the President of the Philippines, the deluge of hate was overwhelming. I was a CNN war correspondent for two decades, but nothing in the field prepared me for the orchestrated, misogynistic attacks on me and our women-led news outlet, Rappler.

Then, at a pivotal moment in 2017, when I was still just learning to deal with this new networked weapon against journalists, this study's lead author - Dr. Julie Posetti - convinced me that I should be speaking up about these attacks – and within 24 hours I did an interview with her for a book she was working on for UNESCO, aptly titled *An Attack on One is An Attack on All*. So, I took a risk and went public about the online violence I was living with. That was when I learned that the only defence a journalist has against these attacks is to do our jobs - to shine the light.

That same year, Philippine government propagandists tried to get the hashtag #ArrestMariaRessa trending. They failed, but they kept at it, and two years later I was arrested – twice – in a little more than a month. At the time of writing, I am fighting a conviction for criminal cyber libel and seven other cases being prosecuted by my government.

In the past six years, I have been repeatedly threatened with rape and murder, dehumanised in ways I couldn't have imagined in a disinformation-laced campaign designed to shut me up, and shut Rappler down. Our data and investigations show these attacks are fuelled by the State, enabled by the technology platforms, and directly linked to the legal harassment I experience in the Philippines that threatens me with decades in prison.

---

1 Maria Ressa is a Filipino-American journalist who is the CEO of Rappler in the Philippines, a news site she co-founded in 2012. Ressa's case is regarded as emblematic of the convergent contemporary threats faced by women journalists. See the big data case study presented in Chapter 3.1

Women journalists targeted by online violence at the nexus of racism, homophobia and other forms of discrimination are even more vulnerable to networked disinformation and gaslighting designed to silence critical reporting and bury the truth, as you'll read in this book.

Crucially, it delivers unassailable data to demonstrate its findings and recommendations, including analysis of over 2.5 million social media posts directed at me and my UK colleague Carole Cadwalladr. That data analysis proves that we are being prolifically attacked online because we are women, and because we are journalists who refuse to be intimidated into silence.

Even worse, these networked attacks are changing the people who see them and who take part in them: social media has become a behaviour modification system that is bringing out the worst in all of us, creating emergent behaviour that strips out the best of humanity.

The Chilling is a vital, at times gut-wrenching call to action - one that we cannot afford to ignore.

Manila, July 2022

Maria Ressa at the Nobel Peace Center in Oslo, Norway, December 2021.
Picture credit: Johannes Granseth / Nobel Peace Center

Executive summary and key findings

Online violence against women journalists is one of the most serious contemporary threats to press freedom internationally. It aids and abets impunity for crimes against journalists, including physical assault and murder. It is designed to silence, humiliate, and discredit. It inflicts very real psychological injury, chills public interest journalism, kills women's careers and deprives society of important voices and perspectives.

This ground-breaking three-year global study on gender-based online violence against women journalists represents collaborative research covering 15 countries. It is the most geographically, linguistically, and ethnically diverse scoping of the crisis conducted up until late 2022. The research draws on:

- **The inputs of nearly 1,100 survey participants and interviewees;**
- **2 big data case studies examining 2.5 million social media posts directed at Nobel Laureate Maria Ressa (The Philippines) and multi award-winning investigative journalist Carole Cadwalladr (UK);**
- **15 detailed individual country case studies.**

The Chilling illuminates the evolving challenges faced by women journalists dealing with prolific and/or sustained online violence around the world. It calls out the victim-blaming and slut-shaming that perpetuates sexist and misogynistic responses to offline violence against women in the online environment, where patriarchal norms are being aggressively reinforced.

It also clearly demonstrates that the incidence and impacts of gender-based online violence are worse at the intersection of misogyny and other forms of discrimination, such as racism, religious bigotry, antisemitism, homophobia and transphobia.

Further, it identifies political actors who leverage misogyny and anti-news media narratives in their attacks as top perpetrators of online violence against women journalists, while the main vectors are social media platforms - most notably Facebook, Instagram, Twitter and YouTube.

The multifunctional role of disinformation in orchestrated online attacks on women journalists is also examined, along with the exacerbating context of rising populism, far-right extremism, and the COVID-19 pandemic.

In addition, the online-offline violence trajectory is mapped. It represents a vicious and self-perpetuating cycle: digital harassment and threats beget offline attacks; and offline abuse (e.g., presidents targeting women journalists during

15 COUNTRIES STUDIED

KENYA

NIGERIA

SOUTH AFRICA

LEBANON

TUNISIA

PAKISTAN

THE PHILIPPINES

SRI LANKA

POLAND

SERBIA

BRAZIL

MEXICO

SWEDEN

UNITED KINGDOM OF GREAT BRITAIN AND NORTHERN IRELAND

UNITED STATES OF AMERICA
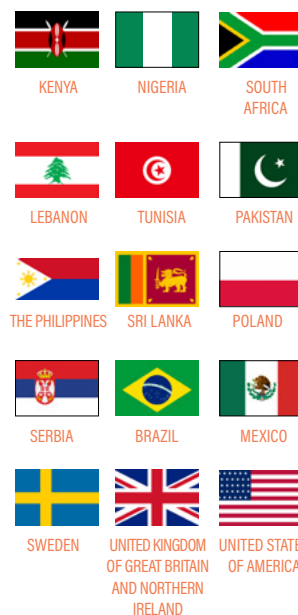
public appearances) can trigger an escalation of online violence which, in turn, heightens offline risks. Targeted online attacks on women journalists are also increasingly networked, and at times State-linked.

The war in Ukraine had only just begun at the time this study was being finalised in early 2022. However, the function of online violence as a weapon of war which is enmeshed in kinetic conflicts was already in evidence.

All this points to the need for responses to online violence to be strengthened in technological sophistication and collaborative coordination.To that end, **in Chapter 7.0, you will find a 25-step tool to assess responses to online violence to help guide freedom of expression-respecting interventions.**[2]

And, in addition to the 35 key findings presented below, the study offers 107 practical recommendations for action[3] tailored for intergovernmental organisations, States, Big Tech, the news industry, legal and judicial actors, and civil society. These recommendations emphasise the following points:

- Intergovernmental organisations, States, social media companies, and civil society organisations have an important role to play in developing early warning systems to monitor, predict and prevent online violence escalation;

- For women journalists to be able to work safely online, Big Tech's business models and algorithms that have been found to drive hate, and prioritise profit over human rights, must be overhauled;

- News organisations need to develop gender-aware protocols to respond to online violence, stop victim-blaming, and avoid disproportionate restraint on the speech of women journalists when they come under attack;

- Instigators and perpetrators of gender-based online violence must be held to account, and be de-platformed and penalised where appropriate;

- A more inclusive approach should be adopted to recognise and call out the intersectional[4] nature of online violence which exacerbates abuse against the women journalists targeted;

- Law enforcement agencies need to develop gender-sensitive digital investigation capabilities, and work more collaboratively, while judicial actors should better understand the ramifications of gender-based online violence as a workplace safety, press freedom and gender equality issue, as well as a potential indicator of future offline harm;

- States need to take action to protect women journalists from online abuse, harassment and threats, recognising the risks posed to their safety and press freedom; but legislative and regulatory interventions must respect international freedom of expression rights and obligations.

---

2 See Chapter 7.

3 See Chapter 7

4 'Intersectionality' is a term which was coined by the scholar Kimberlé Crenshaw, to describe how differing social aspects of a person's identity intersects - for example with race, class, gender - particularly in the experiences of marginalised groups: https://chicagounbound.uchicago.edu/cgi/viewcontent.cgi?article=1052&context=uclf

NOTE: Over the course of the research period leading up to this study's publication, UNESCO and ICFJ have published a research discussion paper (2021), a report presenting the findings of a global survey (2020), and two individual chapters extracted from this study: What More Can News Organisations Do to Combat Gendered Online Violence? (2022) and Assessing Big Tech's Response to Online Violence Against Women Journalists (2022).

UNESCO will publish two additional extracts - *The Chilling: Legal and Normative Frameworks for Combating Online Violence Against Women Journalists* and *The Chilling: Recommendations for Action Responding to Online Violence Against Women Journalists* in 2022.

35 Key Findings

▶ Incidence & manifestations of online violence against women journalists

1. Nearly three quarters (73%; n=456)[5] of the UNESCO-ICFJ survey respondents identifying as women said they had experienced online violence in the course of their work.

2. Threats of physical violence, including death threats, were identified by 25%[6] (n=179) of the women survey respondents, and sexual violence was identified by 18% (n=129). The women journalists interviewed were also plagued by such threats.

3. Threats of online violence against women journalists radiate: 13% (n=93) of women survey respondents and several interviewees described threats of violence against those close to them, including children and infants, as features of attacks.

4. Racism, religious bigotry, antisemitism, sectarianism, ableism,[7] homophobia and transphobia intersect with misogyny and sexism to produce significantly heightened exposure and deeper impacts for women experiencing multiple forms of discrimination concurrently.

5. Black, Indigenous, Jewish, Arab, Asian and lesbian women journalists participating in the survey and interviews experienced both the highest rates and most severe impacts of online violence.

6. Online violence against women journalists verges from large-scale attacks or extreme threats at a moment in time, through to the slow-burn of networked gaslighting,[8] which involves constant lower level abuse.

7. 15% (n=107) of women survey respondents and many interviewees reported experiencing image-based abuse (e.g., manipulated photos or video, stolen images, explicit images shared publicly without permission). 29 survey participants said they

---

5 Expressed as a percentage of 625 woman-identifying respondents who answered this question.
6 Unless otherwise indicated, all statistics presented in this study drawn from the 2020 UNESCO-ICFJ survey are based on a sample size of 714 woman-identifying respondents.
7 Discrimination in favour of able bodied people i.e. those who do not have disabilities.
8 **Gaslighting** is a means of manipulation that uses disinformation tactics designed to generate self doubt and ultimately discredit the target. See: https://www.nbcnews.com/better/health/what-gaslighting-how-do-you-know-if-it-s-happening-ncna890866

had been victims of 'deep fakes' (i.e. manipulated videos, often associated with fake porn, designed to damage reputations).

8. Almost half (48%; n=343) of the women journalists surveyed reported being harassed with unwanted private social media messages, highlighting that much online violence targeting women journalists occurs in the shadows of the internet, away from public view. Several of the interviewees also experienced offline stalking which had begun online via direct messaging.

▶ Impacts of online violence on women journalists

9. One in five (20%;[9] n=119) women-identifying survey respondents said they had been attacked or abused offline in connection with online violence they had experienced. A similar proportion of the interviewees also experienced offline harassment associated with online attacks, including the subjects of both big data case studies. These offline attacks ranged from stalking to physical assault and legal harassment.

10. Online violence against women journalists is integral to the prevalence of impunity for crimes against journalists, including murder.

11. Physical threats associated with online violence caused 13% (n=92) of women survey respondents to increase their offline security. This pattern was also evident among the interviewees, several of whom physically relocated multiple times in response to online attacks.

12. Mental health impacts were the most frequently identified consequence of online attacks, indicated by 26% (n= 186) of the women survey respondents. 12% (n=86) said they had sought medical or psychological help due to the effects of online violence, while a number of interviewees were suffering from Post Traumatic Stress Disorder (PTSD) connected to online attacks, and many more were receiving therapy as a result of the attacks they had endured.

13. When asked "How does the level of online violence you experience affect your journalism practice and your interaction with sources/audiences?", most frequently (30%; n=214) of the women journalists surveyed said that they self-censored on social media. 20% (n=143) described how they withdrew from all online interaction. Self-censorship was also a response noted by many interviewees, who frequently described this as a "chilling effect".

14. Employment and productivity impacts reported by the women journalists surveyed included missing work to recover from online violence (11%; n=79), making themselves less visible (38%; n=271), quitting their jobs (4%; n= 29), and even abandoning journalism altogether (2%; n=14). Linked to this was the professional discreditation of online violence targets. The interviewees confirmed this pattern.

▶ Sources and triggers for online violence

15. Nearly half (49% n=350) of the women journalists surveyed identified gender as the story theme most associated with online attacks, followed by politics and elections (44% n=314), and human rights and social policy (31%; n=221). This was also borne out through the interviews and case studies.

16. 41% (n=293) of survey respondents said they had been targeted in online attacks that appeared to be linked to orchestrated disinformation campaigns. This trend was underlined by many interviewees, and also linked to reporting on far-right extremism and conspiracy networks. Additionally, it was highlighted by both of the big data case studies.

17. The interviewees and case study subjects identified high-level political leaders and other State actors as some of the biggest instigators of online violence. Political actors were also the second most frequently identified sources (37%; n=264) of attacks and abuse after 'anonymous or unknown attackers' (57% n=407), according to the women survey respondents.

18. Partisan news outlets, media actors operating at the fringes of the political spectrum, and individual misogynistic journalists amplify and fuel attacks. This is a trend identifiable across the interviews and case studies.

19. Online violence against women journalists appears to have worsened significantly in the context of the COVID-19 pandemic.

▶ The role of the platforms as vectors of online violence

20. The big technology companies, often generically described as 'the platforms', are the main enablers of online violence against women journalists.

21. Facebook was the most frequently used (77% n=550) service for journalistic work (closely followed by Twitter on 74%; n=528) according to the women journalists surveyed, but it attracted disproportionately higher rates of incident-reporting to the platforms among the respondents: 39% (n=279) compared to Twitter's 26% (n=186). It was also rated the most dangerous of the top five services used, with nearly double the number of respondents rating Facebook "very unsafe" compared to Twitter.

22. Despite fledgling responsive efforts and stated commitments to enhancing journalists' safety on their platforms, the social media companies are still failing to stem the tide of online violence against women journalists, as illustrated through the global trends analysis, big data case studies, country case studies, and interviews conducted for this study.

23. In addition to technical design and business model failures that enable online violence on their platforms, social media companies lack comprehensive gender-sensitive/responsive, and human-focused solutions to the problem of online violence against women journalists. Despite having the capacity to do much more, they have inadequate on-the-ground representation in the countries where women journalists are most at risk, and they lack rapid response units for high-risk cases across the board.

24. These companies also lack capability to respond in all languages that their platforms serve, resulting in further neglect when it comes to dealing with women journalists abused and harassed in languages other than English.

25. There is uneven implementation and enforcement of platform policies across different States and regions.

26. Image-based abuse (e.g., photos, memes, videos) frequently flies under the platforms' radars.

27. Networked attacks often jump across different platforms, a pattern which in turn calls for networked, cross-platform responses.

The role of media employers as responders to online violence

28. Despite progress made by many employers over the past five years, only 25% (n=179) of the women survey respondents said they had reported online violence incidents to their employers, and among the responses they said they received when they did refer incidents to their employers were: no response (10%; n=71); and advice like "grow a thicker skin" or "toughen up" (9%; n=64). 14 (2%) respondents said they were asked what they did to provoke the attack.

29. Evidence of victim-blaming by employers is accompanied by the emergence of a pattern of restricting women journalists' speech as a response to managing online violence through social media policies introduced by news organisations, as illustrated by the testimonies of the interviewees and contextual research.

30. In some cases, women journalists reported being actively discouraged by their employers from speaking about their experiences of online violence or engaging with those attacking them. They are also sometimes told to avoid discussing 'controversial' topics on social media as a preventative measure. In the worst instances, there are cases of suspension or dismissal during a major attack.

31. Gaps remain in news industry responses, with many media organisations still lacking formal online violence response protocols and the integration of holistic strategies that blend digital security, physical safety, psychological support, upward referral mechanisms and gender-sensitive/responsive editorial policies.

Legal remedy and law enforcement

32. Only 11% (n=79) of the women journalists surveyed said they had reported instances of online violence to the police, and very few interviewees reported making a complaint to law enforcement agencies, highlighting a lack of confidence in policing, and prosecution processes.

33. Only 8% (n=57) of the women journalists surveyed and just a few of the interviewees had taken legal action, highlighting a reluctance among those targeted to pursue legal remedy due to various impediments, including a lack of digital investigative capability and gender-sensitive/responsive mechanisms.

34. Two interviewees[10] were pursuing legal action against State actors (foreign and domestic) at the time of writing. Several other journalists' abusers were the subject of ongoing police investigations or legal action.[11]

35. There are gaps in hate speech legislation which could otherwise be used to act against misogynistic online violence towards female journalists. In several countries, gender and/or sex-based discrimination are excluded from protections afforded to those suffering discrimination on the basis of race, religion and sexual orientation.

10 They are: Lebanese journalist Ghada Oueiss and Brazil's Patricia Campos Mello.
11 They include: Jovana Gligorijević in Serbia, and Patricia Devlin and Marianna Spring in the UK.

1

Introduction and methodology

Julie Posetti
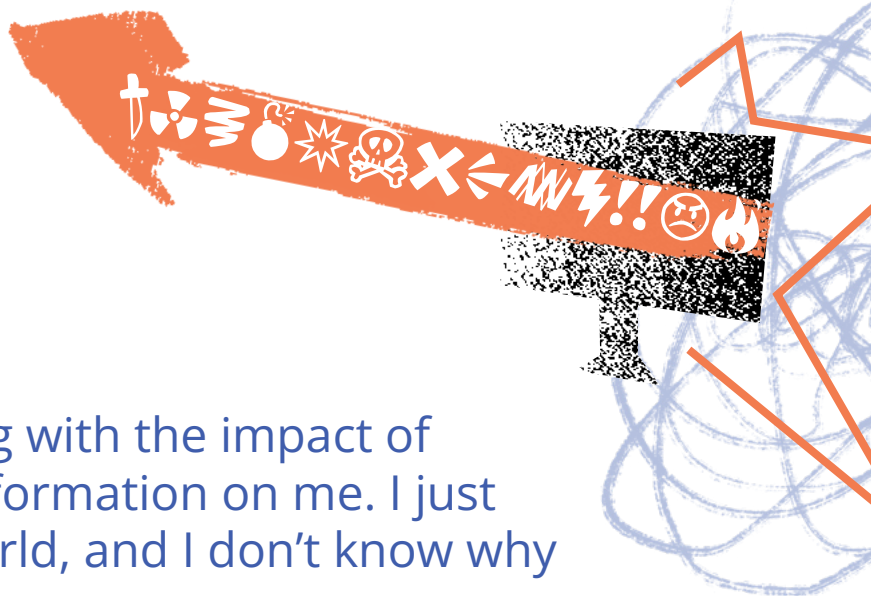
> "The easiest part is dealing with the impact of online violence and disinformation on me. I just see the impact on the world, and I don't know why we're not panicking."

Maria Ressa

There is nothing virtual about online violence. It has become journalism safety's new frontline - and women journalists sit at the epicentre of risk. Networked misogyny and digital gaslighting intersect with racism, religious bigotry, sectarianism, antisemitism, homophobia and other forms of discrimination to threaten women journalists - severely and disproportionately.

Gender-based online violence knows no boundaries. It is a global phenomenon that afflicts women journalists from Nairobi to New York, from Lagos to London, from Manila to Mexico City, from Belgrade to Beirut and Belfast. Around the world, the threats reach them via social media, email, chat apps, voice and text messages, and they are typically delivered by mobile devices. They are also frequently triggered or fuelled by political actors.

The threats radiate too - sometimes being extended to the women's children, partners and parents. The psychological, physical, professional, and digital safety and security impacts associated with this escalating freedom of expression and gender equality crisis are overlapping, converging and frequently inseparable. They are also increasingly spilling offline, sometimes with devastating consequences.

The research suggests that online violence against women journalists is designed to:

- **Belittle, humiliate, and shame;**
- **Induce fear, silence, and retreat;**
- **Discredit them professionally, undermining accountability journalism and trust in facts; and**
- **Chill their active participation (along with that of their sources, colleagues, and audiences) in public debate.**

This amounts to an attack on gender equality and press freedom and it harms the public's right to access information on a scale significant enough to trigger responses from the United Nations.

"Online violence against women journalists not only violates a woman's right to live free from violence and to participate in online discourse, but it also undermines the exercise of democracy and good governance, and as such creates a democratic deficit," the former UN Special Rapporteur On Violence Against Women reported to the UN Human Rights Council in 2020 (A/HRC/44/52) (UN GA, 2020a).[12]

Gender-based violence against women journalists online cannot afford to be normalised or tolerated as an inevitable aspect of contemporary audience-engaged digital journalism, nor online discourse more broadly. This is a view shared by UN Secretary-General António Guterres, who tweeted his concerns in March 2021, in connection with the research underpinning this study: "There should be no room for misogyny and violence in journalism. Social media platforms and governments must protect women journalists from online violence".[13]

In October 2021, the UN Special Rapporteur for Freedom of Expression presented a report to the Human Rights Council which applied a gender lens to freedom of expression issues and cited research published from this study. She placed significant emphasis on online violence against women journalists, noting that:

> *The harm caused by online violence, sexist hate speech and disinformation are real and diverse, affecting the mental and physical health of those targeted, undermining their confidence and autonomy, stigmatising them and generating fear, shame, and professional and reputational damage. In extreme cases, online threats can escalate to physical violence and even murder* (UN GA, 2021a).

One of this study's major goals was to identify effective methods for countering this threat to press freedom, journalists' safety, and women's active participation in journalism, while making recommendations to help overcome impediments to action.

However, while the researchers surfaced several examples of creative responses to gender-based online violence against women journalists, what they ultimately found was an escalating and rapidly morphing crisis, establishing that it is a global problem, but one with varied manifestations and uneven offline impacts, along with complex and intersecting challenges that inhibit effective responses.

The research findings lead to the conclusion that the epidemic of online violence against women journalists is systemic, and that multilayered structural change is required to address the toxic information ecosystem that sustains relentless digital abuse and attacks. Key among the obstacles to more effective responses is the enabling role of the largely unregulated social media companies - often generically referred to as 'the platforms' - and the role of political actors (including high office holders) as instigators and amplifiers of online violence against women journalists (see Chapter 2.3).

---

12 See the comprehensive bibliography accompanying *The Chilling* for details of all in-text citations, which can be found here: https://www.icfj.org/media/31615
13 Tweet from António Guterres from March 2021: https://twitter.com/antonioguterres/status/1375783912614490112?s=20

As this study demonstrates, online attacks on women journalists appear to be increasing significantly, particularly in the context of the 'shadow pandemic' of violence against women during COVID-19 (UN Women, 2020b). The emergence of the 'disinfodemic' (Posetti and Bontcheva, 2020), and its intertwining with conspiracy narratives, populist politics and far-right networks has also increased the toxicity of the online environments within which women journalists work. Thus, they are now more exposed to online violence than ever. They have become "sitting ducks" in the words of Michelle Stanistreet, General Secretary of the UK's National Union of Journalists.

Original, multidisciplinary, gap-filling research

Unique aspects of this research include its coverage of understudied developing countries, intersectional threats and impacts, and novel big data case studies.

The research goals included mapping the incidence, manifestations and impacts of online violence against women journalists internationally, noting commonalities and disparities as they appear in different geographical regions, and according to their intersectional exposure.

The research underpinning this study consists of:

- **A literature scan** covering over 1,000 scholarly and civil society research sources (Gober and Posetti, 2020).

- **A global survey of 901 (self-identified) journalists** from 125 countries conducted in five languages (Arabic, English, French, Portuguese and Spanish) fielded by UNESCO and ICFJ.

- **Long form interviews with 184 journalists,** editors, and experts in the fields of freedom of expression, human rights law, digital safety, and civil society responses. The interviewees span 17 countries and every region. They are ethnically and linguistically diverse. They also include a number of men (e.g., senior editorial leaders, digital security, legal and freedom of expression experts), people of different abilities and a range of sexualities and gender identities.[14]

- **Two big data case studies,** using Natural Language Processing (NLP) and network mapping techniques, which involved the assessment of over 2.5 million posts on Facebook and Twitter directed at two prominent journalists:

  - Maria Ressa, CEO and co-founder, Rappler (The Philippines);
  - Carole Cadwalladr, investigative journalist and columnist, *Observer* (UK).

184
Journalists, Editors and Experts interviewed

901
SURVEY RESPONDENTS

2 BIG DATA CASE STUDIES
ASSESSING OVER

2,5 MILLION
FACEBOOK AND TWITTER POSTS

---

14 All quotes featured in this study are drawn from original research interviews conducted by the researchers identified on the inside cover unless otherwise indicated.

- **Extensive research into:**

    - Manifestations, impacts, trends, and responses to gendered online violence in 15 countries that span the six UNESCO geographical regions[15]

    - News industry strategies/responses

    - Social media companies' policies and responses

    - Trends in legal and normative frameworks.[16]

i. Defining and describing online violence

In the same manner that domestic violence is understood to encompass physical, psychological, digital, and economic violence, online violence includes assault in the form of threats of rape and murder, wide-ranging acts of digitally-enabled harassment, economic and psychological abuse, along with multiple digital privacy and security threats. It is at its worst when it intersects with other forms of discrimination. And it is often bound up with disinformation. Additionally, this study has identified a dangerous trend that links online violence with offline attacks, harassment and abuse.

Former UN Special Rapporteur On Violence Against Women, Dubravka Šimonović, defined the term 'online violence' as extending to "...any act of gender-based violence against women that is committed, assisted or aggravated in part or fully by the use of ICT [Information and Communications Technologies]...because she is a woman, or [which] affects women disproportionately" (UN GA, 2018a).

The Council of Europe's Expert Group on Action against Violence against Women and Domestic Violence provided an additional explanation in a Recommendation published in November 2021 on the digital dimensions of violence against women:

> *...individual acts of violence that are not criminalised may reach the threshold of psychological violence when combined with the mob mentality and repetition facilitated by the internet: a teasing comment may ascend to cyberbullying when made repetitively or by a large number of people. Certain groups of women, including politicians, journalists and human rights defenders and activists are particularly exposed to such violent acts* (GREVIO, 2021).

Gender-based online violence can further be understood as individual components, or a combination of:

- **Misogynistic harassment, abuse and threats**
    This includes patterns of targeted, sexualised abuse and harassment, ranging from threats of violence (such as sexual assault, rape and murder) against the women journalists (and their children or other family members), through

---

15 Identified to cover a range of countries across all UNESCO regions, these are: Kenya, Nigeria, South Africa (Africa); Lebanon, Tunisia (Arab States); Pakistan, The Philippines, Sri Lanka (Asia and the Pacific); Poland, Serbia (Central and Eastern Europe); Brazil, Mexico (Latin America and the Caribbean); Sweden, United Kingdom of Great Britain and Northern Ireland, United States of America (Western Europe and North America).
16 See the detailed discussion of research methodologies at the end of this chapter

to gendered swearing and insults targeting their sex or gender, physical appearance, along with their sexuality, ethnicity, age and professionalism, which are designed to diminish confidence and tarnish reputations.

- **Digital privacy and security breaches that increase physical risks associated with online violence**

   Methods of attack designed to compromise women journalists' online privacy, security and safety include malware, surveillance, hacking, doxxing and spoofing.[17][18] They escalate the physical threats faced by women journalists because these acts can involve revealing their phone numbers and residential and work addresses, along with their patterns of movement. They can also expose identifying information endangering sources, colleagues and family members.

- **Coordinated disinformation campaigns leveraging misogyny and other forms of hate speech**

   Women journalists are frequent targets of digital disinformation campaigns (Posetti, 2018a), including orchestrated efforts with links to State actors (RSF, 2018e; Bradshaw and Howard, 2019; Bontcheva and Posetti, 2020; Posetti, Shabbir, Maynard, Bontcheva, and Aboulez, 2021; Di Meco and Wilfore, 2021b; Jankowicz et al., 2021; Posetti, Maynard & Bontcheva, 2021). Disinformation tactics, such as falsely accusing them of professional misconduct, spreading smears about their character designed to damage their personal reputations and professional standing, and malicious misrepresentation (e.g., deepfake[19] porn videos, abusive memes, manipulated images) are typical features of these attacks. The objective is to undermine the journalists' credibility, threaten them into retreat, and chill critical journalism.

Seven Features of Online Violence

Online violence targeting women journalists manifests itself in a variety of ways, but it has a number of common characteristics:

1. It is frequently networked

Online violence is often organised, coordinated or orchestrated. It can include State-sponsored 'sock puppet networks'[20] along with acts of 'patriotic trolling',[21] and involve mobs who seed hate campaigns within one fringe network (e.g., 4chan), before pushing it into more mainstream networks.

---

17 **Doxxing** is the process of retrieving, hacking and publishing other people's personally identifiable information such as names, addresses, phone numbers and credit card details in an environment that implies or encourages intimidation or threat (Techopedia, 2021a; McCully, 2019).

18 **Spoofing**, in general, is a fraudulent or malicious practice in which communication is sent from an unknown source disguised as a source known to the receiver. Spoofing is most prevalent in communication mechanisms that lack a high level of security (Techopedia, 2021b).

19 **Deepfakes** involve the sophisticated manipulation of audio, video or images designed to misrepresent the target. See: https://www.theguardian.com/technology/2020/jan/13/what-are-deepfakes-and-how-can-you-spot-them

20 The term **'sock puppet'** refers to a user account 'controlled by an individual (or puppetmaster) who controls at least one other user account,' often for 'malicious and deceptive' purposes, and 'to manipulate public opinion' (Kumar et al., 2017).

21 Appropriation of notions of national loyalty in order to discredit other actors as being "traitors".

2. It is generally misogynistic

Misogyny is the dominant feature of online violence targeting women journalists.

3. It radiates

Perpetrators of online violence often target women journalists' families, sources, audiences, colleagues and supportive online communities, too.

4. It is intimate

In detail and delivery, the threats are personal. They arrive on mobile phone screens first thing in the morning and last thing at night, in private spaces as well as at work, and they are often highly sexualised.

5. It can be extreme, intense and prolific

This often results in targets describing attacks with descriptors associated with extreme weather events, disease, natural disasters, and war, such as: "torrential", "epidemic", "tsunami", "flood", "avalanche", "a barrage", "trench warfare", "bombardment".

6. It can behave like 'networked gaslighting'

Constant moderate-low volume abuse and harassment that burns slowly but can be cumulatively devastating and undermine the target's confidence in her understanding of reality.

7. It is more extreme in the context of intersectional discrimination (e.g., race, sexual orientation, faith)

These factors appear to attract increased exposure and worse impacts.

What online violence is not

It is important to distinguish fair criticism from attacks, abuse, harassment and threats. Fair critique of a woman journalist's reporting does not constitute an act of online violence. All journalists need to accept reasoned, fact-based challenges to their reporting, including those that correct errors, suggest alternative sources, or provide diverse perspectives on issues. Robust debate is a feature of healthy democratic dialogue and it enriches journalism, including through collaborative fact-checking and reporting projects involving online communities.

The cross-cutting threat of gendered disinformation

Disinformation is now inextricably linked with gender-based online violence against women journalists (Posetti, 2018a; Jankowicz et al., 2021; Posetti, 2021). The weaponisation of misogyny in the global tilt towards populism and the rise of far-right extremism is also bound-up with gendered disinformation (Di Meco, 2020; Di Meco and Wilfore, 2021a).

This study highlights a threefold function of disinformation in gendered online violence:

1. *Disinformation tactics are routinely deployed in targeted multiplatform online attacks against women journalists;*

2. *Reporting on disinformation and intertwined issues, such as digital conspiracy networks, conflicts, and far right extremism, is a trigger for heightened attacks on women journalists;*

3. *Disinformation purveyors operationalise misogynistic abuse, harassment and threats against women journalists to undercut public trust in critical journalism and facts in general.*

ii. Research context

Gender-based online violence against women journalists is a pernicious problem - the study of which is slowly evolving as the incidence increases, and the methods of attack grow more sophisticated.

In parallel with this project, a detailed literature review and environmental scan conducted in 2020 spanning academic, industry and civil society research focused on the issue was undertaken (Gober and Posetti, 2020). The most notable finding from that process was the distinct absence of culturally and regionally diverse research into the problem.

Most published research prior to this study has focused on developed Western and Anglophone contexts, with a particular emphasis on the US. The most practically useful finding from the literature review and environmental scan was the conclusion that the internet was littered with out of date or defunct resources, initiatives and services designed to respond to the problem.

Here, to situate this study, is a brief overview of the pre-existing quantitative research conducted into gendered online violence, by intergovernmental organizations, civil society and industry bodies, as well as a summary of UN-level responses to the problem relevant to research in the area.

Quantitative research overview

- In 2014, a survey of 921 women journalists conducted by the International Women's Media Foundation (IWMF) and the International News Safety Institute (INSI) found that 23% of women respondents (n=212) had experienced 'intimidation, threats or abuse' online in relation to their work (Barton and Storm, 2014).

- The Organisation for Cooperation and Security in Europe (OSCE) surveyed 186 women journalists from Europe and North America in 2015, finding that most respondents had experienced online threats and abuse (OSCE, 2016a).

- In 2017, the Council of Europe (COE) published a survey of 940 journalists across 47 Member States, which found that 53% of the journalists polled had faced some form of cyber harassment, including personal attacks, public defamation and smear campaigns (Clark and Grech, 2017).

- Another survey conducted by IWMF in 2018, involving a smaller but still substantial sample (n=597) of mostly North American women, found that 63% (n=376) of respondents had been harassed or abused online at least once.

- Two years later, the UNESCO-ICFJ survey (2020) underpinning this study recorded that 73% (n=456) of women journalists surveyed at the global level who answered a specific question regarding incidence said they had experienced online violence in the course of their work.

- A 2019 survey of 115 female journalists in the US and Canada conducted by the Committee to Protect Journalists (CPJ) found that 90% of respondents in the US and 71% in Canada rated digital harassment as the most significant threat affecting them (Westcott, 2019).

- In a Canadian industry-wide survey in late 2021, 69% of women respondents (n=504) reported experiencing online threats and harassment over the past year (Ipsos, 2021). 78% (n=393) said that the frequency of online harassment had increased over the previous 2 years; 74% (n=373) said the onus for protecting the journalists from potential online harm rested with employers, while 69% said it rested with social media platforms.

2021 CANADIAN INDUSTRY SURVEY

Who is responsible for protecting women journalists from online violence?

74% SAID IT SHOULD BE **EMPLOYERS**

69% SAID IT SHOULD BE THE SOCIAL MEDIA PLATMFORMS

The COVID-19 pandemic changed journalists' working conditions, making them yet more dependent on digital communications services for 'socially distanced' reporting, and requiring them to conduct most audience engagement activity via social media channels. Another recent global survey conducted for the Journalism and the Pandemic Project by ICFJ and Columbia University's Tow Center for Digital Journalism (Posetti, Bell and Brown, 2020) found that 16% (n=490) of the women

journalists surveyed said the incidence of online harassment experienced during the pandemic was "much worse" than before COVID-19.

This was a finding echoed by a 2020 study from African Women in Media (AWiM, 2020), which found that almost two-thirds of more than 100 East African women journalists surveyed said online violence had increased in the context of COVID-19.

Although these surveys cannot be directly compared, when viewed collectively the pattern suggests that gendered online violence against women journalists has worsened significantly over the past decade, and in particular in the context of the pandemic.

UN context

This study was conceived by UNESCO amidst mounting concern at the UN-level about the incidence and impacts of online violence against women journalists around the world, and the need to develop more effective responses to the problem.

The conceptual origins of what this study describes as gender-based online violence were first explicitly explored by the UN in a 2006 study investigating violence against women more broadly. The then UN Secretary-General, Kofi Annan, recommended an investigation into online technologies so that emerging forms of violence could be recognised and addressed (UNSC, 2006). Gender sensitivity is incorporated in the UN Plan of Action on the Safety of Journalists and the Issue of Impunity, adopted by the UN in 2012, but it does not reference digital forms of violence.

The current UN Secretary-General, António Guterres, addressed the issue in a 2017 report to the UN General Assembly on 'The Safety of Journalists and the Issue of Impunity' (A 72/290):

> *Women who cover topics such as politics, law, economics, sport, women's rights, gender and feminism are particularly likely to become targets of online violence. While men journalists are also subject to abuse online, abuse directed against women journalists tends to be more severe* (UN GA, 2017a).

The same year, the UN General Assembly adopted a Resolution (A/C.3/72/L.35/Rev.1) on the safety of journalists with a particular gender focus. This Resolution built on other gender-specific resolutions from the UN Security Council in 2015 (UN SC, 2015), and the UN Human Rights Council in 2016 (UN HRC, 2016).

The UN General Assembly passed another Resolution on the Safety of Journalists (A/RES/74/157) in 2019 that further condemned unequivocally:

> *...the specific attacks on women journalists and media workers in relation to their work, such as gender-based discrimination and violence, including online and offline sexual harassment, intimidation and incitement to hatred against women journalists.*

iii. Key responders to gendered online violence

States are the main duty-bearers to protect journalists, with a responsibility to legislate accordingly and ensure law enforcement agencies respond appropriately. But social media are the main vectors for online violence against women journalists and they have a major role to play, while media employers are responsible for ensuring a safe working environment for the journalists who work for them.

Awareness of the gender-based online violence crisis has been rising slowly among internet companies, news organisations and States, which are gradually sharpening their responses.

Problematically, however, women journalists are still both the primary targets of online violence and the first responders to it. On the unmediated social media platforms they use in the course of their work, they are the ones required to 'report', 'block', 'mute', 'delete', and 'restrict' their attackers, potentially compounding the effects of the abuse, and creating unrealistic and sometimes unbearable pressures when the attacks come at scale.

Platform capture and neglect

Online violence against women journalists is a significant feature of what has been called 'platform capture' (Posetti, Simon and Shabbir, 2019b), which involves the weaponisation of social media by particular actors, in combination with the structural tendencies (and failures) of the platforms' business models and algorithmic design, as well as the phenomenon of platform-dependency and the enmeshment of news organisations and journalists, which has been exacerbated by the pandemic.

While making fledgling attempts to improve their products to enable easier abuse reporting and filtering, and engaging in regular public relations exercises designed to promote their commitment to creating safer platforms, the social media companies have so far failed in their efforts to respond effectively. Central to this failure are: tardiness; attempts to use a particular interpretation of 'free speech' as a shield against accountability; and a continuing reluctance to assume responsibility for content and behaviour on their sites in terms of international standards for protecting human rights. There have also been few attempts to hold these companies accountable through legal redress or regulation.

New challenges faced by media employers

Many media employers have made significant progress in addressing the problem of gendered abuse in the comments sections of their websites, and a growing number of news outlets have introduced an online violence response protocol. However, the journalism safety threats posed in the social media environment and the risks intersecting with disinformation appear to be either poorly understood, or too overwhelming to manage in many cases. There are also significant failings with regard to addressing the elevated intersectional risks and impacts associated

with gender-based online violence against women journalists, who concurrently experience other forms of discrimination, such as racism, religious bigotry and homophobia.

Additionally, too often employers respond to the problem and impacts of gendered online violence against their journalists by restricting women's speech (e.g., by introducing punitive social media policies that preclude them from responding to this issue) and victim-blaming (e.g., by suggesting a woman's speech was the cause of an attack).

Another point highlighted by this research is that most women journalists do not report or make public the online attacks they experience, in line with low levels of reporting when it comes to violence against women more broadly (UN GA, 2017c). As many interviewees also stated, media employers still often appear reluctant to take online violence seriously.

Finally, there is a vital need to ensure that news publishers also extend the protections and provisions offered to women journalists in their employ to freelance reporters who come under attack in the course of their work for the outlets concerned.

Ongoing impediments to effective legal and judicial responses

There are also gaps in law enforcement and judicial responses to online violence against women journalists. Primary among these are: the need to ensure that laws dealing with gender-based violence apply online as well as offline; the need to recognise misogyny as hate speech; the need for improved digital capability among police investigators; and the need for law enforcement, prosecutors and judges to be trained to investigate, prosecute and determine cases of online violence targeting women journalists.

See Chapter 4.0 for a detailed analysis of newsroom responses to gendered online violence, while Chapter 5.0 systematically critiques the platforms' responses, and Chapter 6.0 addresses the legal and normative frameworks for dealing with the problem at a global level.

Who are the main perpetrators?

The main perpetrators of online violence against women journalists range from misogynistic mobs seeking to silence women, through to State-linked disinformation networks aiming to undercut press freedom and chill critical journalism via orchestrated attacks. The role of political actors - including senior office holders, elected representatives, and party officials/members - in instigating and fuelling disinformation-laced online violence campaigns against women journalists is of particular concern. Additionally, partisan and fringe news outlets, and misogynistic individual journalists can be shown to amplify such attacks, triggering 'pile-ons' that escalate the risks of online violence morphing into physical assault, or causing significant psychological injury.

What more can States do?

While States sometimes condemn online violence against women journalists through their joint involvement in international organisations, there are few examples of national governments speaking out about domestic cases. On the other hand, a number of governments stand accused of not only failing to fulfil their responsibility to protect women journalists, but of being actively part of the problem. However, there are instances cited later in this study where gender-sensitive national strategies have been adopted to protect journalists and perpetrators of online crimes against them have been prosecuted.

Some political leaders have also spoken out about the need for the platforms to do more, and various legislative or regulatory initiatives are underway. But such interventions often contain implicit risks for freedom of expression which need to be carefully navigated. That is why this study provides a 25-point protocol to guide States wishing to respond to online violence against women journalists through legislative or regulatory measures while respecting freedom of expression rights (see Chapter 7.0).

iv. Methodologies and research approach

This study blends qualitative and quantitative methods to enable assessment of indicative incidence and scale, along with deeper assessment and analysis of the causes and responses associated with gender-based online violence as it affects women journalists around the world.

1. The survey

Between 24 September and 13 November 2020, ICFJ and UNESCO fielded a global survey about the incidence and impacts of online violence against women journalists, and responses to the problem (Posetti, Aboulez, Bontcheva, Harrison and Waisbord, 2020). The survey was offered in five languages: Arabic, English, French, Portuguese and Spanish.

The absence of common denominators for journalists at both the country and global levels ruled out demographically representative survey methods (e.g., randomised sampling). So, the non-probability method of 'purposive sampling' was adopted, with 'snowballing' techniques used to generate non-randomised responses (Patton, 1990) within the field of journalism internationally.[22]

To avoid illegitimate or inauthentic responses and ensure data integrity, the survey was distributed digitally via the closed networks of UNESCO and ICFJ, this project's research partners, civil society organisations that are focused on media development, journalism safety and gender equality, groups of professional journalists, and networks of editors. Given the non-probability, non-randomised nature of the survey, the results are not demographically representative nor

22 Purposive sampling involves the expert identification of relevant groups of research subjects for polling, while 'snowballing' is the term used to describe the process of research participants sharing the survey with other relevant respondents, thereby growing the sample size organically.

generalisable. However, they enable extrapolation of many patterns that may have wider applicability beyond the sample set.

After the removal of invalid responses, 901 self-selecting survey respondents remained, representing 125 countries. The data was then disaggregated along gender lines and a subset of **714 women-identifying respondents** was isolated for analysis. The descriptor 'journalist' is used generically throughout this report to refer to respondents who self-identified as journalists and media workers with a broad range of job descriptions.

Six of the top 10 countries (ranked by number of survey participants) were in Africa, Asia and Latin America, and the dominant great racial/ethnic group participants identified with was 'Hispanic or Latino'. The survey respondents were offered the opportunity to be identified, but most chose to remain anonymous.

Ethical considerations resulted in several questions being made optional. Some questions also allowed the selection of multiple responses to capture a multiplicity of experiences. As a result, the number of responses varies between questions and the results do not necessarily add up to 100%.[23] The survey also included a blend of qualitative and quantitative questions with opportunities for open text responses to elicit nuance and empower participants to share their experiences and challenges.

2.Qualitative interviews

To help deepen the research, identify emerging trends, and understand the impediments to more effectively addressing the problem of gender-based online violence against women journalists around the world, the researchers conducted 184 in-depth semi-structured interviews between July 2020 and June 2021. The interviewees were: journalists, editors and other media workers; civil society responders; and experts in the legal, technology and policy arenas. They were identified through the survey and desk research focused on the 15 countries selected for detailed study, along with the networks of researchers and partner organisations attached to the project.

These interviews were conducted face to face (where COVID-19 restrictions allowed), and via digital channels. The vast bulk were conducted synchronously, using phone, video or audio platforms (recognising the sensitive nature of the topic and the need for human engagement with potentially traumatised research subjects), and these were supplemented by email communications with some interviewees where necessary. The interviews were undertaken and overseen by the researchers identified on the inside cover of this study. The vast bulk of interviewees chose to be publicly identified.[24]

The interviews were subjected to critical thematic analysis and contextualised through additional desk research, allowing the surfacing of key themes and trends at the global, regional and country levels, and comparative analysis between countries and regions.

714

WOMEN-IDENTIFYING RESPONDENTS FROM

*125*COUNTRIES

---

23 All percentages presented have been rounded off to whole numbers.
24 All quotes contained in this study are drawn from original interviews conducted by members of the research team unless otherwise indicated. Quotes should be understood as contemporaneous with the timing of the interview, not the date of publication.

The global-level thematic analysis of the manifestation and impacts of online violence presented in this study emphasises the experiences of women journalists attacked on the basis of their sex and gender, and at the nexus of racism, ethnicity, religious bigotry, antisemitism, sectarianism, homophobia, transphobia, and other forms of discrimination. It also identifies **10 key global trends** (see Chapter 2.0) in gender-based online violence against women journalists.

3. *Fifteen country case studies*

This study is underpinned by 15 detailed country-based case studies produced by an international team of researchers, to provide essential context. Summaries of these country case studies are published in Appendix 1. The case studies were developed through wide ranging desk research and qualitative interviews with journalists and experts in the identified countries.

A note on research ethics

The University of Sheffield provided ethics clearance for the English language survey and interviews conducted in Western Europe and North America. This process also guided the conduct of research taking place in other regions.

Translations of the survey into other languages were conducted by UNESCO and reviewed by ICFJ.

The University of Sheffield also provided ethics clearance for quantitative data gathering and analysis associated with the big data case studies featured in this study.

2

Global Overview:
comparative trends analysis

Julie Posetti, Nabeelah Shabbir, Diana Maynard, and Nermine Aboulez

This chapter presents a global-level analysis of trends pertaining to the incidence, manifestations, impacts and trends associated with online violence against women journalists. It is based on the uniquely diverse data corpus underpinning this study. It emphasises the experiences of women journalists attacked based on their sex and gender, and at the nexus of other forms of discrimination. It also identifies **10 key international trends in the area of gender-based violence against women journalists online.** These trends emerged through a process of synthesising the qualitative and quantitative datasets and conducting a comparative thematic analysis.

The purpose of this chapter is to inform responses to online violence against women journalists by illustrating impacts and highlighting emerging challenges that demand attention. A more granular country-level analysis of these trends and gaps in redress was conducted for the 15 State-level case studies (summarised in Appendix 1) which emphasise previously understudied countries, and provide an essential lens through which to understand the diverse manifestations of gendered online violence around the world.

International research team members identified on the inside cover also contributed data (e.g., interviews and contextual research) and country-level analysis to this chapter.

Global thematic analysis: Key trends

1. Intersectional threats

Women journalists who are also disadvantaged by forms of discrimination that intersect with sexism and misogyny (e.g., racism, religious bigotry, sectarianism and homophobia) face additional exposure to online attacks, with worse impacts.

2. Online violence moving offline

There is increasing evidence that online violence against women journalists can spill offline with significant impact. This includes offline abuse and harassment that is seeded online, as well as physical attacks. There is also evidence of a reverse trajectory, involving political actors, partisan media outlets or other antagonists engaging in targeted attacks on women journalists offline, which spur online 'pile-ons' and create an environment for the enabling and reinforcement of judicial harassment.

3. Disinformation as a multi-pronged and intersecting threat

The weaponisation of false and misleading content functions both as a method of attack (e.g., the deployment of disinformation tactics in misogynistic 'pile-ons') and a lightning rod for attacks (i.e. reporters covering disinformation and associated beats). Simultaneously, orchestrated disinformation campaigns operationalise gendered online violence to chill critical reporting.

4. The COVID-19 factor

The problem escalated during the pandemic, linked to a combination of a) the rise of conspiracy narratives associated with COVID-19 denialism and anti-lockdown/anti-vaccination movements, b) increased exposure to toxic online communities necessitated by reliance on digitally-mediated reporting and audience engagement practices due to social distancing, and c) political crackdowns on journalists scrutinising government responses to the pandemic.

5. The role of political extremism, nationalism and populism

Misogyny is weaponised (along with networks of misogynists) in populist mobilisation and women journalists are clear targets - particularly those reporting on far-right extremist networks.

6. Political actors as perpetrators

Political actors, including presidents and party leaders, are implicated as major instigators and amplifiers of online violence against women journalists.

7. Instrumentalisation of partisan media outlets

The role of fringe/partisan media in instigating and fuelling online violence against women journalists is noteworthy.

8. Social media as vectors of online violence

Social media platforms are the major vectors for online violence against women journalists, who, in turn, also see these companies as failed responders to the problem.

9. News organisations are still struggling to respond

Many news organizations' responses to online violence against women journalists remain ineffective. There is a need to place attention on contextual threat modelling and response protocols, in particular at the nexus of disinformation and diverse forms of discrimination.

10. Gaps in legal and judicial responses, and obstacles hindering their implementation

These include the need to ensure that laws dealing with gender-based violence offline apply equally online; the need for hate speech legislation to cover sex and gender-based violence (online and offline); and the need for law enforcement officers, public prosecutors and judges to be educated so they can respond more effectively to the problem.

Trends 8-10 above are covered in depth in Chapters 4, 5 and 6 of this study.

2.1 Types of online violence and their manifestations

According to the 2020 ICFJ-UNESCO survey, online violence against women journalists is extremely prevalent internationally. Almost three quarters (73%; n=456)[25] of respondents identifying as women said they had experienced online violence in connection with their work in the field of journalism. This result reflects a problem that appears to have significantly worsened over the past decade.

Participants in this research described experiencing online violence across every social media service used for digital reporting: from Facebook, Twitter, Instagram, YouTube, WhatsApp, and email, through to Telegram, TikTok and ClubHouse[26] (see Chapter 5.0 for a detailed discussion). The women journalists surveyed and interviewed said they had been subjected to a wide range of online attacks, from threats of violence, through to cyberstalking, digital security attacks and privacy breaches.

This enabled the production of a taxonomy of gendered online violence against women journalists.

12 main types and methods of attack evident at the global level

1. *Threats of sexual assault and physical violence, including death threats (some of which were extremely graphic e.g., including images of dead bodies). These were generally received via open social media platforms, closed messaging, email, and text messages. Some journalists also received death threats that came via analogue means after being doxxed.*

2. *Threats to harm family members, including young children.*

3. *Hateful and sexually explicit abusive language (especially misogynistic, sexist, racist and homophobic abuse).*

4. *Harassing private messages (including text messages and the closed messaging systems of social media services, along with encrypted chat apps) which often involved soliciting for sex.*

5. *Threats to damage professional or personal reputations. These include reputational smears designed to undermine public confidence in their journalism, or hold them up for censure within their communities.*

6. *Digital privacy and security-based attacks, including hacking, doxxing, interception and surveillance.*

---

25 Expressed as a percentage of the 625 women journalists who responded to the question.
26 Clubhouse is an invitation-only social audio app that is a hybridisation of talkback radio and networked eavesdropping.

7. *Misrepresentation via spoof accounts and manipulated images, video and audio (including deepfakes). Such synthetic[27] media are shared openly via social media and on fringe sites that facilitate coordinated attacks, or via email and encrypted chat apps.*

8. *Harassment in the form of fraudulent online orders being sent to women journalists' homes after doxxing, and other financial attacks like credit card fraud.*

9. *Manipulated search results on YouTube and Google which were flooded with hateful and/or disinformation-laden content designed to malign and discredit the targets, drowning out their professional journalistic content.*

10. *Coordinated 'pile-ons'/ 'brigading'/ 'dogpiling',[28] often linked to "dog whistling",[29] involve not just fringe networks of misogynists and disinformation agents, but also publicly identifiable political actors and male journalists. This is a key feature of the 'networked gaslighting' of women journalists discussed in this study.*

11. *Orchestrated attacks involving State actors (including foreign State actors).*

12. *Body shaming and associated sexist and racist critique of women journalists' appearance.*

i. Networked misogyny

Threats of physical and sexual violence plagued the women journalists interviewed and surveyed for this study. **A quarter (25%; n=179) of the survey respondents had received online threats of physical violence, including death threats.** 18% (n=129) of participants experienced threats of sexual violence, and many interviewees also testified about experiencing chilling threats of this nature.

In the US, White House correspondent April Ryan received a bomb threat from a man who was ultimately imprisoned for targeting CNN journalists.[30] She has also experienced multiple other online death threats, including one warning that she would be shot from overhead, and would know what was coming: "I had one person tell me that they were a former Army Ranger - quote unquote, 'death from above'".

Another example came from Lebanese journalist Ghada Oueiss,[31] a principal presenter at Al Jazeera, who was portrayed as a prostitute by her attackers. She described receiving threatening emails after being doxxed:

1/4

OF THE WOMEN JOURNALISTS SURVEYED

RECEIVED ONLINE THREATS OF PHYSICAL VIOLENCE

---

27 **Synthetic media** is the collective name for tools such as 'deepfake' videos, as well as authentic-looking fake photos and writing, and which spread disinformation (Bateman, 2020).

28 **'Brigading'** is synonymous with the term 'piling-on'. It refers to the function of online mobs called to attack an individual. **'Dogpiling'** is a sexist term which describes a method of loosely organising targeted attacks against women online.

29 **Dog whistling** can be defined as a call to respond or attack that is understood only by the target audience.

30 United States Department of Justice (2018) press release - "Cesar Sayoc Pleads Guilty to 65 Felonies for Mailing 16 Improvised Explosive Devices in Connection with October 2018 Domestic Terrorist Attack": https://www.justice.gov/opa/pr/cesar-sayoc-pleads-guilty-65-felonies-mailing-16-improvised-explosive-devices-connection

31 Oueiss is currently pursuing legal action in the US against various officials and individuals, including foreign State political actors, in relation to the coordinated online attacks she experiences.

One of them that I can never forget [said]: 'You will be looking at the camera to talk to your audience and you will start reading the bulletin and reading the autocue in front of you. You will notice that there is a gun and [a] bullet, that bullet will go straight to your head.' Then I started getting emailed pornographic pictures... they put my head on naked women. And then, they made another email in my name and they started sending to my colleagues pictures of [my] head on a naked body, also porn pictures.

Classic misogynistic tropes in evidence globally include 'witch', 'hag', 'whore', 'bitch', and the pernicious hashtag #presstitute, which implies both sexual and professional misconduct.

The branding of women journalists as 'witches' is associated with some of the worst examples of online violence recorded in this study, including the harassment of assassinated Maltese journalist Daphne Caruana Galizia prior to her death. Media freedom lawyer Caoilfhionn Gallagher KC, who represents Caruana Galizia's family, pointed out that this kind of comparison aims to "dehumanise" the journalist, and make them "fair game".

Many other journalists interviewed for this study also described incidents in which they had been portrayed as witches in online attacks. For example, Mexican reporter Ana Luz Solís was Photoshopped as the fictional witch Hermelinda Linda and NBC-MSNBC disinformation reporter Brandy Zadrozny was also subjected to abuse deploying the 'witch' trope in the US.[32]

In South Africa investigative journalist Pauli Van Wyk was called a 'witch' and a 'white bitch', and threatened with extreme sexual violence, causing humiliation and shame.

I was told that I should be 'shot in the pussy', 'hanged and burned'. They talked about 'witch-burning'. At the start, you think 'This cannot be real. This cannot be happening to me.' But then it turns into a strange feeling of unworthiness.

Such slurs are frequently associated with networked misogyny, which can be defined as a virulent strain of orchestrated gender-based hostility and violence in online environments (Banet, Weiser and Miltner, 2016). Attackers are often self-appointed vigilantes who loosely organise with others and aggregate their attacks via hashtags to silence women journalists (Waisbord, 2020).

But there are also signs that such attacks are increasingly orchestrated or coordinated. For example, swarms of abusers often attack when political actors, high-profile misogynists or partisan journalists issue cues to 'brigade' or 'dogpile' an individual woman journalist. Similar dynamics were in evidence during 'Gamergate' in 2014 - an online harassment campaign against female game developers and

32 In Mexico, 'Hermelinda Linda' is a fictional witch (with warts and exaggerated features) depicted in a comic book series from the 1970s and then in films from the 1980s.

a feminist media critic - that eventually also targeted other women who came to defend them (Dewey, 2014).

This pattern was particularly evident in research about the US situation conducted for this study. The language echoes the anti-press discourse of populist and far-right politicians. It also mirrors classic misogynistic tropes, and the kind of attitudes associated with offline violence against women, including digital manifestations of gaslighting.

One example of this phenomenon involved triggering a campaign of online violence against the former *New York Times'* technology reporter Taylor Lorenz. On International Women's Day in March 2021, Lorenz published a Twitter thread about her own experiences of online violence, in which she also shared a link to research about orchestrated attacks against Filipino-American journalist Maria Ressa that was produced in parallel with this study (Posetti, Maynard and Bontcheva, 2021). Lorenz tweeted: "The scope of attacks has been unimaginable. There's no escape" (Lorenz, 2021). In response, she was trolled at scale, ridiculed, humiliated and mocked for talking about the serious impacts of online violence on her own mental health.[33]

Journalists and commentators from the far-right to the far-left triggered a further pile-on against Lorenz by pseudonymous and anonymous trolls. Fox News' Tucker Carlson broadcast a segment mocking Lorenz, thereby amplifying and prolonging the attack against her (Sullivan, 2021). It lasted for several days, and the lead author of this study was also targeted by some of the antagonists, demonstrating the ways in which online violence radiates.

Figure 1: Screengrabs of tweets published by Taylor Lorenz in March 2021 discussing her experiences of online violence and the ways in which she was targeted in response.

Coordinated gender-based violence can be seeded within misogynistic fora like Kiwi Farms before moving to mainstream social media.[34] A US-based journalist,

---

33 Lorenz auto-deletes her tweets, but in an email exchange she confirmed she tweeted this.
34 Kiwi Farms is a far-right internet forum which was founded out of its 'parent board' on 8chan by Joshua Moon in 2013. It is known for facilitating online abuse/ its active targeting and harassment of trans people.

who asked not to be identified, was targeted in what she described as her "most frightening" online experience during one Kiwi-Farms-generated pile-on:

> *It's kind of like a 4chan,[35] but exclusively for people whose hobby is harassing people that they don't like online. They call them 'LOL cows'. They often select trans women or women they perceive to be vulnerable and will just constantly harass them...I believe that in some cases they have harassed people to suicide.*

The journalist was subjected to a two-week campaign of abuse after Kiwi Farms members compiled a dossier on her and her immediate family members. "Very quickly and very effectively...using online clues...[they] figured out who all of my siblings were and what their jobs were." Kiwi Farms' targets are often subjected to doxxing and offline stalking (Pless, 2016). "They were connecting dots in a way that made me quite scared that they were going to go after my family," the journalist said.

During research for this study, Kiwi Farms also lit up with hateful discussion about another US journalist - Taylor Lorenz - in the context of the attacks described above.
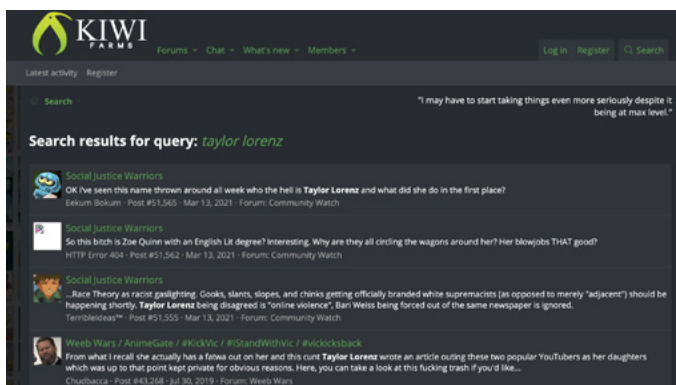

Figure 2: Screengrab of Kiwi Farms search results (March 22, 2021) for the name 'Taylor Lorenz'.

In 2021, the Coalition For Women In Journalism (CFWIJ, 2021b) identified 87 major trolling campaigns against women journalists globally - a 22.5% increase compared to their 2020 data.

The individual big data case studies analysing the experiences of Carole Cadwalladr (UK) and Maria Ressa (The Philippines), featured in Chapter 3 also illustrate the relentless ferocity and intensity of such misogynistic and coordinated abuse, along with the manifestation of networked gaslighting as a form of online violence.

In Tunisia women journalists interviewed explained the political dimensions of orchestrated campaigns intended to intimidate them out of covering 'controversial' stories. Journalist Najoua Hammami has come under attack from political actors:

---

35 4chan, founded in 2003, is an online space on the internet which is home to over 70 'boards', such as /pol/ ('politically incorrect'). The site has cultivated a community of 'free-extremists' with its dedication to 'extreme free speech', and provides a home to the hacker collective Anonymous. It has more recently been charged with more radicalised language cultivating the alt-right. Other such sites include 8chan, Gab.ai and parts of Reddit (Colley and Moore, 2020).

When they want to get to you psychologically, they start a campaign against you on social media from day to night... They block you from commenting – especially when they find out that you are a journalist – on their pages, then they start the harassment, the insults...

Coordinating abuse in the shadows

Almost half (48%; n=343) of the women journalists surveyed reported being harassed - often sexually - with unwanted private messages. This highlights the fact that much online violence targeting women journalists occurs in the shadows of the internet, away from public view. These attacks tend to go hand-in-hand with networked pile-ons in open online spaces.

Direct one-to-one communication via private messaging may appear to be perpetrated by an individual, but it may also reflect an organised and networked effort at domination and intimidation. The impact of online violence on the targeted woman journalist via private messaging can be as severe as being targeted in public spaces, especially when the attacks are coupled with public facing social media-based abuse.

48%
*OF SURVEYED WOMEN
JOURNALISTS*
REPORTED BEING HARASSED
WITH PRIVATE MESSAGES
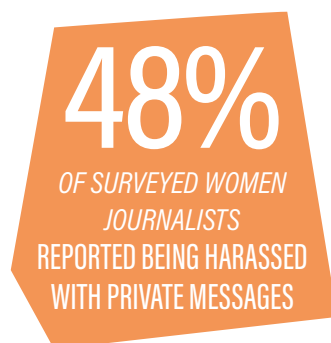
ii. Online threats that radiate

The online violence experienced by the women journalists surveyed and interviewed sometimes radiated to family members, including children, along with bystanders who rallied to their defence. Sources and colleagues were also targeted in parallel. **13% (n=93) of the survey respondents said that in the midst of attacks, they had also received threats of violence against those close to them.** These radiating threats not only place women journalists' professional associates and personal connections in danger, they can also have the effect of further chilling their participation in online communities.

Al Jazeera's White House correspondent, Kimberly Halkett - who was sexually assaulted while reporting on International Women's Day in 2008 - was targeted in a brutal campaign of online violence in 2020 which resulted in her abusers tracking down her teenage daughter and threatening her as well:

They wanted to kill me. They were going to come after my family. My address was posted all over the internet. But the part that really upset me, and that rattled me to my core - they went after my 15-year-old daughter.... They found a post, where I talked about her... and I had tagged her, and so they began to blow up her Instagram, saying all kinds of horrible things about her, about me... My daughter had done nothing except be my daughter.

Similarly, Patricia Devlin[36] - then a reporter with *Sunday World* in Northern Ireland - received a rape threat against her newborn child in October 2019, which came via Facebook Messenger to her personal account: "It was absolutely traumatising… I have never felt disgust and fear like it. I went to the police. That was the turning point for me. I said, I can no longer brush it off. It's getting worse. My children are being threatened".

Dozens of women journalists from the BBC Persian language service have experienced prolific coordinated online violence over the past decade, and in two cases children of the women were also targeted. In the first instance, a picture of one of the presenters was Photoshopped onto a pornographic image, which was then sent to her 14-year-old son at his school. In the second case, a threat was made against the son of another BBC Persian language journalist and also delivered to his school.

In June 2021, one of those targeted, Rana Rahimpour, spoke at a UN Human Rights Council event about the hybrid online-offline threats to her and her family. "They have threatened to cut [our] throats. And I have also been reminded that my parents are in Iran and they are within reach" (Doughty Street Chambers, 2021). The BBC further reported that a journalist for the service had received death threats via Skype after her 27-year-old sister had been arrested in Iran. The journalist was reportedly told her sister would be released if she agreed to spy on her colleagues or leave her job (BBC, 2021a).

According to ARTICLE 19's Itzia Miravete, the online violence which Mexican women journalists experience is highly sexualised and it radiates to family members: "their bodies [are] objectified and disapproved of, the work they do [is] discredited… Their families are also attacked, making women journalists' experience twice as difficult".

The triple bind of 'family' as a cultural value

In some conservative cultures, 'family' as a traditional value (Fargues, 2005) is weaponised so that the targeted women journalists face a triple bind: they want to prove that they can work as independent professionals outside the family; yet they need to protect themselves as they practise journalism in part for the sake of their families; and finally, they need to protect their families from the radiating impacts of the harassment they are facing. Several interviewees in Asia and the Arab States also described family pressures to withdraw from public view for their own 'protection', or to avoid 'shame' and 'dishonour', following threats that radiated to others.

This pattern of radiating violence is exploited by those targeting women journalists online. In Lebanon, journalist Myra Abdallah was sent photos of her sister and friends on Facebook documenting their whereabouts. The purpose: to let her know that her family was being watched.

Journalists who have spoken up in Tunisia have also become targets for online attacks which spread to their family members, particularly when they do not

13% OF THE WOMEN JOURNALISTS SURVEYED

HAD RECEIVED THREATS AIMED AT PEOPLE CLOSE TO THEM

---
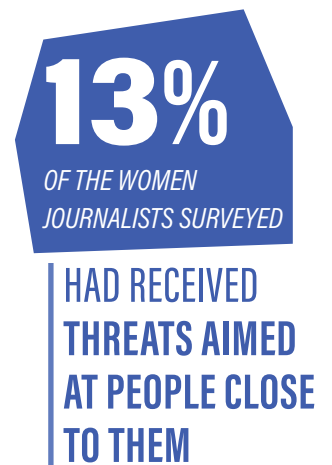
36 See a case study about Devlin's attempt to get Facebook to act against the accounts involved in Chapter 5.0

respond to, or comment publicly about, the abuse levelled against them. In the case of Najoua Hammami her family was targeted with "obscene pictures of genitals and money"; it harmed her relationships. "When this happens to me, I don't respond, I just delete them, but when it goes to your mother, brother and sister…They would send my sister threats that they will kill me if I don't step back. So, my brother wanted me to quit my job."

The attackers want family members of a targeted woman journalist to persuade her to avoid particular stories or sources as a protective measure, according to Tunisian editor Khaoula Boukrim who said this had been her experience. Both she and Hammami told the researchers that they were scared that their mothers could be hurt as a form of retaliation against them. This study surfaced similar challenges for women journalists working in some Asian countries, including Pakistan and Sri Lanka.

iii. Sexualised and sexist abuse

The most commonly reported online threat experienced by the research participants came in the form of abuse laced with hateful language designed to denigrate them. Lilian Okonkwo-Ogabu, journalist and Chairperson of the Nigeria Association of Women Journalists (NAWOJ), explained that women journalists in highly patriarchal contexts are often targeted in ways in which their gender, and sexist perceptions of their secondary social status, underpins the abuse:

> *Women are not to be heard and [their] opinions are less valued. So, when a woman makes strong statements…whether on social or traditional media, she's vulnerable…because it's believed she lacks the traditional authority … it's not expected that any responsible woman should be doing such jobs. I've had a case of some persons who asked my husband about my posts and appearance on Facebook, wondering why he allows me such freedom.*

Kenyan women journalists interviewed described the driving ideology behind the anonymous and pseudonymous trolls who target them. In addition to structural sexism that manifests itself in highly sexualised attacks, Kenyan interviewees spoke of a "false and pernicious" narrative about the "plight of the boy-child", with the perception that women's equality has advanced to a degree where men have been left behind. The interviewees saw this as part of an anti-feminist backlash.

South Africa's Qaanitah Hunter left her job as investigative political reporter at the *Sunday Times* after being the "target of press statements from political parties, and rape threats: you won't find similar harassment of males in terms of rape and body shaming", she said in an interview. Hunter was commended by the South African media industry for speaking out about the depression and anxiety that she suffered as a result of her intimidation by political actors connected to the ruling ANC party, and being body-shamed. This was a thread connecting many interviewees' experiences.

Body shaming as a feature of online violence

In Kenya, a former TV news anchor said that pictures of her during pregnancy were circulated on social media with a view to body-shaming her: "The experience was traumatising because I was participating in a television show and all the attention online became about my appearance and weight gain. As someone in the public eye, you are already self-conscious, then that coupled with the body changes that come with being pregnant, it is all too much." Her self esteem was severely damaged, with significant career impacts, and it led to her decision to quit her job. "There was no apology…that could help me recover from that. I knew I had to leave, so I turned in my resignation before my child was born," she said.

"They make memes, they make fun of my physique because I'm overweight," said Gladys Rodríguez Navarro, a Mexican freelance journalist who has been targeted with fake Facebook profiles for her critical political reporting since 2015.

The cumulative nature of these attacks is also important to note. US columnist and author Lindy West has been prolifically body-shamed with some men telling her they would "…gladly rape me if I weren't so fat" (West, 2017). She also noted the impacts over time:

> *And so each one of those tweets individually you could look at and be like, 'Oh, that's just some idiot trying to scare you.' But taken cumulatively — and even in combination with more innocuous tweets, just people insulting you or trying to hurt your feelings or trying to waste your time — it's very, very, very emotionally draining (Gross, 2017).*

West ultimately quit Twitter in response to the online violence she experienced.

In March 2020, Fadoia Chtourou was trolled on Facebook for not wearing makeup while reporting on a terrorist attack for the Tunisian public broadcaster - so much so that civil society organisations, citizens and fellow journalists came to her defence. Chtourou's experience of online harassment – for coming on television, "looking like that!" - was publicised by the National Journalist's Syndicate (Mosaique FM, 2020). Fellow journalists and citizens launched 'no makeup days', organised via the hashtags #FadoiaChtourou (فدوى شطورو#) #WithoutMakeup and #AgainstYourStandardsAlwaysForever, to denounce the misogynistic bullying campaign against her (Darwish, 2020; Jouini, 2020).

Chtourou told the researchers: "This bullying act shows that there is a problem with our minds: the perception that the woman is just a body, an image. There is a stereotype that women have to be fully 'dolled up'. It needs to be socially addressed!". Meanwhile, fellow Tunisian Malak al-Bakari was abused online for reporting while wearing a short dress (Jouini, 2020).

42%
OF SURVEYED WOMEN EXPERIENCED **REPUTATIONAL ATTACKS**

Inflicting reputational damage

Professional and reputational threats constituted a significant proportion of incidents reported by the women journalists surveyed by ICFJ and UNESCO. 42% (n=300) had been targeted in reputational attacks, such as abuse designed to demean their intelligence, mental capacity, ethics or morality. And 23% (n=164) identified professional threats, including false allegations about misconduct drawn to the attention of their employer, along with spurious claims about the unreliability of their journalism.

Image-based abuse is also used to smear the reputations of women journalists. This includes manipulated photos and videos, stolen images, and sexually explicit images shared publicly without permission. Image-based abuse was reported by 15% (n=107) of the women-identifying survey respondents at the global level. One example, shared by Tunisian journalists' union lawyer Ayoub El-Ghadamsy, involved a woman radio host being subjected to disinformation-laced smears. In an attempt to trigger a misogynistic pile-on against the presenter of the radio programme همس الليل (Whisper of the Night) people used her name and photo to share obscene stories and referenced threats of sexual violence.

In Lebanon Youmna Fawaz, an investigative journalist who has been the target of online harassment for 10 years, was subjected to sexualised trolling by a fellow journalist who often bullies her, she said. While she was live streaming from a beach to report on a terrorist attack in 2020, he took a screenshot to show "the slightest hint of bikini bottoms" and then circulated them to question her morals and damage her reputation.

Deepfakes and other forms of synthetic media attack had also been experienced by 29 survey participants. A further 36 reported 'shallow fakes' (i.e. decontextualised videos or images, such as the misrepresentation of a crime scene) as a technique used to target them. Yosra Frawes from the Association Tunisienne des Femmes Démocrates (Tunisian Association of Democratic Women), an independent feminist and human rights organisation, said that smear campaigns, digital privacy breaches and acts of blackmail are designed to "wipe women out of public spaces".

iv. Digital security, surveillance and privacy-based attacks

Many of the women survey respondents reported exposure to: surveillance 18% (n= 129); hacking 14% (n=100); doxxing 8% (n=57) and spoofing 7% (n=50). This heightens their susceptibility to physical violence because of the public identification of personal details, including physical addresses, as well as risks to their sources, especially at-risk confidential sources (Posetti, 2017a), and to their family members.

Among the interviewees for this study, dozens also said they had been doxxed, with their addresses, emails and phone numbers being exposed. In the US, this has occurred on extremist websites, and sites designed to facilitate the networked trolling of women. Doxxing has also been a tactic for online violence in Kenya

TYPES OF HARASSMENT

18%
OF SURVEYED WOMEN EXPERIENCED
Exposure to surveillance

14%
OF SURVEYED WOMEN EXPERIENCED
Hacking

8%
OF SURVEYED WOMEN EXPERIENCED
Doxxing

7%
OF SURVEYED WOMEN EXPERIENCED
Spoofing

(Ngigi, 2020), where journalists Caroline Mutoko and Julie Gichuru were targeted with invasive comments about their sexuality, ethnicity and physical appearance after their private details were exposed online (ARTICLE 19, 2016).

Another Kenyan journalist said in an interview that when she came to the defence of a woman being harassed "for how she was dressed", she was doxxed and called a prostitute. In Brazil, Thais Arbex of *Folha de São Paulo* experienced doxxing and harassment after reporting on Brazilian President Jair Bolsonaro and his ex-wife (Arbex and Uribe, 2019).

South African political editor Karima Brown[37] briefed a reporter before a national election in May 2019, but mistakenly sent the editorial brief to the WhatsApp group of a populist political party called the Economic Freedom Fighters (EFF). The party leader posted a screenshot of the message containing Brown's cell phone number on his Twitter account, which then had three million followers, alleging Brown was trying to send a "spy" to EFF's event. Within hours, Brown was subjected to rape and death threats (Maphanga, 2019). Graphic sexualised images tagging Brown were posted on social media, particularly Twitter, but abusive messages were also sent through voice and WhatsApp messages.

Serbia's Jovana Gligorijević, a journalist at *Vreme*, was also doxxed - in the comments section of a YouTube video showing a 'men's rights activist' being removed from a feminist conference. Her home address and personal identification number were published, among other details which she believes could not have been obtained without access to a State-run registry that is not publicly available. Gligorijević reported the incident to Serbia's Prosecutor for High-Tech Crimes, who she said then suggested to her that she rely on Google support for assistance (Google owns YouTube).

A number of interviewees - in the US, Sri Lanka, the UK and South Africa - said they had to move house, region, or even country as a result of such threats connected to doxxing (see detailed discussion in Chapter 2.5 about the links between online violence and offline attacks).

Targeted surveillance as online violence

Mexico has an extensive history of surveillance abuses. Research has shown that government agencies regularly abused NSO Group's Pegasus spyware[38] to surveil reporters such as Griselda Triana, a journalist and widow of Javier Valdéz, the assassinated veteran journalist who founded the magazine *Ríodoce*. In 2017, she received text messages containing disinformation and malware links just days after her husband was murdered, a killing condemned by the international community (Grillo, 2018). It was a period during which she was cooperating with a police investigation and speaking out about the assassination (Scott-Railton et al., 2019).

---

37 Brown died from COVID-19 complications in 2021.
38 Forensic Architecture, an academic unit at Goldsmiths University, has created a bespoke platform to map this digital violence in collaboration with Amnesty International and The Citizen Lab: https://www.digitalviolence.org/#/ . See Chapter 2.3 (Top sources and triggers of online violence against women journalists) for more about how the Pegasus surveillance software was installed on the phones of over 180 journalists.

One of the most emblematic cases of digital violence in Latin America involves Mexico's Carmen Aristegui (Editorial Director of Aristegui Noticias and a CNN presenter) who was also the target of digital espionage in 2017. "It activates your own phone, it activates your own camera, it activates your own microphone," she said about the Pegasus malware, which has reportedly been used in more than 50 countries worldwide to infect the phones of activists, journalists and human rights defenders, amongst others (Weizman et al., 2021; Forbidden Stories, 2021).

One European freelance journalist told the researchers about being hacked in an episode that affected both her mobile phone and her laptop, resulting in URLs associated with a foreign State actor appearing randomly on her devices. Another journalist from the Arab States and one from Asia said they were sent photos taken from inside their homes as a warning that they were being watched.

In Lebanon, journalism veteran Diana Moukalled has been a victim of phishing and spoofing attacks, involving an email from an imposter disguised as a friend. When the email did not open, she asked the sender if their email was working and got the reply: "Just as we are sure that you're Shi'ite". Within a second, she said, she lost control of all her social media accounts, indicating a malware attack. This case also demonstrates the function of sectarian attacks as an aspect of gendered online violence against journalists.

v. Economic abuse

Sixty-four survey respondents (9%) reported experiencing digitally-enabled financial threats, such as attempts to extort money or assets, threats to expose spending behaviours, hacking into bank accounts, and misrepresentations made to financial institutions. Interviewees from Brazil, Poland, the UK and the US told the researchers that doxxing attacks had resulted in unwelcome orders being delivered to journalists' homes, including food and sex toys, as well as manure and other offensive material. In one case, a gas canister was sent, and one interviewee reported a "transvestite" prostitute being dispatched to her.

Sometimes payment on delivery of fraudulently ordered items was demanded. "This is a strange invasion of privacy - it's a veiled threat; it makes you feel insecure if a meal arrives at your address," said Ana Freitas, a freelance Brazilian journalist who was doxxed after publishing an article about misogyny within the local gaming community.[39] She also received packages of worms. "It has an immediate chilling effect. I keep telling myself... 'ah, but this is a bunch of kids, there's no courage... they just stay on the internet talking shit'".

Freitas also received threats in a month-long harassment campaign and was forced to leave her house for several weeks, fearing attacks against her family. The cost of relocating - even temporarily - for safety reasons can be costly and several interviewees described having to bear the financial burden themselves.

Other manifestations of economic abuse identified by the research participants include targeting employers (e.g., tagging them in posts designed to discredit the journalist) with a view to triggering the woman's dismissal, suspension or

9%

OF THE WOMEN JOURNALISTS SURVEYED

EXPERIENCED DIGITALLY-ENABLED FINANCIAL THREATS

---
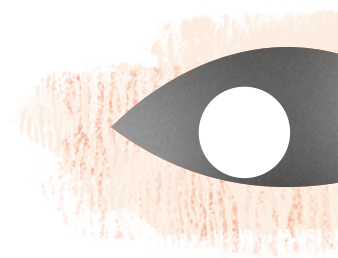
39 HuffPost Brazil closed in 2020, so there is no access to the original article: https://techcrunch.com/2020/11/24/huffpost-india-brasil-shutdown/

demotion. For example, Nigerian documentarian Ruona Meyer was targeted in a campaign involving the tagging of her employer after the BBC broadcast her award-winning investigation into Nigeria's cough syrup cartels.[40]

Other examples of this tactic to draw the attention of media employers , with the chilling effect of potential unemployment, include US journalists Amy Wilder at AP, Lauren Wolfe and Quinn Norton at the *New York Times*, and Felicia Somnez at the *Washington Post.*[41] Guardian US's Julia Carrie Wong also made the point that in the US, when a woman journalist loses her job in the context of an online violence campaign she is frequently exposed to further financial hardship through the loss of fundamental protections, such as health insurance. In the US access to adequate healthcare is extremely limited for the uninsured.

40 See further details in Chapter 2.3 (Top sources and triggers of online violence against women journalists).
41 See Chapter 4.0 and the United States country case study summary Appendix 1.

# 2.2 At the intersection of misogyny and other forms of discrimination

Research for this study highlights the ways in which other forms of discrimination - such as racism, religious bigotry, homophobia and transphobia - intersect with sexism and misogyny to worsen and deepen women journalists' experiences of online violence. Other researchers have also found that such women are more likely to be subjected to intersectional digital harassment (Knight, 2018; Glitch-EVAW 2020; Jankowicz et al., 2021).[42]

The women journalists surveyed for this study clearly demonstrated this pattern. The incidence rates for minorities and additionally marginalised ethnic communities of women were significantly higher than the 64% (n=129) of white women journalists who indicated that they had experienced online violence in the course of their work (compared to 73% overall). For example, 81% (n=55) of respondents identifying as black women, 86% (n=12) of Indigenous women, 77% (n=156) of Latino women, and 88% (n=7) of Jewish women survey participants reported experiencing online violence.

Arab-identifying women respondents were also significantly more likely to experience offline attacks, harassment, and abuse associated with online violence than other ethnic groups. Over half (53%; n=21) of the women identifying as Arab[43] said they had experienced offline attacks that they believed were seeded online, compared to 11% (n=22) for white women and 20% (n=118) overall (see a detailed discussion of the online-offline trajectory in Chapter 2.5).

GENDERED HARASSMENT AT THE INTERSECTION OF RACE AND ETHNICITY

The incidence of online violence among the women journalists surveyed

| WHITE	LATINO	BLACK	INDIGENOUS	JEWISH
61%	77%	81%	86%	88%
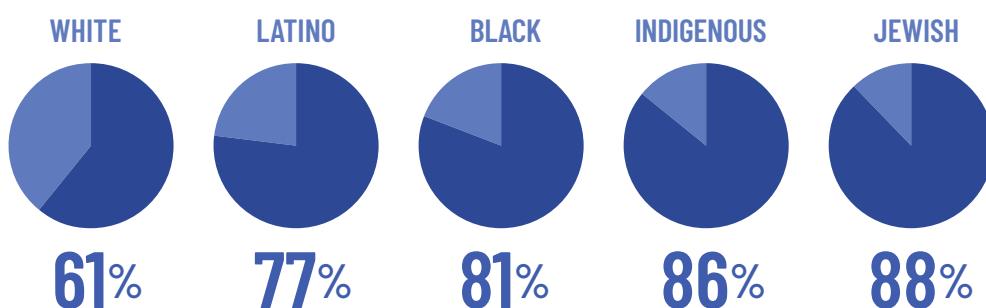
Data derived from a 2020 ICFJ-UNESCO survey of 714 women journalists

42 For a discussion of intersectional feminism, see: https://www.unwomen.org/en/news/stories/2020/6/explainer-intersectional-feminism-what-it-means-and-why-it-matters

43 These participants identified as ethnically Arab but they were not necessarily located in the Arab States.

SEXUAL ORIENTATION AND HARASSMENT

How sexual orientation is interwined with gendered online violence experienced by the women journalists surveyed

HETEROSEXUAL   LESBIAN   BISEXUAL

72%   **85**%   **88**%
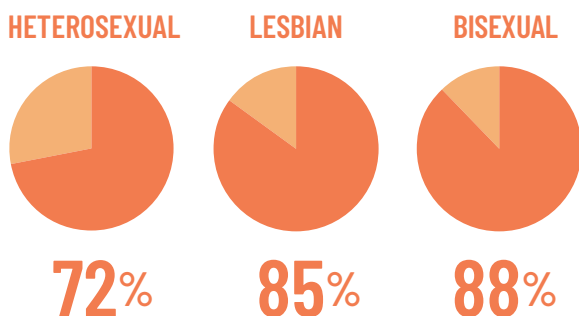
Data derived from a 2020 ICFJ-UNESCO survey of 714 women journalists

A similar pattern can be seen when analysing the survey data through a sexual orientation lens: while 72% (n=357) of heterosexual women indicated they had been targeted in online attacks, the rates of exposure for those identifying as lesbian and bisexual women were much higher - standing at 88% (n=7) and 85% (n=34) respectively (Posetti et al., 2021b).

This pattern of heightened online violence risks faced by women journalists at the intersection of multiple forms of discrimination was confirmed by at least 30 interviewees from around the world, highlighting the functions of racist, antisemitic, xenophobic, religiously bigoted, and sectarian abuse.

i. Racial vilification and structural racism

Women journalists from diverse ethnic and racial groups participating in this study described experiencing an additional layer of attack because they are "low hanging fruit" or "easy targets".

Two black women journalists interviewed in the UK demonstrated how structural racism is amplified through online violence. Both said that being called the 'N-word' online is normalised. Rianna Croxford, an award-winning BBC Investigations reporter, described the repetitive pattern of racist abuse: "It's not the first time somebody called me the N-word. It won't be the last time. I've had racial abuse... 'monkey', mocking my appearance in different ways".

This was a form of abuse also familiar to *The Independent*'s Race Reporter, Nadine White. In June 2020, a tweet from journalist Leona O'Neill, targeted in general for her reporting in Northern Ireland, noted: "If abuse I have received merely for wanting to speak to black people is anything to go by, racism & intolerance is indeed rife".[44]

---

44 This tweet from June 2020 has since been deleted: https://twitter.com/leonaoneill1/status/1268844093762867200?lang=de

In Brazil, eight out of 10 girls and young women have suffered from online harassment, according to a survey conducted by NGO Plan International.[45] Brazilian women journalists interviewed for this study said they experienced discrimination and abuse in ways that are dependent on various intersectional aspects of their identity.

Those who identified as black also reported online attacks such as "monkeys" or "black cows". Journalist Gabi Coelho, formerly of Voz das Comunidades, said: "I was not attacked just for expressing myself…I was also attacked for being black". Broadcast journalist Rita Batista received similar abuse in response to her election reporting. As a result, took an identifiable Bolsonaro-supporting online troll to court and won (Redação Correio 24 Horas, 2019).

In Mexico, colleagues of investigative reporter Kennia Velázquez at PopLab were victims of racist or derogatory slurs, including being branded as "sell-outs to the drug cartels, fat, short, Indigenous", she said. Another Mexican journalist, Reyna Haydée Ramírez, said she is often called "Indian" and "ugly" online. She said there were "worse insults than those - not to mention the death threats - all because I questioned the Mexican president".

South African editor Ferial Haffajee was maligned and vilified in an orchestrated online violence campaign that began in 2016, in response to a corruption investigation she undertook when she was Editor-at-Large for the local edition of *Huffington Post* (Haffajee, 2017). The attacks on Haffajee leveraged "raw racism and misogyny", she said. She was accused of being a journalist working for 'White Monopoly Capital', an attempt to discredit her that was particularly potent in post-apartheid South Africa.

Haffajee experienced taunts that she had never experienced before, such as "Go back to India, this is not your country". Her compatriot Pauli Van Wyk was also racially vilified - for being white - in a brutal campaign of highly misogynistic online violence connected to her investigative reporting.

Kenyan women journalists said that they were insulted online for their ethnicity as well as on the basis of their sex, marital status, and reproductive capacity. They also described facing greater intersectional discrimination for not demonstrating shared ethnically aligned political views, or purportedly traditional tribal perspectives.

In the US, Julia Carrie Wong said she is specifically targeted because she is a Chinese-Jewish woman. "I think it's coming because of what I'm doing, which is trying to expose white nationalists who think that I don't deserve to live in this country and that my racial makeup is disgusting," she said. However, racist online abuse is also not the exclusive preserve of the political right in the US. *Washington Post* White House correspondent Seung Min Kim was racially abused by apparently left-leaning trolls in 2021.[46]

Technology reporter Sarah Jeong, who writes for the *New York Times*, Verge and MotherBoard, has described the ways in which ethnically and racially diverse

---

45 14,000 girls and young women aged 15 to 25 dealing with harassment on social networks were surveyed in 22 countries, including Brazil (Martins et al., 2020): https://www.internetlab.org.br/en/inequalities-and-identities/violence-against-women-online-and-the-courts-preliminary-observations/
46 The lefttist trolls had allegedly once sent online abuse to Alaska's moderate-Republican Senator Lisa Murkowski. See: https://www.thedailybeast.com/liberal-twitter-mob-heaps-racist-sexist-abuse-on-wapo-reporter-seung-min-kim-for-doing-her-job?via=twitter_page

women journalists are expressly and baldly targeted in the US context because they are visible. "People frequently send me hateful messages using their real names and sometimes even their work emails. It is just a part of my daily life, as it is for many women of colour with a visible online presence," she said (DeVito, 2020).

The *New York Times*' Apoorva Mandavilli pointed out the disproportionate levels of hostility she experiences compared to white male colleagues reporting on the same beat: "It seems like white male journalists get more of a pass and are treated with respect.[47] And I've noticed that with other women as well, and certainly the other women of colour. I was just absolutely raked over the coals."[48]

In Europe, VICE UK Editor-in-Chief Zing Tsjeng faces increasing abuse for being 'Chinese'[49] since the pandemic began. "As an Asian woman...I'm perceived as confrontational and not compliant, submissive... That really, really irritates [online abusers]. It adds a dimension to the abuse that I find really uncomfortable... It's quite aggressive."

Greek-born Swedish journalist Alexandra Pascalidou described being targeted by neo-Nazis and abused with a wide array of racist and bigoted slurs: "I've been called a dirty whore, a bloody Gypsy, Jewish, Muslim slut, a Greek parasite, a disgusting migrant, stupid psycho, an ugly liar, a biased hater. They keep telling me to go home, to kill myself or they will shoot me, cut my tongue off, break my fingers one by one. They keep threatening me with gang rapes and sexual torture" (Posetti, 2016).

ii. Antisemitism and far-right extremism

Racialised misogynistic online abuse against women journalists coming from the far-right and other groups is a globally experienced phenomenon, and many of the interviewees linked it to elected officials creating the enabling environment for this pattern to become entrenched. The broader role of so-called 'patriotic trolling' and the weaponisation of social media platforms to advance far-right ideologies has also been explored by other researchers (e.g., Bradshaw & Howard, 2017; Nyst and Monaco, 2018; Beirich and Via, 2021; Jankowicz et al., 2021).

Guardian US' Julia Carrie Wong has been the subject of multiple attacks orchestrated by far-right groups which abused her with antisemitic slurs. In an early manifestation of this trend, she "...caught the attention of the Daily Stormer[50] back when [it] was on the surface web and had a pretty significant audience," she said.

Being Chinese, being Jewish, being a woman just becomes part of the content of the harassment... That became a fixation for the writers and commenters who then proceeded to Photoshop me with horns and a Jewish star, and discussed my racial makeup with quite intense and disgusting fervour.

In 2017, one hour after her story about a Discord server[51] being used to coordinate the alt-right and white nationalists involved in the 'Unite the Right' riots was published (Wong, 2017), Wong experienced what she described as "threatening" and "horrifying" online harassment: "The [Discord Server] had changed all of the infrastructure to use my name and my photo... Constant jokes about the Holocaust, jokes about the Shoah...my name and picture was being used as the welcome bot... They [had] me in this digital format, saying really disgusting antisemitic jokes about the Holocaust."

Crister Ohlsson, head of security at Bonnier News, a privately held Swedish media group which owns the newspaper *Expressen* (among others), said much of the harassment against journalists working for their company is based on antisemitic abuse emanating from the racist right. "Our Editor-in-Chief is Jewish, the Bonnier family [is] Jewish. So that's one dimension of harassment from the extreme right-wing. But while they are not that many, they make a lot of noise", he said.

In a 2017 study on xenophobia and radical nationalism online, the Swedish Defence Research Agency (FOI) found that the prevalence of xenophobic and nationalist groups was widespread in the Swedish digital environment (Kaati et al., 2017), with hostility and prejudice expressed towards women and ethnic minorities, as well as public figures, journalists and politicians (ibid).

Similarly, Reporters Without Borders (RSF) Sweden has reported that the rise of political populism and white supremacy movements has resulted in a growing online presence and surge of alternative media with a clear far-right political agenda. A survey of Swedish Union of Journalists' members in 2016 confirmed this trend, with 43% (n=366) of respondents who said they were exposed to online violence identifying their perpetrators as belonging to a right-wing extremist and/ or racist group (Löfgren Nilsson, 2017).

In Serbia, journalist Jovana Gligorijević was the subject of a 28-minute video by a young YouTube influencer and two representatives of far-right organisation Leviatan (Leviathan) who accused her of being "the main source of Serbia's downfall" (Vasić, 2019).

A US journalist who wished to remain anonymous was among a group of 15 journalists who cover the far-right featured in a video presenting "threatening imagery...of mass shooters intercut with images of the reporters" in April 2020. The video was reportedly produced by a person connected to the neo-Nazi Atomwaffen Division group (Ross, Bevensee, and ZC, 2019).[52] "It looked like a kill list," the journalist said, which could be read as: "nobody would be mad if you took this into your own hands".

---

51 Discord is a video game chat application (which rebranded as a more mainstream conversation portal). A Discord server is a channel where users congregate to discuss a particular video game or topic (Wong, 2017).
52 The YouTube video was circulated via Telegram in April 2020. It followed a viral YouTube video titled 'Sunset the Media' created by an extreme right group who adulate a white supremacist who shot and killed nine African Americans in Charleston, South Carolina, in 2015, and share his 'bowl' haircut (Holt, 2019).

In the UK, the divisive 2016 EU referendum known as Brexit, and manifestations of the 'culture wars', have also fanned nationalism associated with online abuse against women journalists, especially those reporting on topics related to identity and colonialism. Vice UK Editor-in-Chief Zing Tsjeng said: "I'm [attacked as] a journalist who's talking about the downside of colonialism." Two other UK interviewees said they received antisemitic abuse and threats based on the perception of their surnames as Jewish, despite the fact that were not actually being Jewish.

iii. Sectarianism

Online violence based on religious and associated cultural factors amplifies the risks to women journalists in countries where perceptions of immorality or blasphemy can trigger backlash and even offline attacks.

In Pakistan, for example, blasphemy can attract the death penalty. Amber Shamsi, a freelance reporter formerly with the BBC's Urdu language service in Pakistan, has reported an increase in the targeting of women journalists from the Shia Muslim sect. When Shamsi published a story about Qandeel Baloch,[53] a social media influencer who was murdered by her brother in a so-called 'honour killing', Shamsi and BBC Urdu "were accused of having 'nothing better to cover' for giving space to a 'slut' who was disgracing the country".

Shamsi noted the irony in being abused for reporting about a woman who had been attacked online and, fatally, offline. She was sent pictures of daggers when she reported on sectarian violence in Pakistan. "The fear in Pakistan always is that the digital hate could quickly become very real," she said (BBC, 2016).

Yusra Jabeen was accused of blasphemy and 'defaming Islam' when she shared a interview in a Facebook group that she had conducted with a veteran feminist activist in Pakistan, which discussed the fact that women heroes are missing from Pakistan's school text books. In the aftermath, a troll contacted her editor, and Jabeen's colleagues asked her to take down the post as "these accusations can get you traced, killed or kidnapped...It could have affected my workplace and everything," she said in an interview.

Pakistani women journalists are often forced to steer away from covering hard issues such as politics, war, terrorism and religion (Jamil, 2020; Kamran, 2019; Khan, 2020). Aima Khosa, a Punjabi journalist at *The Friday Times,* is often asked why she is writing about Pashtuns, or Sindis, or Baloch, and she finds herself misrepresented in online attacks like this: "Then you must be an 'NGO Auntie' who is a paid prostitute for all these suppressed nationalities...they want to cause divisions when there shouldn't be [any]."

This pattern manifests itself in the US too. A US-based reporter who asked to be anonymous received online abuse for interviewing local Muslim groups in what she called a "straightforward" local news story. "I tweeted this story and I looked online, and I had hundreds and hundreds and hundreds of replies about [being] a terrorist, or look at this Muslim writing the story about Muslims, or promoting Islam or Sharia law", she said.

Yusra Jabeen
Picture credit: Anwar Zeb

---

53 Qandeel Baloch's 2016 murder has been linked to caste dynamics in Baloch society in Pakistan (Human Rights Watch, 2019; Abdul Aziz Patel, 2020).

Similar stories came from women journalists in Sri Lanka from minority backgrounds, and those who report on religion and other "taboo" subjects. One Sri Lankan editor interviewed described a typical reputational attack where a post she made on Facebook about a contentious political issue was copied by the page administrator, edited so that it appeared that she had insulted the Lord Buddha, and re-posted. Others then created a 'mud poster', soliciting revulsion for this religious offence, which was re-shared thousands of times on different pages.[54]

In Northern Ireland, online violence against women journalists also has a sectarian aspect. Journalist Patricia Devlin reports on families affected by the ongoing criminal activities of paramilitary groups in the country.[55] She has recently watched "usual levels" of misogynistic abuse become sectarian. "Smear campaigns, orchestrated harassment campaigns, high levels of anonymous Twitter accounts abusing me… I was being accused of only writing about loyalist paramilitaries because I was perceived as coming from a nationalist background, even though I also write about republican paramilitaries." These online attacks became a daily ordeal for Devlin.

In Lebanon, sectarianism and caste are overlapping intersectional elements of abuse. Luna Safwan, a freelance journalist and journalism safety trainer who is a Shi'ite Muslim, said she was subjected to smear campaigns because she is critical of Hezbollah, and she is falsely accused of being an Israeli agent. "I was attacked online, like other famous journalists who are also Shi'ite, facing campaigns and death threats, because we were perceived to be against Shi'a, we were perceived to be traitors to the religion," she said in an interview.

In a similar case, Lebanese TV presenter Dima Sadek said that being a Shi'ite herself and publicly criticising the Shi'ite Hezbollah has made her the target of many smear campaigns, with trending hashtags such as "ديما الواطية" or ("Dima the vile"). Her daughter, who has a disability, has been targeted too. "If you weren't such a slut, your daughter wouldn't have turned out to be a handicap (sic)," and "your daughter is wondering who her father is", are just two examples of the sort of abuse she is exposed to online.

These examples highlight the 'triple whammy' effect of intersectionality: female journalists are targeted as women; for doing critical journalism; and for not conforming to sectarian prejudices.

iv. Gender identity and sexual orientation

In many countries gender equality is not mandated in law. In some, commitments to the rights of women and girls are being wound back, in others and governments condemn LGBTQ identities (Ash, 2020; Amiel, 2021). These approaches further intensify the digital risks facing many women journalists.

CNN International reporter Tara John experienced a three-day wave of abuse after publishing a story on trans rights focused on Scotland's Gender Recognition Act

54 In Sri Lanka 'mud poster' is shorthand for mud-slinging poster, a form of political attack used extensively throughout south Asia. Originally the term referred to street posters, but now encompasses memes and other derogatory digital visual messages.
55 A Northern Ireland peace agreement was signed in the late 1990s.

laws. The harassment became so intense that John felt afraid to leave her home at the beginning of the COVID-19 pandemic:

> *The moment the piece went out, it exploded. And this was around the start of the lockdown... I'm not really too fussed when people start trolling me... but this time it felt very personal. I had, for about three days, people scrutinising my tweets. One group threatened to sue me. Others told me to fuck off; many encouraged pile-ons and others called me a men's rights activist, a moron, and it just did not stop... Perhaps the intensity of it was, for me as a woman, that these were also women.*

Sabahat Zakariya, a freelance journalist and YouTube vlogger[56] points to the intersectional abuse faced by women journalists in Pakistan, where sexism and misogyny blend with homophobic vitriol, and where the word 'khusra',[57] a slur for trans, is used against women journalists interviewed for this study.

> *One thing I get a lot is [abuse] about my short hair. They write comments like, 'oh, you're khusri', or something either homophobic or transphobic. You know, they try to find a way to get to you or something that will make you feel insecure about yourself.*

There is also evidence of homophobic abuse in the big data case study analysing attacks on Filipino-American journalist Maria Ressa in Chapter 3.1. There, these slurs intersect with extremely racist tropes and particularly vile expressions of misogyny.

Mini case Study: **Ghada Oueiss, Al Jazeera**

> *I'm an Arab, so I'm not supposed to ask tough questions. I am also a journalist, which is bad enough in itself. And to top it off, I'm a woman.*

Lebanese journalist Ghada Oueiss, a principal Al Jazeera Arabic presenter based in Qatar, has been the target of a prolific online violence campaign which she said started with the Arab Spring in 2011. It continued with her reporting on the brutal murder of her friend, US-based Saudi journalist Jamal Khashoggi, in 2018 inside the Saudi Arabian consulate in Istanbul (Warwick, 2020; Shilad, 2021b; CPJ, 2021e).

---

56 Zakariya is founder of Feminustani, a YouTube channel on feminism in Pakistan: https://www.youtube.com/channel/UC9izag27NtE1DIaXg1i9VhQ

57 Khusri (female) or khusra (male) is an Urdu language word for eunuchs or trans-people. It is a very socially derogatory word and objectifies a person, in the Pakistani social context, to someone who is not accepted by society because they either look different or defy the gender they are born with. These are mostly outcast, and not accepted into the education, social or professional sectors.

This coordinated digital campaign against Oueiss is expressed through intersectional abuse focused on her gender, religion and age. She has been portrayed as a prostitute by her attackers, and in one instance she received 40,000 abusive tweets in just a few hours, she said in an interview.

In one episode, stolen photos of her wearing a swimsuit were shared using the hashtag #Ghada_Jacuzzi (Vered, 2020); and she also received chilling death threats while live on air.

Oueiss is one of our research participants who identified political actors as top sources of online violence. "If I was working with CNN or MSNBC or NBC or ABC or even BBC, they [her attackers] wouldn't dare try to kill my character... But when it comes to an Arab media outlet, it is more irritating for the regimes in this part of the world."

Oueiss' harassment is coloured by rivalries and perceived loyalties in the region:

> *I was attacked by the Israeli government because I covered the war on Gaza... They said she is not a journalist. 'She is the spokesperson for Hamas, which is [labelled] a terrorist group'. I was attacked by the Syrian regime because I covered the demonstrations and the revolution in 2011... Then I was attacked by the Bahraini regime because I covered the demonstrations, the revolution in 2011. And then I was attacked by Emiratis because of the blockade against Qatar... They wanted Al Jazeera to shut down. And when it didn't shut down, they started to harass the journalists who work for Al Jazeera.*

Oueiss expressed serious concern that the online violence she experiences will spill offline and result in physical violence. She recalled a conversation with Jamal Khashoggi shortly before he was brutally assassinated:

Ghada Oueiss
Picture credit: GO & Ebrahim Serhan

> *When he saw how I was attacked, and because he was also attacked by the same accounts, he sent me a message saying: 'Don't be upset, ignore and block them'. And I said, 'I don't think I should ignore and block, I think I should tell the world about it'. And then, one month later, they killed him in the [Saudi Arabia] consulate [in Turkey]. So he was attacked online and then he was attacked physically. What if I'm attacked now online and then I will be attacked physically? So, I chose to talk about it in case something happens to me so that the world knows who did it.*

In December 2020, Oueiss initiated legal action in a US court against a number of high profile international political actors and their associates including the crown princes of both Saudi Arabia and the United Arab Emirates, for allegedly spearheading coordinated online attacks against her (Shilad, 2021a). The same month, a federal court in Florida issued a summons to the de facto leaders of Saudi Arabia and the UAE. The case was ongoing when this study was being finalised in mid-2022. Credible safety threats continue to prevent her from travelling outside Qatar.

2.3 Top sources and triggers of online violence against women journalists

i. States and political actors as perpetrators

The correlation between populist politics, misogynistic narratives and the demonisation of journalists and journalism, exhibits an expansion of politically-motivated attacks on journalists and independent journalism around the world. Political actors - including politicians, government officials, political party representatives, party members, political operatives - frequently act as instigators and primary perpetrators of online violence against women journalists.

This is an alarming trend confirmed by the survey respondents, interview participants and big data case studies featured here. For example, over a third (37%) of the survey respondents identified political actors as top sources of online attacks against them.

Prime targets of such attacks among the interviewees include: Lebanese journalist Ghada Oueiss; Brazilian journalist Patricia Campos Mello; multiple journalists in Mexico,[58] Poland, Pakistan and Serbia; Maria Ressa in the Philippines; Kimberly Halkett, April Ryan, Julia Carrie Wong, and Brandy Zadrozny in the US; Carole Cadwalladr, Rianna Croxford, Laura Kuenssberg and Nadine White in the UK; and Ferial Haffajee and Pauli Van Wyk in South Africa.

As noted by the 2021 Joint Declaration on Politicians and Public Officials and Freedom of Expression by special rapporteurs of the UN, OSCE, OAS and the African Commission on Human and People's Rights, there is a

37%
OF THE WOMEN
JOURNALISTS SURVEYED

SAID **POLITICAL ACTORS** WERE TOP SOURCES OF ONLINE VIOLENCE

> *...growing incidence of online and offline 'hate speech', disinformation and dangerous rhetoric against and scapegoating of the media, human rights defenders and groups at risk of discrimination, including by politicians and public officials, which chills freedom of expression, thereby reducing the diversity of information and ideas in society and misleading citizens (Khan et al., 2021).*

This research confirms that online violence against women journalists also comes from foreign State actors, and it is increasingly associated with State-linked legal harassment (OHCHR, 2021b). Examples of State-linked campaigns of online violence against women journalists surfaced through this research include the

---

58 One of the most prominent cases in Mexico is that of journalist Lydia Cacho, who also experienced online abuse after being abducted by Mexican police officers (El Universal, 2019), and who was tortured, and illegally jailed. Identified as a site of entrenched impunity by the Committee to Protect Journalists' Global Impunity Index (2021), the country has seen at least a prosecution of a businessman and the former state governor of Puebla (Justice in Mexico, 2020). In her book, *The Demons of Eden: The Power that Protects Child Pornography* (2005), Cacho had exposed a conspiracy to protect a paedophile ring involving several prominent businessmen and politicians.
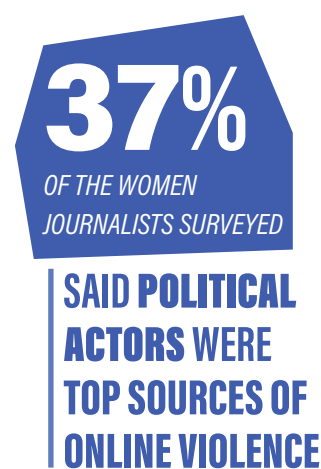
cases of Maria Ressa (the Philippines),[59] Al Jazeera's Ghada Oueiss (Lebanon), Ferial Haffajee (South Africa), Jessikka Aro (Finland), and in the UK the cases of multiple BBC Persian language service reporters.

The attacks on the BBC Persian journalists are suspected of being connected to "Iranian authorities" according to a statement from four UN Special Rapporteurs (OHCHR, 2020). In the statement, the rapporteurs noted that:

> *Journalists working for the BBC Persian Service and other Farsi-language news outlets outside Iran have faced threats, criminal investigations, unlawful surveillance, freezing of assets, defamation and harassment by Iranian authorities. Several journalists have also been targeted for going public about the harassment and seeking protection from the UN...Reports also indicate a pattern of gender-based harassment, targeting women journalists since 2009, and including the dissemination of false stories, spreading of rumours and slander, usually with highly misogynistic content and threats of sexual violence* (OHCHR, 2020).

The journalists' lawyer, Caoilfhionn Gallagher KC, described the gendered online violence her clients receive on social media platforms such as Twitter and Instagram - including rape and death threats - as: "...horrendous threats of sexual violence either to the journalists themselves, or to family members. And sometimes they're misogynistic threats to male journalists singling out female family members, or false allegations about male journalists raping their female colleagues".

One of those targeted, Rana Rahimpour, received a threat in February 2020 to assassinate her, her husband, her children, and her elderly parents (who remain in Iran), within the following month. "The message also said that Ms. Rahimpour would be the first employee of the BBC to be killed, and that, after her assassination, it will be the turn of other BBC Persian service employees," the UN Special Rapporteurs said of the reported message. The rapporteurs also stated:

> *These allegations are extremely concerning and if confirmed, would indicate that the Iranian authorities are prepared to use force extraterritorially, in violation of international law. Harassment, surveillance, death threats against journalists, within and outside domestic boundaries violate international human rights law, including the right to physical integrity, the right to life and the right to freedom of expression. We reiterate our earlier calls to the Iranian Government to cease the intimidation, harassment and threats, including death threats, against BBC and other journalists working outside Iran for Farsi-language news outlets, as well as reprisals against their family members in Iran* (OHCHR, 2020).

---

59 See Maria Ressa big data study for a detailed analysis.

In Mexico, those responsible for abuse against women journalists are often powerful men. Official reports indicate that many cases named public servants, and municipal and state officials as the culprits (DataLab, 2017). Intimidation and harassment, threats, physical and digital assaults, arbitrary detentions and prosecution, and worse are everyday experiences for many Mexican women journalists. And these threats are exacerbated by online violence.

The Mexican State is largely accountable for this situation according to lawyer Yesica Sánchez Maya, who is a board member of the Consortium for Parliamentary Dialogue and Equity in Oaxaca: "Authoritarian and populist politicians view women journalists as uncomfortable subjects just like anyone who criticises the system and its structural failures and believe that their truth-seeking work needs to be stopped" (Raynor, 2020).

Lidia Alejandra López Castañeda, a journalist for online news site Pregoneros. com.mx in Veracruz, has experienced continuous online harassment from the son of a city mayor and has filed a case against him for cyberbullying and discrimination to a state-level human rights commission.[60] Since 2018, at least five other women journalists have also made official complaints against the man for online harassment. One of them, María Elena Ferral, a reporter for *Diario de Xalapa* and *Quinto Poder*, was murdered in March 2020 (Martin, 2020). López Castañeda said:

> *He's used hateful words and misogynistic messages. The harassment has been going on for two years, since his father was elected mayor. I am concerned about the stance the authorities are going to take. Out of five journalists who denounced his attacks, one was murdered - how could I not be worried?*

Alma Delia Murillo, a *Reforma*[61] newspaper journalist and writer, experienced severe online harassment and abuse after reporting on rape allegations against Salgado Macedonio, a former mayoral candidate of Guerrero state. She received threats on Twitter and by telephone (Cerdeira, 2021b; Murillo, 2021). She also alluded to the chilling effect of these attacks. "The first thing that I thought was to never write about that subject… The fear unchains dark fantasies when you live in a country where the levels of violence are like Dante's Inferno," she said in an interview.

Gabriela Rasgado, National Journalism Prize winner and reporter for Código Veracruz, received rape and death threats on Twitter in 2019. She has said that a former mayor, who was jailed for nine years on charges of criminal association, money laundering and embezzling public funds, continued to send her threats via

60 Commissions or Comisiones Estatales de Derechos Humanos, CEDH (DAV / 0779/2020). Castañeda's case was dropped by the local government of Veracruz in the context of declining funding for investigations and for the protection of journalists. The Veracruz State Commission for the Attention and Protection of Journalists (CEAPP) also declined to act. Castañeda says that while the harassment of her and other citizens from the mayor's son stopped due to her complaint, she now self-censors as a result of the experience.

61 In 2019, journalists of *Reforma* newspaper said they had been targeted by the Mexican president in his press conferences at least 13 times after they published the private address where he was hiring extra security, and which was already in the public domain. ARTICLE 19 Mexico reported that general editorial director Juan E. Pardinas was the victim of death threats, harassment and attempted doxxing on social media, and target of a smear campaign linking the newspaper with organised crime, a common threat used by Mexican trolls. ARTICLE 19 called upon the president to not threaten press freedom, on the Office of the Special Prosecutor for Crimes against Freedom of Expression (FEADLE) to investigate online threats, and on the Mechanism to protect Pardinas: https://articulo19.org/autoridades-deben-proteger-a-director-de-reforma-por-agresiones-en-su-contra/

Twitter (Reuters, 2018; Economía Hoy, 2019). And in 2020, Adela Navarro, editor of the weekly newspaper *Zeta*, accused a local mayor of attacking the newspaper and herself during a Facebook livestream (Zeta, 2016).[62]

Still in Mexico, Patricia Montiel García, editorial director of the newspaper *El Imparcial de Tlaxcala*, was falsely accused of having a sexual relationship with two male colleagues in a smear campaign in February 2020 in retaliation for her reporting about the wife of a mayor:

> *Fake photos were published online of me and my two male colleagues, saying that I was a lover of I don't know who, that I was linked to a colleague, that the other was a rapist and murderer. And that I protect them. A month later we were targeted with death threats.*

In 2021, the Coalition For Women In Journalism (CFWIJ, 2021b) identified incidents in Pakistan and India connected to government-sponsored trolls.[63] Nighat Dad, founder of Pakistan's Digital Rights Foundation, herself a target of serious online violence described the role of political actors and the State in targeted online violence against women journalists:

> *The attacks are made by people declaring affiliation with the ruling party, and in the coordinated campaigns women journalists are referred to as peddlers of 'fake news', enemies of the people, and accused of taking bribes. Some journalists [have] shared that, after official harassment, their social media accounts are bombarded with gendered slurs and abuse by accounts displaying the ruling party's flag or the Prime Minister's picture.*

In mid-2020, a group of prominent Pakistani journalists issued a statement condemning coordinated attacks on women reporters. "Vicious attacks through social media are being directed at women journalists and commentators in Pakistan, making it incredibly difficult for us to carry out our professional duties," they wrote. "The online attacks are instigated by government officials and then amplified by a large number of Twitter accounts, which declare their affiliation to the ruling party" (CFWIJ, 2020).

On behalf of Pakistan's ruling party, the Minister for Human Rights Shireen Mazari condemned the harassment described by the women journalists during a session of the National Assembly's Standing Committee on Human Rights: "There should be no abusive language on social media... If anyone has proof, we will take action.

---

62 A tweet posted by Adela Navarro in July 2020 translates as: "Again, the mayor [Baja California] attacked Zeta. Yesterday he attacked my personal life. He is bothered by #InvestigativeJournalism, that researches further, that doesn't conform to the information that he controls. We will continue because #Wearefreeasthewind": https://twitter.com/adelanavarro/status/1283135553236463616
63 Pakistani journalists who queried the representation of women in school curriculums, Asma Shirazi (criticised by ruling party Pakistan Tehreek-e-Insaf for allegedly insinuating insults against Prime Minister Imran Khan's wife, IFJ, 2021) and Gharidah Farooqi, were both attacked in coordinated online abuse campaigns.

Those who harass others using the PTI's name should be identified… If someone believes any PTI account is engaged in harassment, they should point it out. Action will be taken against any PTI account engaging in harassment" (Zaman, 2020).[64]

Kenyan editor Catherine Gicheru alleged that political players in her country have employed groups of 'patriotic trolls' which she compares to "soldiers in a war" who are paid to attack political rivals or negatively influence public opinions or debate. In Gicheru's experience, these are typically "legions of young people who will be mobilised to attack at the request of some dumb political person who has an axe to grind with whatever a journalist has written… they will mobilise them to attack". She estimates that 90% of the abuse experienced by women journalists is political, "since online spaces are used as a tool to silence people here", with a direct line from politician to troll:

> *The politician will recruit individuals for that type of work. If [a woman] is too vocal or too strongly opposed to something, he will just send those on his payroll to attack her. [They are] extremely well organised… They have department heads who sit and direct, saying things like 'you do that, you do that.*

In Nigeria, Oluwatosin Alagbe, former Judiciary Correspondent for the *Premium Times*, now a freelancer, said she received threats on Twitter from members of the Indigenous People of Biafra (IPOB) separatist movement when she covered the trial of their leader Nnamdi Kanu in 2017.

The patterns are similar in Eastern Europe. A local mayor of the small Polish town Przedbórz was sentenced to one year in prison in 2019 and lost his office for inciting another person to send threats and abuse to local journalists, bloggers, and political opponents of the mayor (Piotrków Trybunalski Nasze Miasto, 2019).

In Serbia, TV Prha journalist Tatjana Vojtehovski was subjected to a series of death and rape threats on Twitter (which included threats made against her daughter), for reporting on anti-government protests. Associated with these attacks was a fraudulent tweet using Vojtehovski's name and purporting to show her supporting a convicted pedophile. A man was sentenced to eight months of house arrest in connection with these threats (Council of Europe, 2019).

Vojtehovski has said such attacks were a daily reality: "There's usually a trigger for a threat to be issued – for example, the last trigger was Serbian President Aleksandar Vučić stating on some television channel that I do everything against the State, so the threats via Twitter began that same evening" (Slavko Ćuruvija Foundation, 2018). Other researchers have also noted these patterns on a global scale (e.g., Nyst and Monaco, 2018; Jankowicz et al., 2021), and the algorithmic bias that boosts such narratives online has also been acknowledged by Twitter (Milmo, 2021; Huszár et al., 2021).

In the Philippines, digital 'red-tagging' is another form of abuse used against women journalists. It involves falsely identifying a journalist as a member of the

Communist Party of the Philippines (CPP) or the party's military wing, the New People's Army (NPA). Individual community journalists are more vulnerable to these attacks but in some cases entire news outlets, such as Bulatlat, Northern Dispatch, Rappler, and CNN Philippines, have been tagged as being the media of the CPP-NPA.

In 2017, then President Rodrigo Duterte declared the CPP-NPA a terrorist organisation (Presidential Communications Operations Office, 2017). In line with these efforts, he created the National Task Force to End Local Communist Armed Conflict (NTF-ELCAC) in 2018, which has been accused of red-tagging journalists, news organisations NGOs, human rights defenders, politicians, and even celebrities (Lopez, 2020; Lalu, 2020d; Posetti et al., 2021).

According to Kim Quitasol, Editor-in-Chief of online community news outlet Northern Dispatch, journalists who are red-tagged on social media become an open target for offline violence. The Philippine Commission on Human Rights (CHR), Amnesty International, and Human Rights Watch have all linked recent killings of Filipino rights activists to red-tagging (Gavilan, 2020; Lalu, 2020b; Amnesty International, 2020e; Robertson, 2020).

In early 2021, nine red-tagged activists were killed two days after the president ordered the military and police to kill insurgents without regard for human rights if they were carrying weapons (Talabong, 2021). The killings also followed the red-tagging of Maria Ressa's Rappler by the National Task Force to End Local Communist Armed Conflict (NTF-ELCAC) in a threat published on Facebook associated with a warning about what was to come. That post was removed by Facebook (Posetti et al., 2021).

The "fake news" factor

The function of such anti-press rhetoric as a trigger for gendered online violence against journalists is very familiar to women reporters in the US, who experienced exponential online attacks during Donald Trump's presidency. Trump frequently incited his supporters to take action against journalists (Feller and Minutaglio, 2021). "At the heart of [Trumpism] is an anti-journalism agenda, an anti-free and independent media agenda, because accurate information is like Trump's kryptonite," former UN Special Rapporteur, David Kaye said in an interview for this study.

Expressions of online violence often echoed his distinctive anti-journalism rhetoric, including labelling credible journalists and their work "fake news", demonstrating the functions of disinformation and gaslighting[65] in his attacks, and using sexist or sexually explicit language to refer particularly to women journalists from minority backgrounds (CFWJ, 2021).

Like a number of counterparts elsewhere, Trump regularly smeared, disrespected and insulted women reporters during official briefings and press conferences, especially those who challenged him with questions designed to hold him to account. The objective was to chill their critical reporting and to undermine their

---

65 For a discussion of gaslighting works in practice to destabilise or discredit a target, see: https://www.bostonglobe.com/2021/03/18/opinion/when-hate-crime-is-clear-gaslighting-often-follows/

confidence. For example, in the context of the presidential debate that NBC News' Kristen Welker moderated in 2020, she was attacked by the Trump campaign machine almost daily in an effort to destabilise her (Rupar, 2020).

In July 2020, Al Jazeera's White House correspondent Kimberly Halkett was falsely accused of calling then President Donald Trump's Press Secretary Kayleigh McEnany a "lying bitch" during a press briefing, after earlier being mocked on Twitter. Audio recordings misrepresenting the exchange between Halkett and McEnany were ultimately picked up by sections of the mainstream press, despite Halkett's denials, before being debunked (Baragona, 2020). This experience resulted in a flood of misogynistic abuse, including death and rape threats against her and her teenage daughter, leading to a need for increased physical security.

US academic Dr. Michelle Ferrier - herself a victim of online violence while a reporter - founded Trollbusters. She highlighted the escalating online-offline threats faced by women journalists in the US as a result of former President Trump's 'dog whistling': "There has always been a core element inside of that that is very organised, very violent and very intent on racial and gender violence. And using all means to be able to organise and deploy those strategies online and off".

Trumpism has cast a long shadow and the end of the Trump presidency in 2021 did not materially change US women journalists' lived experience of online violence. It has also been exported as a weapon of populist leaders internationally.

In Brazil, President Jair Bolsonaro told Talita Fernandes, formerly of *Folha de São Paulo,* to "shut up" during press conferences and described her questions as "not interesting". In this context, she said she is "insulted" online by trolls every time she writes something critical of the president (Sakamoto, 2020). Fernandes has also been booed at the presidential palace. Fraudulently manipulated video of these incidents has been circulated online to undermine her credibility, make her questions "seem stupid", and "build a narrative" that she "wanted to attack" the president. One pile-on after the then-minister of state retweeted the manipulated video, even reached her "shocked" grandmother in the São Paulo countryside via WhatsApp (ibid.).

Since President Bolsonaro's election in 2018, black and Indigenous Brazilian women journalists have become bigger targets for intersectional online violence (Lehman, 2018; De Sousa, 2020; Beirich and Via, 2021). In 2020, the president called Maria Júlia (Maju) Coutinho, a presenter on TV Globo group, a "liar", and this was followed by a torrent of insults and misogynistic and racist threats (the journalist is black) by Bolsonaro supporters on social media (see the hashtag: #MajuMentirosa) (RSF, 2020i).

In South Africa, as previously mentioned, there have been large scale online attacks on women journalists instigated by the leader of the populist Economic Freedom Front (EFF) party, Julius Malema, and his associates. The South African Editors' Forum (SANEF) made the direct link between the EFF leader's online or on-air comments and the journalists being targeted online and offline by "EFF leaders on social media and at various political rallies".

SANEF highlighted the "chilling effect in newsrooms" of the case of Daily Maverick's Pauli Van Wyk which is referenced throughout this study, and involved online
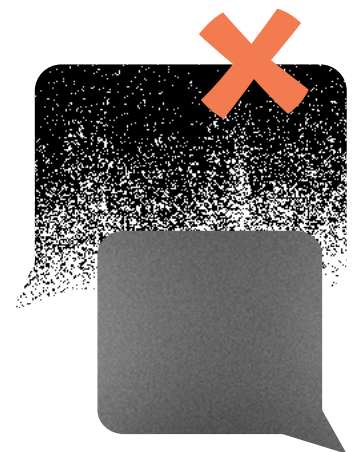
threats of extreme sexual violence and murder, "with specifically younger journalists feeling fearful and intimidated and thus withdrawing from critical reporting on politicians" (SANEF, 2018).

Bringing political perpetrators to account

Since 2018, Brazilian journalist Patricia Campos Mello has endured a torrent of disinformation-laden online attacks instigated by political actors connected to her critical reporting on President Jair Bolsonaro's election campaign. She described the attacks as "very aggressive" and saw them "snowballing" on Instagram, Facebook and Twitter when "several Bolsonarist deputies, right-wing bloggers and then-deputy Eduardo Bolsonaro [President Jair Bolsonaro's son]" said that they did "not doubt that she has sexually insinuated herself to get information".

Campos Mello sued and won moral damage cases against President Bolsonaro and his son for the claim (Neder, 2021; BBC, 2021b; Fundamedios, 2020).[66]

Bianca Santana, a columnist for the largest Portuguese language website UOL, was falsely accused by President Bolsonaro of spreading "fake news".[67] Like Campos Mello, she pursued civil charges for moral damages and won. The YouTube account featuring the attack was de-platformed[68] and the president was ordered to pay her damages in December 2020 (FENAJ, 2020).

While the UK government announced a new framework to address safety of journalists (among other forms of attack) in 2021, two black journalists interviewed for this study experienced severe online harassment after their journalism was discredited by the UK's Equalities Minister. Nadine White was a reporter at HuffPost UK when the Minister named her on Twitter, described her as "creepy and bizarre" and said her conduct provided "...sad insight into how some journalists operate". A flood of abuse followed - via email, phone and social media (Council of Europe, 2021).

In response, White and her editor called for a public retraction, and then sent a letter to the minister, asking for an apology. White said:

> *As a black woman journalist, coming into this industry as the vast minority within this white-dominated elite base, it's daunting on so many levels just to get up in the morning each day and do what I do, much less to know that you've been targeted by a minister and that's been effectively sanctioned by the people that run this country.*

In March 2021, the UK's Permanent Secretary to the Cabinet Office replied to an alert issued by the Council of Europe about the incident - triggered by a complaint lodged by a number of press freedom organisations - stating that:

Patricia Campos Melo
Picture credit: Felipe Campos Mello

---

66 Brazilian journalist Patricia Campos Mello has written a book about her online abuse: The Hate Machine - A Reporter's Notes About Fake News and Digital Violence (Companhia das Letras). Her legal action against the Bolsonaros is discussed in further detail in Chapter 6.0.
67 Santana connected Bolsonaro to the 2018 assassination of local politician Marielle Franco before she was targeted in the President's weekly YouTube address: https://www.uol.com.br/ecoa/colunas/bianca-santana/2020/05/26/por-que-querem-federalizar-as-investigacoes-do-assassinato-de-marielle.htm
68 Santana tweeted the video apology from the Presidential Palace spokesperson: https://twitter.com/biancasantana/status/1288990447629737985

The Government agrees that the harassment of journalists and those working in the media is unacceptable and that is why the United Kingdom has established the National Committee for the Safety of Journalists [and]... a National Action Plan for the Safety of Journalists, which looks at measures to address the abuse experienced. The Government notes that these tweets were not issued from an official government Twitter account, but instead from a personal Twitter account. The Minister is personally responsible for deciding how to act and conduct herself, and for justifying her own actions and conduct (ibid).

ii. Sources, colleagues and competitors as perpetrators

Sources and contacts were also identified as top perpetrators of online violence experienced by 15% (n=107) of the women journalists surveyed for this study. This is a pattern that reflects women journalists' exposure to offline violence (NBC, 2018). Whether delivered virtually or physically, violence against women journalists perpetrated by sources or contacts is often used as a method to silence their critical reporting. Reflecting the ongoing issue of workplace sexual harassment experienced by women journalists (Schmall, 2021; McCurry, 2019) in the newsroom, incidents of online violence perpetrated by colleagues were identified by 14% (n=100) of the woman survey respondents.

The survey results and the interviewees also pointed to the role of journalists from other news organisations. Fifty-seven of the women journalists surveyed identified "staff of rival news organisations" as among their regular abusers. A senior UK political journalist who chose to remain anonymous said "some of the worst stuff" came via other journalists on the political left. Interviewees also referenced other cases in the UK where journalists have been subjected to gaslighting attacks and 'dogpiling' from journalists and commentators featuring pernicious sexism and misogyny (see the big data case study on Carole Cadwalladr in Chapter 3.2).

In the US, far-right media have regularly launched campaigns against women journalists who confronted former President Trump, or reported on 'lightning rod' issues such as politics, religion, white supremacy, guns, hate speech, feminism, internet culture, and disinformation. In 2019, partisan outlets such as Breitbart accused Guardian US' Julia Carrie Wong of racism, after she reported on links between it and white nationalists, triggering attacks by other extreme sites.

When Trump's spokesperson called questions by The Grio's White House correspondent April Ryan "ridiculous", a segment would make it onto Fox News that very night mocking her, as "the most ridiculous question of the day at the press conference". Fox News' Tucker Carlson has amplified attacks. In October 2020, he provided a platform to a white nationalist troll engaged in online abuse of Brandy Zadrozny, who covers disinformation at NBC News-MSNBC. The man

15%
OF THE WOMEN JOURNALISTS SURVEYED

SAID **SOURCES AND CONTACTS** WERE TOP ONLINE VIOLENCE PERPETRATORS

falsely accused Zadrozny of doxxing Trump supporters,[69] and "hundreds" of threatening voicemail messages and emails to the journalist followed, mentioning her children; stating that her throat would be cut; and that trolls were outside her home. This led her employer to hire armed guards to protect her at home.

In another case, after Carlson profiled then-*New York Times* journalist Taylor Lorenz on his show, she said she was "overrun" by attacks coming via multiple platforms, including Clubhouse. She was "inundated" by six Clubhouse conversations naming her in one week, "including one saying I should be in jail". As a result of sustained attacks across platforms and her then employer's failure to take steps to protect her, she said, she had a "mental health crisis" and was "contemplating ending [her] life".[70]

Journalists and editorial leadership from Notimex, the Mexican State News Agency, also feature as a 'top trigger' for online abuse for journalists in Mexico (Kansara and Martinez, 2020). Harassment against Reyna Haydée Ramírez of Pie de Página on Twitter and Facebook turned threatening when she criticised President Andrés Manuel López Obrador during a press conference on live television (Sandoval, 2020). The threats were traced back to Notimex by the Laboratory for Technological Innovation and Applied Interdisciplinary Studies (Signa_Lab) at the Universidad Jesuita de Guadalajara (ITESO).

Signa_Lab has also documented attempts by officials from Notimex to pay to gain access to a reporter's social media accounts to obtain images, blackmail attempts, smear campaigns and other forms of harassment (Signa_Lab, 2020). "In 27 years as a journalist, this is the first time I have been threatened with death - even if it is on social media, that is still a threat," said Ramírez. "My first reaction was to complain to the federal government, because they are responsible for the owner of Notimex. They never answered me, but they sent people from the Mechanism [for the protection of journalists] and told me to file a complaint, about the threats and for discrimination, and to declare whether I wanted psychological care".

iii. Targeted for reporting on gender issues

Gender was the story theme most frequently associated with heightened online violence according to the survey results and many interviewees. Nearly half (47%; n=336) of the women survey respondents rated reporting or commenting on gender issues as a top trigger for online attacks. Themes included reporting on feminism, domestic violence, sexual assaults, femicide, reproductive rights (especially abortion), and transgender issues.

In 2019, Mexican broadcast journalist Montserrat Ortiz received online death threats in response to her reporting on sexual violence, and said: "I am in fear, I do not feel safe, I do not feel well" (Chocarro, 2021; ARTICLE 19, 2019). In addition to chilling such reporting, the threats and abuse are designed to dehumanise the

14%
OF WOMEN JOURNALISTS SURVEYED IDENTIFIED
COLLEAGUES AS ONLINE ATTACKERS
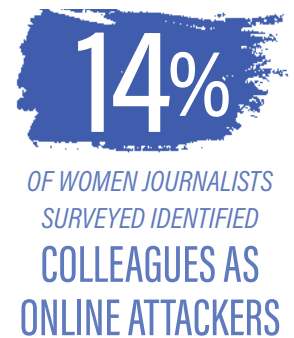
REPORTING ON GENDER IS A TOP TRIGGER FOR ONLINE VIOLENCE ESCALATION *ACCORDING TO*

47%
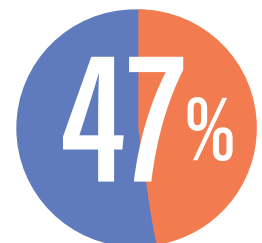
WOMEN JOURNALISTS SURVEYED

---

69 Beattie took a chapter written by Brandy Zadrozny in a book designed to help journalists investigate disinformation networks, and misrepresented it as doxxing on her behalf: https://datajournalism.com/read/handbook/verification-3/investigating-actors-content/how-to-analyze-social-media-accounts
70 Another Tucker Carlson prolonged pile-on against Taylor Lorenz from early 2021, in which he mocked the journalist for discussing her experiences of online violence (Sullivan, 2021).

women journalists covering gender-based violence. Also in Mexico, Alma Delia Murillo, a *Reforma* newspaper journalist and writer, experienced online abuse on Twitter (and by telephone) after reporting on rape allegations against Salgado Macedonio, the former mayoral candidate of Guerrero state: "[It] feels like going to work and, as you walk down the hallway to your office, fifty people are shouting epithets at you: 'prostitute', 'idiot', 'sell-out', 'mad as a hatter'"(Chocarro, 2021).

Those who write from a feminist perspective are also commonly targeted. *Daily FT*'s Marianne David, a veteran Sri Lankan business journalist, was called a traitor to her country when she spoke out against women being sexually harassed on the street. In Lebanon, writer Zahra Hankir said that a caricature of her in a rubbish bin circulated online, implying she was worthless, after she replied to a tweet in support of an imprisoned women's rights activist in Saudi Arabia.[71]

Khaoula Boukrim, Editor-in-Chief of Kashf Media[72] in Tunisia, was attacked on Egyptian social media pages for supporting the hashtag #FreeSolafa concerning a jailed woman journalist.[73] Natalia Żaba, a journalist based in Serbia for a decade, followed the abuse levelled against a prominent journalist covering the #MeToo[74] movement: "I was watching on Twitter how they attacked [them]...literally threats of rape and death threats."

Other human rights and social justice issues were identified by nearly a third (31%; n=221) of survey respondents as a category of stories attracting high levels of abuse, while immigration issues were selected as a top trigger for attacks by 17% (n=121) of survey respondents.

iv. Political reporting as a top online violence trigger

The reporting theme second most likely (44%; n=314) to be met with online harassment and abuse was 'politics and elections', according to the ICFJ-UNESCO survey. This underscores the role of political attacks on the press in exacerbating threats to the safety of journalists.

Brankica Stanković is the complainant in a number of ongoing cases regarding targeted online violence in Serbia. She had reported on the ties between gangs and political figures for B92 television's Insajder programme. The police later discovered that the barrage of on- and offline threats against her came from a group of the most notorious criminals in Serbia, who assassinated the first democratically elected Serbian Prime Minister, Zoran Đinđić, in 2003 (IWMF, 2021; CINS, 2009). Stanković previously lived under police protection for five years between 2004 and 2009 (Insajder Video, 2015).

Catherine Gicheru of the Africa Women Journalism Project (AWJP) said that in Kenya much online violence was triggered by "(t)he usual nonsense...that people don't like women being opinionated about politics". This is something that Nancy Kwamboka said she has also experienced as the head of radio Kenyan station

REPORTING ON POLITICS AND ELECTIONS IS A LIGHTNING ROD FOR ONLINE HARASSMENT ACCORDING TO

44% OF THE WOMEN JOURNALISTS SURVEYED

---
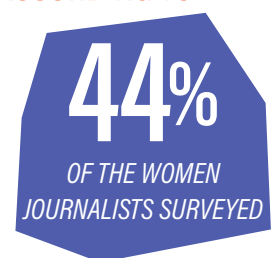
71 See footnote 79.

72 Kashf Media (meaning 'Reveal' Media) is a non-profit media platform created in 2020 reporting on topics related to the northwestern part of Tunisia (Zamit, Kooli and Toumi, 2020).

73 Solafa Magdy, detained in Egypt, testified about being sexually assaulted and severely abused in prison in 2021: https://ifex.org/egypt-imprisoned-journal-ist-solafa-magdy-subjected-to-grave-violations-in-prison/

74 Brankica Stanković, journalist with Insajder TV, broke the story of Serbian actress Danijela Steinfeld, who started the #MeToo movement in Serbia in connection with multiple allegations of sexual assault by a prominent actor and director.

Egesa FM. She is harassed online by "keyboard warriors" or "goons" who react to news stories about local political developments (Maundu, 2020a). Broadcast journalist and digital security trainer Cecilia Maundu said online violence "becomes very rife, and it peaks in connection with elections". This was a view echoed in a 2021 tweet by Judie Kaberia, Executive Director of Association of Media Women in Kenya (AMWIK) who pointed to the worsening of online attacks in the lead-up to the 2022 Kenyan elections.[75]

In Sri Lanka, an anonymous interviewee described being attacked with "hundreds of messages" on Facebook Messenger, as well as via the pages of political parties: "Then other admins reposted it...and sent me messages, voice recordings saying that...I would rape you, kill you, throw acid at you, kidnap you - beware when you are taking a bus...and that has happened a few times."

A BBC reporter highlighted the fact that politically motivated online violence in the UK comes from both the far left and the far right:

> *If I had to kind of measure it, the far-left has been by far the worst for me by a million miles...the most alarming bit to me is about how... [respectable] people...have started using attacks on journalists as part of their campaign. That's much more alarming to me than actually the sort of mob stuff either online or in real life, which is horrible. I still get abuse online and emails saying, "why don't you fuck off, and die of ovarian cancer".*

v. Targeted for disinformation-busting

Attacking women journalists is a favoured tactic of disinformation agents; disinformation methods are co-opted in misogynistic attacks; and reporting on disinformation and associated themes (e.g., digital conspiracy communities; far-right networks) is a lightning rod for attacks (Posetti, 2018a; Posetti, 2020b; Bontcheva and Posetti, 2020; Jankowicz et al., 2021).

In the words of UK columnist and investigative reporter Carole Cadwalladr who writes for the Observer: "If you report on disinformation, you become a target of disinformation". At a global level, 16% (n=114) of the women journalists surveyed identified disinformation as a beat likely to lead to an increase in online violence.

Most of the women journalists interviewed for this study had experienced disinformation-based attacks designed to smear their personal and professional reputations, hold them up to ridicule, and expose them to increased offline risk. They were targeted in the context of reporting on interlinked issues like right-wing extremism, conspiracy and disinformation networks, the manipulation of electorates by political actors deploying disinformation tactics, and their accountability reporting focused on social media companies as corporate facilitators of disinformation and hate speech.

75 Kaberia also tweeted this, in November 2021: https://twitter.com/AMWIK/status/1461599887620034561

But while many women journalists participating in this study have been subjected to smear campaigns, covering disinformation appears to be a particular trigger. For example, in late 2018, Brazil's Patricia Campos Mello published a series of articles investigating disinformation allegations associated with the Brazilian election that swept President Jair Bolsonaro to power in 2019.

The orchestrated attacks against the *Folha de São Paulo* journalist have involved significant disinformation elements - including false allegations that she traded sexual favours for information. Such disinformation was published by right-wing bloggers and Congressman Eduardo Bolsonaro, son of the President (Campos Mello, 2020). As a result, the internet became littered with falsified pornographic images of her, deep fake videos portraying her as a "prostitute journalist", disinformation-encrusted viral memes, and rape threats.[76]

Also in Brazil the founder of the fact-checking outfit Lupa, Cristina Tardáguila,[77] had to flee the country after threats on her life. "Filthy scum," one direct message on Twitter read. "When the time comes we will come looking for you, one by one." A viral cartoon caricatured her and two other prominent women fact-checkers - half-naked and leashed like dogs. The prominent philanthropist George Soros, who founded Open Society Foundations (OSF) and invests in counter-disinformation initiatives, was depicted holding the leash and dangling a wad of dollars above the women (Palau, 2021).

In the US, Brandy Zadrozny, who covers disinformation and the far-right, said the pattern of attacks against US journalists who report on disinformation is becoming very uniform. This trend also involves the increasing "blowback and abuse we get for telling that story", she said.

When the BBC's specialist disinformation reporter Marianna Spring interviews conspiracy influencers, she said she is often intimidated with recordings of the interview which are later preemptively shared online. These trigger pile-ons in which she is, ironically, accused of peddling disinformation. In another incident, a 2,000-word blog post purporting to have been written by Spring, adopting a first-person perspective, was published on a fringe site. Around 3,000 hypersexualised threats and rape threats were posted underneath the 'story'. It was "really, really, really disgusting," Spring said.

vi. The puppeteers: Orchestrated disinformation campaigns against journalists

A high number of women journalists responding to the UNESCO-ICFJ survey (41%; n=293), and of those participating in interviews, have experienced online violence in the context of what they perceive to be coordinated disinformation campaigns which seek to undercut critical reporting (Edström, 2016; Posetti, 2017a).

The orchestrated, disinformation-fuelled campaign against South African editor Ferial Haffajee (referenced in Section 2.2.1) was triggered by her journalism on

41%
OF THE WOMEN JOURNALISTS SURVEYED

HAD BEEN TARGETED IN **ORCHESTRATED DISINFORMATION CAMPAIGNS**

COVERING DISINFORMATION IS A **TRIGGER FOR ONLINE VIOLENCE ESCALATION** *ACCORDING TO*

16%

OF WOMEN JOURNALISTS SURVEYED

76 Watch a panel discussion featuring Campos Mello, Marianna Spring and Rana Ayyub at the 2022 International Journalism Festival
 https://www.journalismfestival.com/programme/2022/online-violence
77 Tardáguila now works for the International Center for Journalists (ICFJ) which was commissioned to conduct this study.

the so-called 'Gupta Leaks' investigation into corrupt business practices that led to what South Africans refer to as 'state capture' (February, 2019). The now disgraced and defunct UK-based public relations firm Bell Pottinger deployed a disinformation campaign linked to the then-Jacob Zuma Presidency in connection with the 'Gupta Leaks' investigation (Thamm, 2020). The strategy involved networks of Twitter 'bots' (both automated and fake human-operated accounts), with the objective of intimidating Haffajee into silence.

Fabricated, sexist and sexualised images of Haffajee, including pictures of her face photoshopped onto dogs' bodies, were used in the Twitter attacks. In one example, she was portrayed in sexually compromising positions with a billionaire businessman she had never met. She wrote about the impacts at the time: "The images make me wince with their distortions and insults. I snap my phone shut and move to another screen. Or make a cup of tea. Images are powerful and the designers have very specific messages. That I am a whore, a harridan, an animal and a quisling" (Haffajee, 2017).

The attacks on Haffajee died down in the period after the demise of disgraced president Jacob Zuma, the downfall of Bell Pottinger, and anti-trolling measures introduced by Twitter in 2020,[78] she said. However, in early 2021, online violence against Haffajee and her Daily Maverick colleagues Pauli Van Wyk and Marianne Thamm began to escalate again in the context of dog-whistling by partisan media. This also demonstrates the traits of networked gaslighting which involve coordinated or interlinked attacks that undermine the confidence and credibility of the abuse target, by twisting perceptions of reality, and falsely attributing the same abusive behaviour (including disinformation-laced attacks) to the subject.

The Zuma-aligned Independent Media news group escalated its disinformation-laced attacks on Daily Maverick's editorial staff and their investigative collaborators at the amaBhungane Centre for Investigative Journalism, triggering death and rape threats against the women journalists in the firing line on social media. Attacks mounted by Independent Media include a discredited article falsely accusing Haffajee, Van Wyk and Thamm of participating in orchestrated disinformation campaigns which the paper fraudulently equated with Bell Pottinger's dark propaganda assault (Makhowana and Dlamini, 2021). This conduct mirrors partisan media amplification and instigation of online violence attacks against women journalists reporting on disinformation in the US, as discussed above.

Kenya is one of the countries where the use of 'cyber troops' has been prevalent in spreading disinformation particularly during election periods (Bradshaw and Howard, 2019; Nitsche, 2019). Reporters Without Borders (RSF) has also documented attacks on major media outlets in Kenya, including the 2019 trolling of the Nation Media Group, East Africa's largest independent media house, by pro-government activists (RSF, 2021a).

There is some evidence that 'reverse swarming' (i.e. mass amplification of supportive messages for women under attack on social media) and mass reporting of perpetrator accounts can be effective in the Kenyan context. Catherine Gicheru, editor, founder and director of Africa Women Journalism Project (AWJP), refers to it as "raining on" trolls via informal womens' networks: "We mobilise each other

---

78 Twitter introduced easier abuse reporting mechanisms, targeted monitoring and support in the midst of swarm attacks, and the ability to limit respondents to public tweets; see Chapter 5.0 (Platforms and vectors: Assessing big tech responses to online violence) for more.

from different regions of the country, and occasionally from beyond the country. We usually respond by calling the perpetrator out and otherwise engaging with these trolls so that you, as the victim, don't feel you're alone."

In Lebanon, according to the interviewees, the perpetrators behind most orchestrated disinformation campaigns associated with online violence against women journalists are State and political actors (including foreign powers), and businessmen. Writer Zahra Hankir faced violent online retaliation involving disinformation after showing her support for Saudi women's rights activist Loujain Alhathloul using the hashtag #freeloujain.[79] She was subjected to a week-long campaign of organised online violence involving rape and death threats, remarks that she was "a Lebanese whore", and being falsely accused of being affiliated with the Muslim Brotherhood.[80] "We're likely talking about troll armies here rather than individual actors. The harassment felt targeted and...too vast for it to be sporadic".

vii. "Foreigners" and "spies"

Coordinated disinformation campaigns which also often involve so-called 'patriotic trolling' narratives designed to question the target's loyalty to her country or culture, are also evident in the research undertaken for this study. In some cases, attackers used doctored video and audio that distorted reporters' work and was widely distributed to smear their reputation. In the cases of Kimberly Halkett in the US and Talita Fernandes in Brazil, manipulated audio and video was edited to make it look like they adopted an anti-presidential narrative.

In the UK, the BBC's specialist disinformation reporter Marianna Spring was the subject of a YouTube video falsely accusing her (among other things) of being a "spy" because she studied modern languages at the University of Oxford. The video still had not been removed in late-2021, despite being reported many times.

In Tunisia, Henda Chennaoui said that she had experienced false allegations that she received money from foreign NGOs, and that she participated in the revolution as an act of conspiracy against Tunisia "with aid from Western countries". Meanwhile, Khaoula Boukrim, Editor-in-Chief of Tunisia's Kashf Media website, says she has been called "a leftist...[with] a secret agenda", "a mercenary", a "foreign agent", a "traitor", "unpatriotic" and "pro-Emirati". When Boukrim said in a 2019 TV interview that it was not a political party's right to incite people during protests, she was targeted on Facebook pages with manipulated video from the interview, used to create the false impression that she was "inciting people to take to the streets to bring down the regime".

Nigerian documentarian Ruona Meyer was targeted in a campaign of extreme online harassment which lasted almost a year after the BBC broadcast her award-winning investigation.[81] This association with the BBC, as well as her marriage to a German citizen, led to anonymous trolls accusing her of being a 'foreign agent'. "I'm also somebody not to be trusted because I'm not patriotic. Why? Because I have...'sexually transmitted citizenship'," she said. Trolls - linked, she suspects,

---

79 Loujain Alhathloul had been jailed for three years for advocating for women's right to drive in Saudi Arabia; the hashtag #freeloujain garnered over 1,500 likes and was retweeted by Amnesty International and other human rights organisations.
80 The Muslim Brotherhood is a religio-political group that has been banned in Egypt where it has been branded a "terrorist organisation". See: https://www.britannica.com/topic/Muslim-Brotherhood
81 https://www.bbc.co.uk/programmes/w3csyhqw

to those the investigation exposed - created a fake blog in her name and posted homophobic content on it.

> *"Whenever they want to turn the Nigerian public against me they'll say things like I'm 'wayward', I 'married a foreigner', my 'husband's penis doesn't work'... What has that got to do with my work as a journalist?"*

These attacks speak to how gender stereotypes may be used to undermine women journalists' credibility and undermine trust in their reporting. Meyer also takes the threats with the utmost seriousness - her journalist father Godwin Agbroko[82] was assassinated for his journalism.

Attacks on Serbian women journalists include them being called 'Albanian' and 'gypsy' as ethnic insults (Djurić & Jović, 2020). In Ana Lalić's case, she was called a 'mercenary', a 'traitor', and 'unpatriotic'. The journalist from Nova.rs, a Serbian news website, has said she was harassed on the streets, including being thrown out of some establishments on the basis that she was 'not a patriot' (ARTICLE 19, 2021b).

Meanwhile, Maria Ressa, the Nobel Prize winning Filipino-American journalist is falsely accused of working for the CIA, and being a "foreigner", amid the torrential online violence to which she is subjected.

viii. The role of the COVID-19 pandemic

The COVID-19 pandemic has exacted a terrible toll on journalists worldwide - both personally and professionally (Posetti, Bell and Brown, 2020). Gendered online violence against journalists has escalated during the pandemic - due in part to the rise of online conspiracy communities associated with pandemic denialism and anti-lockdown/anti-vaccination movements. Further exposure to toxic online communities has also come via increasing reliance on digitally-mediated reporting and audience engagement practices necessitated by social distancing.

In the US, the former *New York Times*' journalist Taylor Lorenz told the researchers: "People I care about died during the pandemic and I was trying to grieve while being smeared and attacked on every platform. Throughout it all, I had to perform at the highest level and deal with this multi-platform harassment campaign, while my own employer stigmatised me and repeatedly blamed me for the harassment I was receiving."

COVID-19 reporting as a trigger for online violence against women journalists

UK journalist Carole Cadwalladr's experience of online abuse and harassment has been influenced "very acutely" by the coronavirus pandemic, she said. "My Brexit trolls converted overnight into Corona trolls." Similar accounts about the exacerbating role of the pandemic were offered by many other research participants internationally.

---

82 Journalist Godwin Agbroko was shot in his car in Lagos, Nigeria, in 2017. He was head of the editorial board of the daily *This Day*, and won the PEN/Barbara Goldsmith Prize in 1997. Agbroko was arrested twice during Nigeria's military rule in the 1990s, as an editor (UNESCO, 2007).

Interviewees in Mexico, Brazil, Tunisia, Kenya, Lebanon, Pakistan, Poland, Serbia and the US described rising attacks on journalists in the context of increased polarisation, eroding trust in journalism and other public services during the pandemic, during which a climate of impunity for offline attacks fed online attacks and vice versa.

For many journalists interviewed, online violence connected to reporting on COVID-19 was among the most damaging because it was designed to undercut their professional credibility, with impact on public trust in their journalism - at a time when facts were under fire in the context of a 'disinfodemic' (Posetti and Bontcheva, 2020).

In Serbia, armed officers searched journalist Ana Lalić's home after she reported on the inadequate working conditions and lack of basic medical equipment at a hospital at the beginning of the pandemic.[83] She was kept in police custody for 24 hours for "publishing texts that cause panic and disorder", a day after the passage of new legislation used to restrict journalists in the name of COVID-19 counter-measures (ARTICLE 19, 2020b). Lalić was subjected to an online smear campaign when she was released from jail after the charges were dropped. She said her attackers act like they "can really do whatever they want to me".

During this episode, a digital advertisement appeared on a popular game on Google Play which showed a photo of Ana Lalić with a caption accusing her of "acting against the interests of the state" and of being "state enemy number one". This advert pointed users to a website carrying defamatory and false information about the journalist's work and private life. Lalić endured months of online harassment as well as offline violence – her car tyres were slashed in the aftermath (Council of Europe, 2020) and security had to be hired.

Lalic described a knock-on effect that further endangers her: she is now publicly recognisable as a result of the exposure associated with her arrest and the online attacks that followed. It felt for her like an "announcement of an open hunt" involving rape threats and other indignities, like "Being hanged on a lamp post, infected with Corona". She was also asked "Does that cow have a child?". On Google Play. Lalic said "my photo was popping out in front of whoever was playing a game, calling me a traitor, a mercenary". Lalić added that a misleading notion that journalists are "public figures" subject to the same kind of scrutiny as politicians or celebrities, creates additional burdens: "instead of writing the news, [you] become the news," she said.

Politically-instigated slurs have also become increasingly common in Mexico during the pandemic. When columnist and academic Denise Dresser tweeted at the Undersecretary of Health, asking for official COVID-19 death toll figures,[84] President Andrés Manuel López Obrador replied in reference to her question in a press conference: "we are living in a vulture season".[85] Thereafter, hashtag #LadyZopilota (#LadyVulture) trended on Mexican Twitter, accompanied by a

---

83 Serbian women journalists Dragana Pećo and Milica Vojinović and social media manager Anđela Mitrović, all from the Crime and Corruption Reporting Network (KRIK), received threats online before their apartments were ransacked in 2017, 2019, and 2020, respectively (KRIK, 2019; 2020).
84 The tweet read: "Attention Hugo López-Gatell. Where are the official projections for the death toll in Mexico? In the US they calculate 100,000-200,000. What model and figures are they using here? It is urgent to know to understand the magnitude of the COVID-19 crisis and how to face it better."
85 President López Obrador said that opponents of his government and some members of the tabloids have not helped in the COVID-19 emergency "because hatred dominates them," including "some journalist" who had asked to know how many deaths will be registered. He concluded that "we are living in a vulture season" (Redacción AN / AL 2020): https://aristeguinoticias.com/0204/mexico/vivimos-temporada-de-zopilotes-dice-amlo-sobre-opositores-y-prensa-amaril-lista-enterate/

meme of a vulture with Dresser's face. The metaphor implies that her critical reporting equates to vulture-like conduct (ARTICLE 19, 2021b). Dresser referenced macho and sexist culture, and argued that over the years, online violence was being sanctioned by authorities and "translated into violence in homes, beds, which forms the other pandemic we are living in" (Cerdeira, 2021a).

At the height of the pandemic, Pamela Cerdeira, radio host for the privately owned MVS Noticias (Mexico City radio station), received this online threat: "I'll rape you and I'll rape your daughter." She said that at that moment it seemed that the threat was not only in her telephone, but also in her head "that could not stop spinning", she wrote (Cerdeira, 2021a).

Yalina Ruíz Chino, a journalist at the Indigenous and Afro-Mexican news agency, Notimia, was sexually harassed via WhatsApp by an elected public official after she interviewed him about COVID-19. He sent her unsolicited nude images of himself via another channel, even after she had blocked him on WhatsApp. When she publicly shared the story of his inappropriate behaviour on social media, he threatened her via phone calls and text message, and insisted that she remove her posts about him (Jiménez, 2020).[86] She resisted.

The Programme for Freedom of Expression and Gender at the Mexican NGO CIMAC[87] has documented attacks against, and the stigmatisation of, journalists seeking to hold health authorities in charge of responding to the pandemic to account. During the early part of the pandemic, they recorded 29 cases of violence against women journalists: saying that seven were directly related to coverage of the health crisis, and that six of those took place online. Of the 29 cases, 14 involved documented alleged violence by public officials (García Martínez, 2020).

In Brazil, *Real Amazônia* journalist Kátia Brasil was harassed after publishing her report on the impact of the pandemic on Indigenous peoples living in isolation, and with limited access to information (ARTICLE 19, 2021c).[88] An Indigenous leader, affiliated with the Bolsonaro government, was dissatisfied with Brasil's reporting and she said "tried to pressure us in every way", attacking her in an hour-long Facebook livestream. In the video, he sought to discredit Brasil's reporting about the lack of infrastructure for combating COVID-19 in the Amazonas as false. Brasil said: "He kept calling us [my colleague Elaíze Farias and me] liars, bandits, and ugly women". She was considering legal action following this episode, despite receiving legal advice that she would likely be subjected to further "guerilla army" attacks if she proceeded. The man who instigated the attacks has previously been arrested for libel and defamation, and required to compensate his targets for 'moral damages' (Brasil, 2014).

---

86 In Yalina Ruíz Chino's case, eventually the Ministry of Environment and Natural Resources (Semarnat), where the perpetrator worked, published a statement condoning harassment, and measures were allegedly to be taken against the man (Jiménez, 2020). Ruíz Chino fears that when she goes back to her community, the case will not only remain unresolved but that it could worsen the situation for her or her family. Professional women are seen as bringing dishonour and shame upon their families in conservative, patriarchal and *machista* (misogynistic) communities.
87 Women's Communication and Information Centre (CIMAC).
88 The Vale do Javari is one of the largest Indigenous lands in Brazil where more than 5,000 indigenous people live, in the Atalaia do Norte region (more than 100 km from Manaus).

COVID conspiracy theorists as assailants

Online violence against UK journalists also further escalated during the pandemic, with the rise of viral conspiracy theories. Many British journalists interviewed spoke about experiencing peaks of abuse in association with reporting on the health crisis, in particular when covering the theme of disinformation connected to COVID-19.

BBC's Marianna Spring described increasing online attacks associated with her COVID-19 mythbusting work, which were also harder to deal with in the context of lockdown:

> At the moment I don't walk into the office and say to someone, 'Oh, so weird. I got an Instagram message for someone saying they want to kill me because I'm a Satanic pedophile who should be killed'… I get this kind of barrage of abuse from QAnon…'we're coming to get you'; 'I don't know how you sleep at night'; 'You're complicit in war crimes'

Spring said that a a "slightly unnerving" 2.5 hour interview she conducted with a 5G conspiracy theorist "who had been accused of shooting someone in the head when he was a bouncer", was counter-filmed by the interviewee's companions and uploaded to YouTube (BBC Panorama, 2021a). This led to another torrent of online abuse against the journalist for doing her job in "questioning the idea that the [COVID-19] vaccine is…part of some…mass genocidal plot". Spring was also physically stalked in connection with her COVID-19 conspiracy debunking reports (see further discussion below).

After she was interviewed for this study, Spring was commissioned by BBC Panorama to make a documentary about gender-based online violence. She focused on her own experiences, along with those of medical doctors and celebrities (Spring, 2021). For the documentary, University of Sheffield computer scientists and International Center for Journalism (ICFJ) researchers conducted a short-range big data case study analysis of approximately 75 thousand tweets directed at Spring from March to August 2021.[89]

This analysis identified three major abuse spikes in April 2021 connected to anti-lockdown protests in London, her commentary on the online abuse she received in the aftermath, and her reporting on a large Facebook takedown of an anti-vaxxer group with 100 thousand members.

---

89 This case study adopted the same methodology applied to the cases of Maria Ressa and Carole Cadwalladr featured in Chapters 3.1 and 3.2.

Figure 3: A tweet from BBC specialist disinformation reporter Marianna Spring cataloguing the abuse she experienced in response to her reporting on COVID-19 conspiracies in the context of anti-lockdown protests in the UK.

The big data analysis also showed that more than half (55%) of the abuse analysed was designed to discredit Spring as a journalist and undermine her public interest journalism, while 27% was sexist and misogynistic (the remainder was classified as generally abusive).

Like Maria Ressa and Carole Cadwalladr, whose case studies are highlighted in Chapters 3.1 and 3.2, Spring was most frequently branded a "liar". But she was also called a "presstitute", a "silly woman", a "nutter", and a "stupid c*nt". The top abusive accounts identified were all anti-lockdown, anti-vaccination conspiracy theorists who sent similarly abusive messages to other prominent accounts.

As discussed, BBC investigative reporter Rianna Croxford and former HuffPost UK reporter[90] Nadine White were both subjected to online violence triggered by a government minister with the portfolio for race and inequality, after seeking comments from her in the course of their reporting on COVID-19. Croxford was reporting on the disproportionate incidence and impacts of COVID-19 on black and ethnically diverse communities in the UK.[91]

The minister wrote a column for the *Mail on Sunday* newspaper, which was highly critical of BBC journalists in general, that was published alongside an attack piece on Rianna Croxford illustrated with a photo lifted from one of her social media accounts. "I remember feeling upset and really intimidated, especially as I was still at the start of my career," she said. This political demonisation of a junior journalist led to a prolonged campaign of racist and misogynistic online harassment, and a message of support for her story issued by the BBC.[92]

Doxxing, death threats and incitements to mobbing against BBC journalists reporting on the pandemic escalated further after an incident in June 2021, when political editor Nicholas Watt was chased and harassed by anti-lockdown protestors outside the British Prime Minister's residence. This incident led the

---

90 Now *The Independent*'s Race reporter.
91 Rianna Croxford was a community affairs correspondent when the minister contacted her superiors at the BBC to complain. She is one of only four black journalists to ask questions during the government's early pandemic press briefings (Tobitt, 2021c).
92 In June 2020, Rianna Croxford tweeted that she would be requesting payment from *The Mail on Sunday* "for using my photo without consent and without offering a right of reply before publication", which she would donate to charity. The BBC Press Office on Twitter defended "our story": https://twitter.com/The_Crox/status/1269564164487421953, https://twitter.com/BBCNewsPR/status/1269562339168305158

director of news and current affairs at the BBC to announce the establishment of a working group to address "a growing problem", acknowledging that women journalists, particularly those from marginalised ethnic groups, were more acutely affected (Townsend, 2021).

In Brazil, Gabi Coelho from daily *O Estado de S. Paulo* said she was seeing more attacks online during the pandemic from people who "say they are journalists, but [who] have no commitment to doing journalism." It is a "dangerous" situation, she said, leading to a proliferation of disinformation and hate speech online, highlighting the ways in which partisan media actors fuel disinformation-related attacks on women journalists in Brazil and elsewhere.[93]

Online abuse during the pandemic also continued to be more pronounced at the intersectional level. In the US, the *New York Times*' Apoorva Mandavilli was told by email "that she should have her head cut off in public" for articles she had written about COVID-19. Mandavilli said her white male colleagues are spared the sort of threats and vitriol she is subjected to. Journalist Carl Zimmer, who covers the same beat, tweeted support for Mandavilli, confirming that he was not exposed to the same sort of threatening direct messages.[94]

Mandavilli said she understands that the online outrage and the anger directed at her for doing her job is not really about her: "I'm hopeful that when some of that anger subsides, that they will stop also being nasty to me on Twitter.

Figure 4: Screengrab of a tweet by the NYT's Global Health reporter Apoorva Mandavilli asking for civility from Twitter users especially during COVID-19.

Sexual harassment by news sources increased in Kenya during the pandemic, despite less physical contact due to COVID-19, according to journalists interviewed for this study and parallel research. Some women journalists said that sources were using the pandemic as an excuse to request "private" meetings with them instead of conducting interviews in public spaces. For others, the abuse escalated online.

In 2020, the Media Council of Kenya identified eight cases of COVID-related press freedom violations involving online attacks on women journalists, attributed in part to a rise in cyberbullying during the pandemic in the context of lockdown policies and increased online activity (Media Council of Kenya, 2020). Kenyan women journalists also described being stigmatised and shamed in broader

---

93 See Chapter 2.4 for discussion about partisan news media attacks on independent women journalists.
94 Tweet from *New York Times* journalist Carl Zimmer in February 2021, recognising his colleague - a woman journalist of colour - received more online abuse: https://twitter.com/carlzimmer/status/1365075035355021315?s=20

society for testing COVID-positive, in a context where the country's first COVID-19 patient was doxxed and abused.[95]

In Pakistan, Geo TV reporter Benazir Shah was harassed and abused online over a prolonged period - "especially [by] government officials" which triggered pile-ons - for investigating the government's response to COVID-19. She said she had to shut down her Twitter account because of the pile-on,[96] and she was threatened with a defamation lawsuit after she took her complaint to the National Assembly.[97]

Women journalists lodging official complaints about online violence are sometimes exposed to legal harassment in response, she said. In this climate, Shah signed an 'anti-digital violence' statement[98] along with 165 other Pakistani journalists in 2020 which was largely met with silence, she said (CFWIJ, 2020).

The pandemic has exacted a terrible toll on journalists worldwide - both personally and professionally (Posetti, Bell and Brown, 2020).

ix. Investigative reporting as a lightning rod for attacks

Fourteen percent of survey respondents (n=100) said investigative journalism prompts high levels of online abuse.

Carmen Aristegui, investigative journalist and founder of Mexico's Aristegui Noticias, has experienced years of harassment directed towards undermining her privacy and personal reputation through coordinated Twitter attacks designed to destabilise and discredit her, as well as surveillance attempts against her and family members and colleagues (Forbidden Stories and Amnesty International, 2021).

As discussed above, Aristegui was one of 19 people in Mexico targeted by Pegasus software[99] installed on her personal phone - allowing her microphone to be used and for her to be spied on between 2015 and 2017 (R3D, 2017). She said she began receiving messages with suspicious links after she published an investigation on former Mexican President Enrique Peña Nieto. In 2020, Aristegui was also targeted for months online via the hashtag #ApagaAristegui (#TurnOffAristegui) after she published an article about illegal work activities connected to Notimex, the Mexican state press agency (Arteta, 2020).

The attacks on Aristegui were noted by the Office of the InterAmerican Special Rapporteur for Freedom of Expression (IACHR, 2021a). In November 2021, the first arrest was made in Mexico (Sorto, 2021) for "illegal intervention of communications aggravated to the detriment of a journalist, using the software known publicly as Pegasus".

The big data case studies on Maria Ressa and Carole Cadwalladr (Chapters 3.1 and 3.2) also highlight the links between online violence and investigative reporting.

95 The case of Brenda Cherotich (Maundu, 2020a; Freedom House, 2020). A nurse and patient were also harassed online and doxxed in the pandemic (Mutahi, 2020).
96More about Pakistani women journalists calling for State action: https://cpj.org/2020/09/as-ruling-party-fans-spew-online-abuse-pakistans-female-journalists-call-for-government-action/
97 Benazir Shah's tweet from 2020 of her official complaint to the National Assembly, about online abuse triggered by MP responses to her reporting on COVID-19: https://twitter.com/benazir_shah/status/1296111405503053826
98 The campaign against "vicious social media attacks, doxing, and hacking attempts by accounts affiliated to the ruling party and conservative, right-wing elements in the country" was accompanied by the Twitter hashtag #AttacksWontSilenceUs: https://womeninjournalism.org/cfwij-press-statements/pakistan-cfwij-supports-the-new-statement-on-digital-violence-against-women-journalists-and-demands-from-all-political-parties-to-take-swift-action-against-online-violence
99 Forbidden Stories, part of an international consortium which investigated the Pegasus Project, describes this as a technology sold by the Israeli company NSO Group to governments around the world.
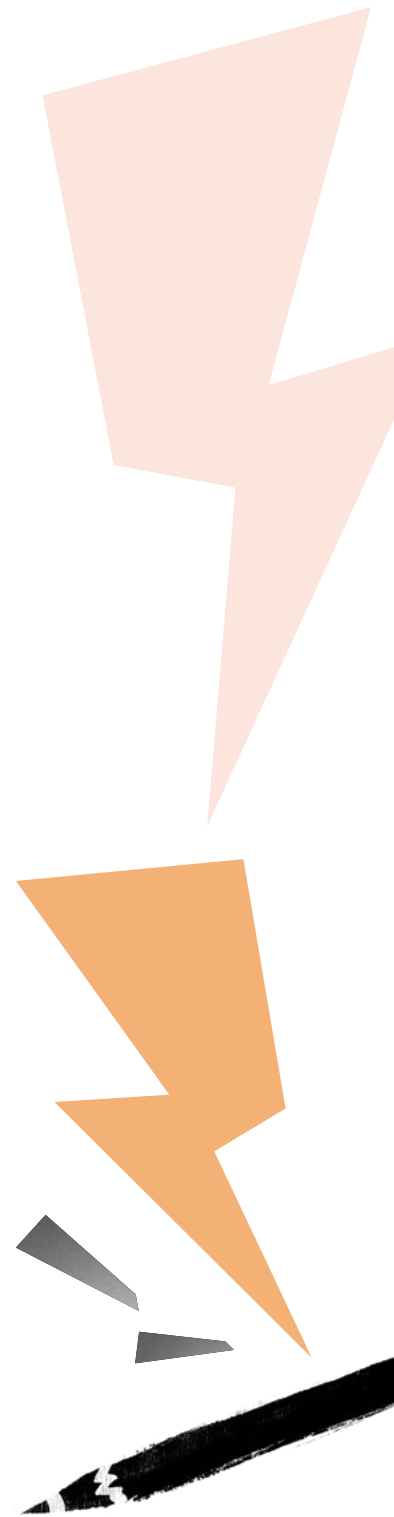
2.4 Impacts of online violence on women journalists and their journalism

"To be a victim of harassment is a traumatic experience. It takes away your peace of mind ...it fills you with doubts... We have been taught to be a dignified victim, and a stoic woman [who] knows to be silent [and] accepts as routine all the insults towards her body... [But] we will not shut up."

Mexican journalist Alma Delia Murillo, columnist for Reforma.

i. Psychological warfare

UN Special Rapporteur for The Right to Freedom of Expression Irene Khan has said: "Online violence against women is not about speech, it's about violence, and violence actually means having an impact on women's health, psychological status, and on their lives" (APCNews 2021). In other words, online attacks inflict real harm. According to our research, impacts include depression, anxiety, stress related ailments, Post Traumatic Stress Disorder (PTSD) diagnoses, damaged careers, and absences from work.

Over a quarter (26%; n=186) of the women journalists surveyed for this study identified mental health impacts as the most significant consequence of online violence that they had experienced. A substantial number (12%; n=86) also said they had sought medical or psychological help in response. And 11% (n=79) said they had taken time off work to recover.

Additionally, several of the women journalists interviewed experienced significant psychological injury, including PTSD, requiring medical intervention and time off work. **Many of them broke down while discussing their experiences with the researchers, describing how important it was to them to be listened to.**

"There were days when I couldn't sleep, couldn't eat. Because to file the complaint, I had to read a lot of ugly things with all the hate online, in addition to the death threats. I couldn't speak, I couldn't work," said Reyna Haydée Ramírez, a journalist for the Mexican news site Pie de Página. She was subjected to a 'pile-on' after participating in televised press conferences with the president. A similar account came from Brazilian journalist Kátia Brasil, who "spent a few days without sleep, shocked by [my] situation...feeling impunity, feeling that desire to scream," after she was attacked by a public figure on Facebook in connection with her reporting.

MENTAL HEALTH IMPACTS WERE THE MOST SERIOUS CONSEQUENCE OF ONLINE VIOLENCE **FOR**

26% *OF THE WOMEN JOURNALISTS SURVEYED*
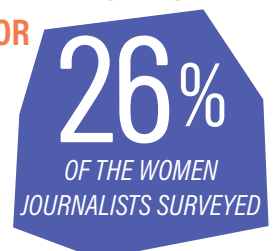
At the time of her interview for this study in early 2021, former *New York Times'* journalist Taylor Lorenz said multiplatform harassment had resulted in: "Weeks where I can't leave my bed and can't function and I'm crying all day, or throwing up all day because of the anxiety and stress that it causes. And there was one point where I definitely didn't want to even live any more".

Online violence has been directly linked to suicide (Dooley and Hida, 2021). Polish journalist Natalia Żaba found she had trouble "with simple things like paying my bills…I understood that the level of violence I am experiencing every day, whether it's offline, online, doesn't matter. You know, it's pretty [much] the same when it comes to how I feel, and how my body reacts. It's just unacceptable. And it has to change, and we have to start talking about it."

12%

OF SURVEY RESPONDENTS SOUGHT MEDICAL OR PSYCHOLOGICAL HELP DUE TO ONLINE VIOLENCE

Kenyan radio journalist Jeridah Andayi has said that a week-long endurance of a multiplatform social media harassment campaign took all the energy out of her. She described struggling whilst on air: "I could not say a thing. I was breaking down… I felt so violated, like I had been stripped and beaten in public" (Maundu, 2020b). Andayi had posted a Facebook photo which garnered almost 3,000 trolling messages, and which she later felt compelled to delete. She decided against blocking trolls or leaving Facebook after speaking to a pastor, who told her he was praying for her and to ignore "those people, they don't know you".

However, in many cases, such advice may not be any more effective than the newsroom lore (frequently dispensed by men in leadership roles) to 'grow a thicker skin'. Liz Orembo of KICTANet[100] said that many women have developed resilience because of continued trolling (KICTANet, 2021).

But an over-emphasis on the requirement for the targets to develop resilience can cause the need for protective and defensive systems, and the imperative to end impunity for trolling to be downplayed.

Cultural taboos can inhibit interventions

Mental health struggles can be taboo in some contexts, making the experience of dealing with online violence impacts harder for a number of interviewees. "Seeking help for mental health here in Kenya is looked down [upon]," pseudonymous TV journalist 'Sarah Jane' said. "So we find that people don't seek help, they deal with it individually."

Kiki Mordi, a Nigerian freelance journalist and founding member of the Feminist Coalition, received online threats after the airing of her BBC documentary, Sex For Grades,[101] which exposed a culture of male lecturers sexually harassing female students at the University of Lagos.[102] In early 2020, the online violence Mordi frequently experienced morphed into threats of physical violence, "one after another", after she defended a woman who had accused a pop star of rape. She

---

100 Kenya ICT Action Network (KICTANet), a multi-stakeholder platform for people and institutions interested and involved in ICT policy and regulation.
101 A feminist collective in Nigeria which campaigns on a range of issues.
102 https://silverbirdtv.com/education/64598/aftermath-sexforgrade-bbc-journalist-kiki-mordi-receives-death-threats/: https://www.bbc.co.uk/news/av/world-africa-49907376/sex-for-grades-undercover-in-west-african-universities

said she had to "pause work" since the online abuse battered her mental health: "I hate that that's something we have to... try to get used to," she said.

Women journalists in the Arab States also described the need for a culture shift, so that they do not feel compelled to endure online abuse and suffer in silence, especially considering what they described as a stigma surrounding mental health issues in their contexts. Fadoia Chtourou, a journalist working for Tunisian TV, who was harassed for not wearing makeup on air, and had a "delayed reaction" to the smear campaign she experienced, shared a familiar pattern among the interviewees. While some of them described the value of support from close family and friends, others said the withholding of such support had exacerbated the pain.

Tunisian editor Khaoula Boukrim said her mother and boyfriend told her "'don't be weak, don't be too sensitive'...and it hurts, it makes me feel that they can't understand how I feel. I know that they say that because they love me, and want me to be strong, but I am human, I get affected."

Some interviewees said they had decided to be open about the mental health impacts of gendered online violence - recognising that talking about it would help bring the impacts into the light. But this was weaponised against them at times. Serbia's Jovana Gligorijević has written about having her looks mocked; being called a "frustrated whore" and a "communist MILF", and threatened sexually. She continues to be targeted for speaking out about the abuse she experiences, with online 'jokes' from trolls suggesting she might kill herself (Gligorijević, 2019).

ONLINE VIOLENCE
MADE
17%
OF WOMEN JOURNALISTS SURVEYED
FEEL PHYSICALLY UNSAFE

Fear of physical attack

The fear and anxiety connected to the threat of physical attack associated with online violence is also relevant. 17% (n=121) of survey respondents indicated that they had felt physically unsafe as a result of online violence, while 29 of them reported missing work due to the potential for the online attacks on them to morph into physical attacks.

For almost a year, Verónica Espinoza from Mexico's *Proceso* magazine was falsely accused by trolls of "being a spokesperson for a criminal organisation," and labelled a "feminazi".[103] She described the multilayered effects: "Digital violence has been risky and dangerous for my integrity...I believe that this could escalate to other types of aggressions, or that it could make me censor my work. There is an emotional toll that has lasted for a while because of the risk...."

It is also important to note the additional exposure to psychological injury associated with attacks on women journalists who experience intersectional abuse. For those women, the compound nature of the online violence they experience is correlated with increased offline impacts on their wellbeing. Women journalists from minority backgrounds working in traditional newsrooms spoke of facing far more significant career hurdles, and greater expectations for emotional resilience.
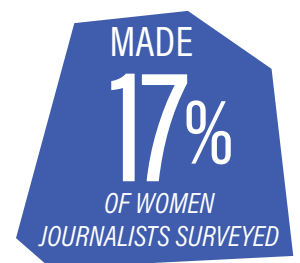
103 "In Guanajuato, there is a conflict between warring criminal groups who use women journalists as cannon fodder," said Verónica Espinoza from Mexico's *Proceso* magazine. "The criminal gangs spread misinformation stating that a journalist supports one side or the other. To cover the story is to explicitly expose ourselves to the conflict."

ii. Chilling professional impacts

The professional toll of online violence against women journalists represents another dominant theme emerging from the survey and interview data. This has significant implications for diversity and plurality in (and through) the news media, not just for individual women's careers. Apart from the aforementioned impacts on work attendance, one in ten (10% n=71) survey respondents said their professional reputations or employment had been affected. Career impacts caused by online violence were severe in some cases.

For example, 29 respondents said they had quit their jobs as a result of online violence, while 14 indicated that they had left journalism for good - as a direct consequence of the online attacks they had endured. While these numbers might appear small, any number of women journalists quitting their jobs or leaving the profession due to gender-based online violence is a disturbing statistic.

The increasing risk of offline attacks associated with online violence also caused several of the interviewees to quit their jobs or re-evaluate their careers. "The day I decided to stop writing was the day I was sent a picture taken from inside my house; that's when I decided that it was not worth it," said former investigative journalist Myra Abdallah in Lebanon.

Concern about the risk of offline harm connected with online attacks was echoed by Lebanese TV presenter Dima Sadek, a journalist for over two decades, who is re-evaluating her career.[104] "When you feel that if anything should happen to you, no one will protect you, you start thinking why have I been doing this for 10 years? Do I want my daughter to grow up without a mother?" But she added that she does not want to give up, even if she has limited support: "I have to do what I have to do, I have to say what I believe in."

Northern Ireland's Leona O'Neill abandoned her 23-year career as a journalist after being diagnosed with PTSD. This followed what she described as a global social media hate campaign against her and her family[105] in the aftermath of the fatal shooting of fellow journalist Lyra McKee, which she witnessed in 2019 (López Beltrán, 2021).

In South Africa, Qaanitah Hunter quit her reporting job at the *Sunday Times* after a death threat sent to her in the form of a picture of a gun. She suffers ongoing mental health impacts but continues her journalism career as deputy editor of the political desk at the online publication, News24. However, she expressed concern about self-censorship triggered by a fear of backlash.

NBC's Brandy Zadrozny said, while she "loves" her job, she said at the time of her interview that she felt compelled to devise an exit strategy and did not see herself covering the disinformation beat in five years' time: "I'm not in a war zone, I'm behind a computer. But the effect of online harassment and bad faith journalism from disinformation agents...is meant to silence us, meant to stop our reporting, meant to scare us." Several months after this interview, Zadrozny said publicly

---

104 In a survey of 250 Lebanese, Arab, and international female journalists working in Lebanon from 2016, 90% had considered quitting their jobs due to gender discrimination and sexual harassment (Melki and Mallat, 2016).
105 See discussion of the threats and abuse directed at O'Neill and her family in Section 2.5 below.
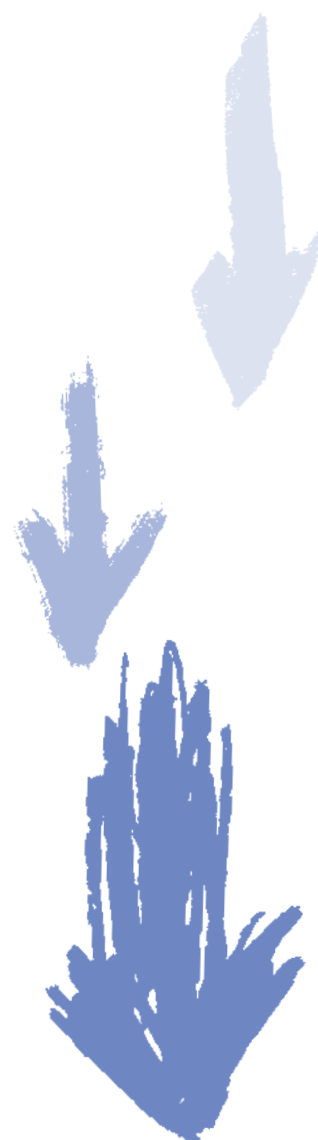
that she had changed her mind about this, and expressed determination to stay with her investigative disinformation beat.

iii. The invisibility cloak: costs of self-censorship, retreat and erasure

Many women journalists are clearly being silenced and retreating from view in response to being targeted online. For example, 10% (n=71) of our female survey respondents had asked to be taken off air, or have their bylines removed, after being attacked.

This pattern of responding to online violence through avoidance was also evident in impacts on the reporting practices of the women surveyed and interviewed for this study. This included negative effects on the stories and sources pursued. For example, 19% (n=136) of survey respondents said they avoid particular beats, and 10% (n=71) said they avoid certain stories because of the likelihood that they will expose them to online violence. Finally, 15% (n=107) of survey participants said that they avoid particular sources or contacts because of their history of abuse. These responses can be understood as a form of self-censorship that chills reporting.

Several interviewees said they had asked, or been asked by their editors, to change beats to focus on less incendiary issues, in the interests of their physical safety and mental health which had been directly affected by prolonged exposure to online violence triggered by their reporting.

Julia Carrie Wong left the technology beat to take on a new role at Guardian US; being made a target for online violence perpetrators on the far-right was instrumental in her decision. "The consequences of reporting on the far-right are real, and they take a toll," she said.

A number of interviewees had also given bylines away to colleagues on major stories to avoid being the target of online attacks in the aftermath. Brandy Zadrozny said she handed the main byline on a big story to a male colleague - who was not experiencing the same abuse and threats - to deflect some of the harassment that would inevitably come her way. Another interviewee, a black journalist in the UK, was encouraged by her editors to hand a major story to another reporter to avoid the avalanche of abuse she was likely to attract if it was published in her name.

"But there's no escape from [it], and this is my career and I don't know how to escape this career that I spent my entire adult life building...This beat is a beat that I made," former *The New York Times*' journalist Taylor Lorenz[106] said.

Self censorship: the "fear that shuts my mouth"

Polish radio journalist Aleksandra Wojciechowska now chooses her stories "more wisely", she said, saying she has "lost confidence" as a reporter.

19%

OF WOMEN JOUNALISTS SURVEYED AVOIDED CERTAIN BEATS WHICH LEAD TO ONLINE VIOLENCE

---

106 Lorenz moved to the Washington Post in January 2022.

I am afraid of people's reaction to my work. I'm afraid of any mention of my name in the media. This is not the kind of fear that shuts my mouth - because I enter the studio, read the news, I still work in the media... I have that fear in the back of my head all the time.

Self-censorship was also occasionally evident at the institutional level. Polish editor Alicja Cembrowska, Editor-in-Chief of a web-based, women-oriented portal said she "let go" of stories where the fear triggered by publishing was just too great. Her magazine "filters" content "quite significantly" and tries "to tone down the topics that ignite the public" – such as coverage of abortion, refugees and migration, and LGBTQ rights, she said.

In Brazil, in the face of racist and misogynistic hate speech, certain news media outlets and reporters self-censor, use coded language, and choose not to report particular events to avoid conflict, according to research for this study. For example, Brazilian freelance journalist Ana Freitas noted that online violence "greatly restrained my willingness to talk about topics [...] to avoid fatigue".

Several interviewees feared that the phenomenon of online violence against women journalists would deter young women from a career in journalism. "If a popular anchor is trolled and shamed, students will be like, 'I want to be an anchor but I do not want to go through that'," said pseudonymous Kenyan TV news editor 'Sarah Jane'. "People are seeing us doing nothing [about it]...we complain that the media is male-dominated but we see women leave." This perspective was also echoed in the UK by a senior political journalist.

Disengagement impacts diversity and democratic deliberation

Making themselves less visible on social media was another instinctive self-defence mechanism for a significant number of the women surveyed and interviewed. Nearly a third (30%; n=214) of female survey participants said that they self-censor in online communities as a result of being targeted in online violence campaigns. One fifth (20%; n=143) said they avoid all interaction and only 'broadcast' to their followers, while 18% (n=129) said they avoided engaging with their audiences in particular. The same number said they had withdrawn from social media for a period, while 11% (n=79) said they had permanently withdrawn from certain online communities.

Although such acts of withdrawal and concomitant self-censorship could be understood as strategic defensive moves, there are significant professional impacts in a profession where visibility is often essential to promotion. They also demonstrate the effectiveness of online attack tactics - designed to chill critical reporting, silence women, and muzzle truth-telling. Additionally, there are significant gender equality and freedom of expression implications associated with the silencing of women journalists, or decreasing their visibility, and these

SELF CENSORSHIP ON SOCIAL MEDIA WAS ACKNOWLEDGED BY

30%

OF THE WOMEN JOURNALISTS SURVEYED

have impacts on media diversity and gender representation in the news (Global Media Monitoring Project, 2020).

When women journalists' voices are silenced in this way, so are their sources'. Press freedom is threatened, and democratic deliberation is eroded. Soraya Chemaly, author and former Executive Director of the Representation Project in the US, highlighted the broader cost of such self-censorship on diversity in and through the news media. "Women journalists have stepped back from covering issues that they are deep experts in, that they are interested [and that] they would have valuable insights into," she said. As women reporters reduce their public engagement on social media and other platforms in response to online violence, their potential contributions to democratic discourse are diminished too.

2.5 When online violence spills offline

A pattern emerged from this research, connecting online violence campaigns and offline attacks. It is evident that online abuse and physical attacks against women journalists are not only correlated, but causally related in many cases. This suggests that online abuse can be a predictor of physical violence - including murder with impunity, as discussed below.

Aside from inflicting very real psychological injury,[107] targeted online violence poses increasing physical safety risks, including mob reactions. Women journalists interviewed for this study had death threats pinned to trees next to their work, messages left on notice boards at train stations on their commute, letters sent in the post with serious threats enclosed, car tyres slashed, and apartments ransacked. They have experienced being assaulted, ostracised in their local communities, and physically stalked.

The online-offline violence trajectory represents a vicious and self-perpetuating circle. Online harassment and threats beget offline attacks; and offline abuse (e.g., political actors targeting women journalists during public appearances) can trigger an escalation of online violence which, in turn, can exacerbate offline risks.

One fifth (20%; n=143) of the women journalists surveyed reported experiencing harassment, abuse and attacks in the physical world that they believe were seeded online. This finding is particularly disturbing, given the emerging correlation between online attacks and the actual murder of journalists with impunity (Witchel, 2017; Posetti, 2020a).

Women in some ethnic groups and regions also appear more likely to experience offline attacks connected to online violence according to this research. For example, 53% (n=53) of Arab-identifying women journalists[108] who responded to the UNESCO-ICFJ survey reported enduring such attacks - a rate more than double the global figure recorded.

Sometimes, analogue death threats have followed online attacks. Mexican TV host and reporter for *Diario de Quintana Roo* Cecilia Solís Martín has been subjected to intimidating messages, online abuse, surveillance, and smear campaigns since February 2021, after she covered an anti-femicide protest (ARTICLE 19, 2021a). Solís Martín's car has been vandalised, and her relatives have also received threats. Also in Mexico, in 2015 *El Universal* columnist Maite Azuela received a death threat in the post in the form of a defaced photo of her (Redacción AN, 2015; ARTICLE 19, 2015).

Around the world, many of the women journalists interviewed for this study said they had resorted to changing their patterns of movement, going into hiding, relocating, and even moving abroad in response to online violence. Several also described increasing their physical security in connection with online attacks,

1/5

WOMEN JOURNALISTS SURVEYED

LINKED OFFLINE HARM THEY'D EXPERIENCE TO PRECIPTATING ONLINE VIOLENCE

---
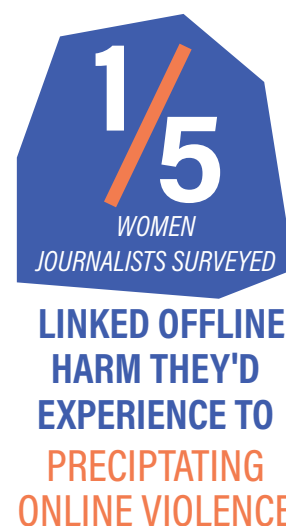
107 See Chapters 2.4 and 4.0 for further discussion about psychological impacts.
108 This group is defined according to ethnicity, not location, so the respondents are not necessarily based in the Arab States.

highlighting their sense of vulnerability, and their awareness of the potential offline consequences of digital attacks.

Then, there is the phenomenon of online violence becoming part of an enabling climate for what Filipino-American journalist and CEO Maria Ressa refers to as 'lawfare' - targeted law enforcement investigations and legal action against women journalists designed to silence them.[109]

In a joint statement issued on the International Day to End Impunity for Crimes Against Journalists in November 2021, three UN special rapporteurs called on States to investigate and prosecute online violence against journalists, linking the phenomenon to murder with impunity:

> *The failure to investigate and address attacks online has real-life consequences for women journalists, affecting their mental and physical health, undermining their confidence and autonomy, stigmatizing them and generating fear, shame and reputational and professional damage. In extreme cases online threats can escalate to physical violence and even murder, as the killing of Daphne Caruana Galizia showed* (OHCHR 2021a).

i. Correlating online threats of sexual and physical violence with offline attacks and impunity cases

In October 2017, the Maltese investigative journalist Daphne Caruana Galizia was killed when a bomb placed under her car detonated close to her home. Before she was murdered, Caruana Galizia endured frequent online threats and spoke about being called a "witch." There were clear gendered aspects to the intimidation she suffered, and the impunity that surrounded those threats and preceded her killing.

The parallels between the State-sponsored harassment experienced by Maria Ressa and the pattern of online violence associated with Caruana Galizia's death prompted the murdered journalist's sons to issue a public statement expressing their concerns. "This targeted harassment, chillingly similar to that perpetrated against Ressa, created the conditions for Daphne's murder," they wrote (Daphne Caruana Galizia Foundation, 2020).

The propensity for misogynistic online attacks on women journalists to morph into offline attacks is particularly serious in countries like Mexico, where femicide levels in general are extremely high (SSPC, 2021; UN Women, 2017), in addition to the serious risks faced by journalists in general (RSF, 2021l).

Mexican journalist María Elena Ferral was attacked for years, and sent death threats by a former local congressman, a former mayor, and the former local member of the Chamber of Deputies. She was targeted by fake Facebook accounts

---

109 Read more in the Maria Ressa and Carole Cadwalladr big data case studies in Chapter 3.0, and Chapter 6.0.

and falsely linked to organised crime in response to her critical reporting about the Poza Rica City Council over a three year period, and also she endured two physical attacks in 2006 and 2012 (Santiago, 2020; CIMAC and ARTICLE 19, 2020).

When Ferral wrote about corruption, she was attacked on social media with threatening comments like "I wish you were dead" and "I wish you were kidnapped" (ARTICLE 19, 2020a).

Between 2016 and 2018, Ferral reported the threats from the former mayor to the Office of the Special Prosecutor for Crimes against Freedom of Expression (FEADLE), and tasked a colleague with making a handful of emails public in the event of her assassination (PEN Canada, 2020). In 2018, she was also abused and threatened online by the son of a mayor who has been implicated in a series of targeted online attacks against journalists.Ferral was murdered in 2020. She was shot by a motorcyclist at least six times as she headed to her car in broad daylight in Veracruz. Months later, her daughter, a local news journalist, was also attacked.[110] This case powerfully underlines the links between online violence and deadly offline attacks in Mexico.

The 2017 murder of Indian investigative journalist Gauri Lankesh drew international attention to the risks faced by women journalists who are subjected to online violence while being openly critical of governments and politicians (Romig, 2019). Lankesh, who was shot dead outside her home, was known for being a critic of right-wing extremism and Hindu nationalism, and she was subjected to significant online abuse before her death. In the days after Lankesh's killing, trolls took to social media to celebrate, describing her as a "bitch."

Pointing to the emergence of a pattern, the case of another Indian journalist – Rana Ayyub – led five UN special rapporteurs to intervene in 2018 following the mass circulation of disinformation-laced attacks designed to shame her, heighten the risks she faced offline, and counter her critical reporting (Taskin, 2020; OHCHR, 2018). The independent journalist was on the receiving end of deepfake videos, as well as direct rape and death threats.

Rana Ayyub
Picture credit: Mohit Varu

The UN experts issuing the statement in defence of Ayyub pointed to the murder of Lankesh following online attacks, and called on India to act to protect Ayyub, stating: "We are highly concerned that the life of Rana Ayyub is at serious risk following these graphic and disturbing threats." The threats against Ayyub have escalated dramatically since 2018.

The *Washington Post* columnist wrote in January 2022 that she had received 26,000 Twitter-based threats on one day (Ayyub, 2022). As these attacks worsened, two UN Special Rapporteurs issued a new statement in support of Ayyub, stating "Relentless misogynistic and sectarian attacks online against journalist Rana Ayyub must be promptly and thoroughly investigated by the Indian authorities and the judicial harassment against her brought to an end at once" (OHCHR, 2022).

The trend of online violence spilling offline with potentially deadly consequences is also evident in the US. The man who killed five staff members of the Capital

---

110 Fernanda de Luna Ferral, editor of the website El Quinto Poder based in Papantla, northern Veracruz, "continued editing the website's Facebook page after her mother was murdered": https://cpj.org/2020/05/gunmen-attack-daughter-of-murdered-mexican-reporte/. UNESCO Director-General Audrey Azoulay condemned the assassination (Deutsche Welle, 2020; UNESCO, 2020).

Gazette in Annapolis, Maryland in 2018[111] had previously harassed the paper's staff by email and Twitter and used Facebook to stalk them. He had targeted the newspaper and its staff after they covered a case in which he was found guilty of harassing a woman online (Robertson, 2018).

When the press is demonised, women journalists are prime targets. The Grio's April Ryan made national headlines in 2017 when the White House Press Secretary told her to "stop shaking [her] head", and inferred she was pursuing a political agenda:

The way they talked to me and viewed me and described me led into the threats, the bullying...be it online, be it letters coming to my home, be it emails...and be it people standing outside of the White House waiting for me to come out. Be it people coming to my old house and sitting outside looking, waiting for me.

Ryan said the Capitol Hill insurrection in the US on January 6, 2021 "was basically a realisation of all threats that everyone has been getting".[112] In 2018, she was sent a bomb threat on Twitter by a fervent Trump supporter who is now serving a 20-year jail sentence for sending explosive devices to CNN and high-profile Democrats; he threatened that she and a colleague would be next. Ryan, who also received threats about being shot by a sniper on her way home from work, is now even frightened about going to the grocery store. "People recognise me... I don't like shopping unless I go early in the morning [so] there's no one there, or very late at night. My life has totally, completely changed."

Another White House correspondent, Al Jazeera's Kimberly Halkett, was the subject of a death threat sent to the White House Correspondents Association, triggering police and Secret Service involvement. But she said she was "scared - in the United States - of being shot", despite the extra security:

Given the fact that the White House has had this battle with the media for the four years with the Trump administration that we had been called 'fake news'... I was very worried when I walked from the White House gates to my car that somebody was going to shoot me, and that I was going to be that person that would be representative of the media. It was terrifying.

Online violence a serious threat in post-conflict Northern Ireland

Sunday World reporter Patricia Devlin in Northern Ireland has also received credible death threats. In February 2020, on a day during which she had been trolled by

---

111 The attack claimed the lives of five US journalists, including two women. The killer had railed against the journalists in tweets and sent them death threats, which he eventually acted upon: https://www.nbcnews.com/news/us-news/woman-harassed-capital-gazette-shooting-suspect-says-he-tormented-her-n888196
112 https://twitter.com/VICENews/status/1412437077917159430

"about 30 accounts set up to abuse", her name was graffitied on three walls in East Belfast next to a picture of rifle-sight crosshairs. Devlin drew a direct line between the online and offline threats: "I do believe that the online violence against me has created real life threats that are inciting criminals and very dangerous loyalist paramilitaries to issue threats against me. And that's put my life in grave danger".

Apart from receiving multiple online death threats, and threats to rape her baby, Devlin has also been formally warned by the police that she was facing imminent assassination and advised not to report from a neighbourhood in the city.

The *Irish News*' award-winning security correspondent Allison Morris, who was branded an "MI5 agent", also had her name graffitied on Belfast walls.[113] The National Union of Journalists, Deputy First Minister of Northern Ireland Michelle O'Neill and the chair of the all-party group on press freedom all responded in support of Morris (Irish News, 2021).

Figure 5: Screengrab of a tweet from Deputy First Minister of Northern Ireland Michelle O'Neill.[114]

In the same month, yet another woman journalist from Northern Ireland, Leona O'Neill, was branded an "M15 tout" and a "shit stirrer"[115] in the same neighbourhood in which journalist Lyra McKee was killed during riots in 2019 (Westcott, 2019, Council of Europe, 2020). In online attacks, O'Neill was falsely accused of fabricating her eyewitness account of McKee's murder, and she was also subjected to conspiracy theories that she was responsible for it (McDonald, 2020). Being told in "hundreds of messages" online that she should be "attacked, stabbed, arrested, set on fire, that my children would burn in Hell, that I was a liar", morphed into offline threats (ibid.).

In June 2021 she tweeted about enduring an eight month-long campaign of abuse: "Some malicious and dangerous individuals were running actual funding pages to buy weapons to hurt me and my family with - and getting donations! Police and Guards said they could do nothing…astonishing."[116]

---

113 This highlights the role of high-level collaborative condemnation in efforts to combat online violence against women journalists.
114 https://twitter.com/moneillsf/status/1362756080883884032
115 Tweet from Leona O'Neill in February 2020 with a picture of the abuse: "I am a journalist, working in my city trying to provide for my family. I consider this a threat to my safety. I call on community leaders to help me get this dangerous slur removed now": https://mobile.twitter.com/LeonaONeill1/status/1224768561471725569
116 Tweet from Leona O'Neill, June 2021: https://twitter.com/LeonaONeill1/status/1407252820433252352.

Explicit death threats a prime indicator for offline harm

Many other journalists interviewed for this study had received death threats online. Among them was Al Jazeera's Ghada Oueiss, who told the researchers about a US $50,000 'contract' for her kidnapping or murder which was issued on Facebook by a Syrian government supporter based in Lebanon. She called the police and the perpetrator was arrested. Incidents like this increase the physical danger she faces as a woman journalist working in the Arab States, she said, and this has led to her decision to avoid travelling as much as possible.

In 2020, in the midst of the COVID-19 pandemic, Serbian journalist Jovana Gligorijević also received a death threat on her personal Instagram account. This was a direct and unambiguous threat, and the perpetrator was arrested shortly after she reported the threat. "His argument was that as I didn't post enough journalistic content on my Instagram account…[so] it was not clear enough that I was a journalist," Gligorijević wrote to the editors in a follow-up email after the alleged perpetrator had been released (see also Kristic, 2021).

ii. Impunity cases used as a threat

Impunity regarding online violence aids and abets impunity for physical attacks on journalists, and vice versa. There are instances of women journalists around the world receiving online death threats in the form of references to journalists murdered with impunity.[117]

For instance, Patricia Devlin received a death threat warning her that she would "end up like" *Sunday World* investigative journalist Marty O'Hagan,[118] who was assassinated in Northern Ireland in 2001. Media freedom lawyer Caoilfhionn Gallagher KC noted a similar incident involving the BBC Persian language service journalists she represents. They have received threatening messages with photographs depicting Iranian journalist Ruhollah Zam, who was renditioned and executed in Iran 2020 (OHCHR, 2020; BBC, 2021a).

In a case of online violence which "shook" her, Brazilian journalist Talita Fernandes received a death threat on social media in the form of being tagged in a well-known dictatorship-era image of a journalist hanging in his cell, which said "a good journalist …doesn't ask a stupid question".[119] Three nights later she was woken by rapping on the door by an unknown man who accused her of making noise, although she had been sleeping. Fernandes connects the offline incident, which she reported to the police, to the online threat of the photo of the dead journalist on which she was tagged: "You start to get paranoid [am I] being chased," she said. "Social media…postpones and extends a terror that you [experience] in person, it [has] physical effects."

---
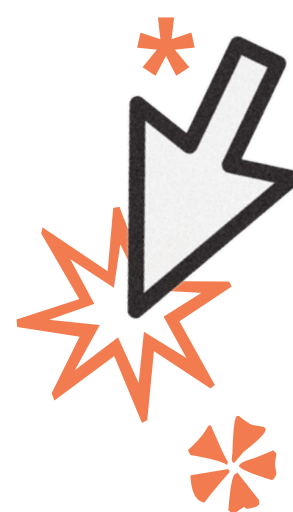
117 See further discussion in Chapter 5.0.

118 Since the assassination of Martin O'Hagan near Belfast in 2001, the first killing of a journalist in the line of duty in the UK was Lyra McKee in Derry in 2019. No one has been convicted of O'Hagan's murder (RSF, 2020). No one has ever been convicted of O'Hagan's murder, which provides a chilling parallel to the online anonymous abuse against Patricia Devlin, employed later at his newspaper, the *Sunday World*.

119 Vladimir Herzog, a critic of the military regime, died in 1975 after being tortured by military intelligence officials, who tried to frame it as suicide: https://www.dw.com/en/brazil-charges-6-ex-military-dictatorship-agents-in-journalists-murder/a-52826777

When Mexican journalist Soledad Jarquín Egda – the mother of María del Sol Cruz Jarquín, a photojournalist who was murdered in Oaxaca in 2018 – kept up the pressure regarding investigations into her daughter's death, she faced constant intimidation, harassment, and death threats, much of it coming via social media (Frontline Defenders, 2019).

In Malta, editor Caroline Muscat has been threatened on Facebook with images of her murdered friend Daphne Caruana Galizia accompanied by messages like "She deserves some bombs too" (Geerdink, 2018).

iii. Protests and punches

Physical violence against women journalists does not occur in a vacuum, but in a wider climate of legitimation (and legal inaction) against private actors using force to achieve their goals. This is a climate in which social media companies are especially implicated in association with mob violence - online and offline.

Two US interviewees and one Brazilian interviewee said they had been physically assaulted in the course of their work. The US journalists were separately investigating radicalised supporters of former US President Donald Trump. In an interview for this study, Taylor Lorenz described being physically assaulted while covering the deadly 'Unite the Right' rally in Charlottesville in 2017 before being targeted by online conspiracy communities. Women journalists in the US have also reported being punched, harassed, touched, called names and other forms of abuse while they covered Trump rallies and other political gatherings, fuelled by online abuse (CPJ, 2019). Several other journalists interviewed had also been attacked or sustained injuries while covering protests in the summer of 2020.

In 2011, while reporting outside a sit-in at a ministerial building during the Arab Spring for Tunisia's Express FM, Henda Chennaoui described how she was beaten after military police clashed with her editors. "Then, I started to see that some people online, some pages, started to put my name and my photo online saying that I had an agenda for this political action, that I was lying about what happened on this day." Najoua Hammami, director of the Media Office at the Arab Institute for Human Rights NGO in Tunisia, said: "In protests, the first targets are always the women journalists, the amount of beating and verbal humiliation is honestly indescribable."

An anonymous political journalist described being subjected to high levels of online harassment in the context of the UK's 2016 Brexit referendum and this moved offline, especially in association with left-wing political rallies, she said. The journalist remembered being "…jeered or surrounded by people" at one rally, "screaming, shouting in my face, you're not welcome here, get out, you're a lying bitch, all of that." Working from press offices inside the UK Parliament, she found herself: "having to walk with my head down or [accompanied] across the road, trying not to be seen because there was a kind of hostile mob that were going to start screaming and shouting, screaming insults to me".

In 2020, rising protests against gender-based violence coincided with the pandemic in Mexico. Female journalists reporting on the movement were harassed, vilified, incarcerated and attacked by the police. For example, journalist Cecilia Solís Martín

was shot by the police in Cancun while covering a feminist protest (BBC, 2020b) and Lizbeth Hernández, a freelance journalist based in Mexico City, was detained by police while covering a sit-in (Hootsen, 2020). These attacks were accompanied by escalating online violence.

And in Pakistan, broadcast journalist Amber Shamsi said she was recognised and followed while reporting on a far-right sectarian group's protest,[120] members of which often target her on social media for her journalism on Pakistan's blasphemy laws.

iv. Offline stalking and harassment linked to online attacks

Cyberstalking that escalates into offline stalking is also a dangerous trajectory. *Sunday Times* journalist Rosamund Urwin, who has received online threats to strangle her newborn baby, has written about her experiences of being physically hunted down in the UK by a man who first started threatening her on Twitter:

> *It is a strange feeling when a fear becomes flesh. For a long time, I have had a nagging worry about a man who harasses me on Twitter. I fretted that he would one day move from being a virtual nuisance to a real world threat* (Urwin, 2018).

The man, in his 30s, began harassing Urwin after she appeared on BBC TV in 2016. He sent sexually threatening messages, "creepy videos", and sometimes sexually explicit images. Then, he did become a physical threat. "And here he was: all 6ft 3in of him, standing in my office's reception, asking to see me. I started to shake with terror, but then later I felt slightly vindicated too. This had not been paranoia — he really did want to hunt me down", she wrote in 2018 (ibid.).

After the police became involved in Urwin's case, the man - described as seriously mentally ill - was moved to a secure facility. However, he broke out twice in 2018, and both times he was found in London, where Urwin works. "Being stalked shrinks your life, and shrinks you. I have hidden away, scared to leave my home, then scared to leave my office. Even now, I dart mouse-like into the Underground station. I have lost my stake in public space. A state of hypervigilance drains you emotionally. My work has suffered" (ibid.). The threats Urwin faces escalated again in May 2021, when her stalker once again escaped confinement:

---

120 The Tehrik-e-Labaiq Pakistan sectarian group called for Asia Bibi, a Pakistani Christian woman jailed for blasphemy and who gained asylum in Canada, to be hanged: https://www.dw.com/en/pakistan-protests-why-the-islamist-tlp-party-is-now-a-major-political-force/a-57214719.

Also in the UK, a message was left for the BBC's specialist disinformation reporter Marianna Spring on a notice board outside the train station she uses to commute to work. "So when I'm walking to and from work, it's me that has to be hyper-aware, me that can't listen to music or podcasts when I'm walking any more. It's me that kind of has to forfeit certain freedoms as a consequence," she said. Meanwhile, *Independent*'s Home Affairs correspondent Lizzie Dearden and her partner were stalked online and offline by a far-right political figure who was attempting to thwart Dearden's critical reporting. He was found to have threatened and harassed them both online and offline.[122]

Serbian journalist Verica Marinčić was physically stalked following online threats and cyber attacks in response to a local news story about men who were speeding in their cars (Djurić, 2019; Mapping Media Freedom, 2018; Apro, 2020). The man who was stalking her allegedly swerved his motorbike into her after she filed criminal charges against him (Stojanovski, 2019). Fellow Serbian journalist Jovana Gligorijević also experienced both on- and offline stalking by a man who impersonated her Twitter account and showed up near her offices (Gligorijević, 2018). According to a survey by the Serbian Independent Journalists Association (NUNS), almost 30% of the 82 women journalists who responded said that the threats they had received turned into months of stalking (Djurić and Jovic, 2020).

Mexican journalist Denise Dresser was shouted at by a man at a supermarket in an incident she said was seeded online. Some of the abuse she experiences is centred on her "mental health, sexuality, image, gender, partners, and the bribe that I have allegedly received," she wrote (Dresser, 2021). In South Africa, there was also an incident in which Ranjeni Munusamy, a senior editor who has since left the profession, was accosted by three men and had her name shouted at while doing her grocery shopping. The South National African Editors Forum (SANEF) stated: "Criticism is part of democracy, but hate speech and sexually abusive threats online and bullying are out of order and can endanger the lives of journalists" (SANEF, 2018).

v. Increasing offline security in response to online attacks

Many research participants also increased their physical security, highlighting their sense of vulnerability, and their awareness of the potential physical consequences of digital attacks, including 13% (n=93) of the women journalists who responded to the ICFJ-UNESCO survey. When harassment turned into non-stop insults and death threats, dozens of the interviewees said they had decided to change their daily routines to feel safer and protect their families. In order to keep doing their jobs, they had to become extremely careful about their offline surroundings as well as their online presence. In some cases, this entailed moving house, cities, or even coutnries.

The case of Sri Lankan author and former journalist Sharmila Seyyid illustrates further the ways in which online violence is accompanied by offline threats. Seyyid was targeted online when she suggested in an interview with the BBC that the legalisation of sex work could provide workers some protections. She was then subjected to a campaign of harassment by self-identified Islamic fundamentalists (NewsMinute, 2015). Seyyid has been "raped and killed online" (Sundaram, 2015), cyber stalked, and had her identity stolen. In connection with the online attacks, she was mocked offline, including while shopping;[123] the English academy that she and her sister ran was vandalised; and her parents have had to relocate several times.

ONLINE VIOLENCE CAUSED 13%
OF WOMEN JOUNALISTS SURVEYED
TO INCREASE THEIR PHYSICAL SECURITY

The online campaign against Seyyid climaxed in a fundamentalist Islamic group circulating Photoshopped images of her apparently mutilated body. At one point, even Seyyid's family was worried that she had been murdered. Seyyid was forced to give up her journalism career and fled the country to India for three years with her infant son (ibid.).

In some cases, women journalists have independently hired security and moved house multiple times to reduce the risk of physical attacks. The Grio's White House correspondent April Ryan told researchers:

> *I had to protect myself by any means necessary, be it putting lighting at my house, making it look like it's daylight at night, having cameras everywhere, having bodyguards, having a certain kind of drill in the house, if something happens, you know, I even park differently, in case somebody tries to attack... . I'm always looking over my shoulder.*

Ryan said her former employer "begrudgingly" paid for security after a bomb squad and the FBI both came to her house to investigate a bomb threat. Brandy Zadrozny[124] at NBC News-MSNBC was also placed under guard by her US employer. And during the UK Brexit debate, the BBC hired bodyguards to protect then political editor Laura Kuenssberg after she received several credible threats.

123 See also references to women journalists being abused while shopping in Mexico and South Africa earlier in this chapter.
124 See Section 2.4 about how white nationalist Darren Beattie led an 'open invitation' to dox Brandy Zadrozny in an interview on Fox News' Tucker Carlson show in October 2020; threats made in "hundreds" of voicemail messages and emails to Zadrozny.

Tunisia's Najoua Hammami was threatened and abused in online comments under her investigative reporting over a six-year period, she said. The threats spread offline, and her newsroom had to introduce security so physical trolls could not gain entry.

Maria Ressa was also physically harassed by online trolls who brought hard copies of their threatening tweets to Rappler's newsroom in Manila before security guards were installed.

In the UK, former chief reporter for *The Mail* Amy Fenton has been subjected to extreme online violence with offline impacts in connection with her reporting on grooming gangs (Pidd, 2020). Fenton's senior editor Vanessa Sims said that 12 men gathered outside the newspaper's offices "shouting intimidating slurs and demands" in connection with the online threats against her journalists (Tobitt, 2020). Fenton has reported more than 100 such threats to law enforcement (Pidd, 2020). She has been given police protection and has moved counties.

In South Africa, Pauli Van Wyk was attacked by populist politicians "on TV, on radio, at their own press conferences, which were often housed in their own buildings and where their supporters [would swarm]". Her editor at Daily Maverick took her off covering press conferences and provided bodyguards; Van Wyk increased security at her home and was forced to move constantly to thwart physical attacks. "At one stage, for a few months... [I moved] to a different location every two weeks," she said.

vi. Online digital security threats with offline impacts

The function of digital privacy and security threats - including surveillance, hacking, and doxxing - also increases the physical risks faced by women journalists, as noted earlier. Doxxing, in particular, has become a standard feature of coordinated attacks against women journalists, with many of those interviewed for this study, and 57 of those surveyed, having had their private phone numbers, home addresses or other personal information published online in the course of major episodes of online violence.

In March 2021, a *New York Times* journalist was doxxed by the right-wing news channel One America News Network (OANN). Her phone number was televised in a segment which encouraged viewers to harass the journalist. When OANN later tweeted the segment, they further exposed her number. The dox-tweet stayed live for several hours before being removed. The *New York Times* told a Gizmodo reporter covering the story: "Journalists should be allowed to do their jobs without harassment. Our reporter will not be intimidated and will continue to follow the facts where they lead" (Cameron, 2021). Two White House correspondents interviewed also received credible death threats associated with doxxing episodes that led to their employers providing increased physical security.

A number of Mexican journalists said the online violence they were subjected to increased the offline risks they faced. However, given the extreme levels of physical violence against journalists in the country - which demands constant

vigilance - digital safety and security have been regarded as less urgent and are somewhat overlooked according to research for this study. Mexican journalists are said to "focus on physical integrity, an interest that explains why in this context digital safety seems to have been neglected or placed in a second order of importance by them" (González and Rodelo, 2020).

Tunisian editor Khaoula Boukrim described a chilling act of intimidation which she connected to the digital security attacks she experienced. After attempts to hack her accounts and her laptop camera, in an episode which she said felt like a "warning to stop writing", Boukrim was woken at 3am - during the COVID-19 curfew - by vigorous knocking on her door. She called the police and the person disappeared. This experience echoes that of the Brazilian journalist Talita Fernandes described above.

vii. Online violence as an enabling environment for offline 'lawfare'

In mid-2021, Indian journalist Rana Ayyub - who has suffered serious online violence episodes, as discussed earlier in this chapter - was facing criminal charges brought by the government over a tweet she shared. The tweet was from a Muslim man who described being assaulted in a racially motivated attack.

In a *Washington Post* column, Ayyub said there was an irony in being charged with "criminal conspiracy, promoting 'enmity', insulting religious beliefs and provoking riots in an attempt to destabilise the country", at a time when India had just signed a Group of Seven summit statement to promote "freedom of expression, both online and offline, as a freedom that safeguards democracy and helps people live free from fear and oppression."

The intent of the legal action, Ayyub said, was to chill her journalism:

> *At a time when I should be reporting on the undercounting of [pandemic-related] deaths, the unemployment in the country, our preparedness for the COVID-19 third wave while continuing to extend help to families in India through my relief work, I am engaged with a battery of lawyers to fight the legal cases and summons slapped on me* (Ayyub, 2021).

Ayyub's case is similar to that of Maria Ressa, as detailed in the big data case study published in Chapter 1. Ressa faces multiple criminal charges brought by the State in connection with her reporting, including in connection with her tweets. She considers the online violence she is exposed to on social media, which has included targeted attacks calling for her prosecution and imprisonment, to be part of the enabling environment for her conviction on a criminal cyberlibel charge and her ongoing legal struggles. British journalist Carole Cadwalladr, the subject of our second big data case study (Chapter 3.2), is experiencing similar offline legal battles (Fitzgerald, 2022).

3.1

Maria Ressa: At the core of an online violence storm

Julie Posetti, Diana Maynard and Kalina Bontcheva

"**First, I'm attacked for being a journalist, second I'm attacked for being a woman.**"

Maria Ressa, Nobel Peace Prize laureate (2021) and co-founder of Rappler

Death threats. Rape threats. Doxxing. DDOS attacks. Racist, sexist, and misogynistic abuse sent via both open and closed social media channels through text, images and memes. These are just some of the features of the online violence that Maria Ressa, the Filipino-American journalist who co-founded the Manila-based news site Rappler, has faced daily since the 2016 election which swept President Rodrigo Duterte to power in The Philippines.

This chapter is a summary of a collaborative big data case study (Posetti, Maynard & Bontcheva 2021)[125] which involved a forensic analysis over of 450 thousand Facebook and Twitter posts directed at the Nobel Peace Prize laureate[126] and 2021 UNESCO World Press Freedom Prize winner[127] between 2016 and 2021. It highlights the intensity and ferocity of attacks designed not only to vilify a journalism icon, but to discredit journalism itself, and destroy public trust in facts. These attacks also created an enabling environment for Ressa's persecution and prosecution in the Philippines. As a result, her life is at risk and she faces decades in jail.

Ressa is an emblematic case study in the global scourge of online violence against women journalists. She lives at the core of a very 21st century storm. It is a fusillade of disinformation and attacks - one in which credible journalists are subjected to online violence with impunity; where facts wither and democracies are put under severe strain.

The former CNN war correspondent says none of her experiences in the field prepared her for the massive and destructive campaign of gendered online abuse, threats, and harassment directed at her over the past six years (Posetti, 2021; UNESCO, 2021g). At one point, in response to an investigative series on State-linked disinformation (Ressa, 2016a), she recorded receiving more than 90 hate messages an hour on Facebook (Posetti, 2017a).

Ressa is not only attacked for being a journalist. She is attacked for being a woman. For the colour and texture of her skin. For her dual citizenship. And for her sexuality. The attacks against Ressa are enabled by a climate in which

125 See: https://www.icfj.org/our-work/maria-ressa-big-data-analysis
126 See: https://foreignpolicy.com/2021/10/11/maria-ressa-nobel-peace-prize-dmitry-muratov-journalism-media/
127 See: https://www.unesco.org/en/articles/world-press-freedom-day-philippines-journalist-maria-ressa-receive-2021-unescoguillermo

former President Duterte publicly condemned her - while also once musing that journalists are not exempt from assassination (Lewis, 2016; Hammer, 2021). The worst attacks against her appear to have been orchestrated (Nimmo, Shawn Eib and Ronzaud, 2020).

Despite belated and largely ineffective attempts to address hate speech within the social media ecosystem, the technology companies concerned serve as the vectors for these attacks, which in turn create a facilitative environment for State-led legal actions against Ressa. This 'lawfare,' as she refers to it, led to her conviction in mid-2020 on a criminal 'cyber libel' charge, and it continues to escalate.[128] Issued with 10 arrest warrants in less than two years, and detained twice in the space of six weeks, Ressa was still fighting seven separate cases in 2022 (Amnesty International, 2019; BBC, 2020a; Maas, 2020).

If she is convicted on all charges, she could spend the rest of her life in jail. Caoilfhionn Gallagher KC, the co-lead of Ressa's international legal team, states that for having published journalism, her client now "faces a barrage of baseless lawsuits that seek to criminalise her work and expose her to a century in prison" (Doughty Street Chambers, 2020). She told the researchers:

> *The authorities vilify her, and President Duterte has helped to amplify online attacks against her. State authorities thus both directly attack Maria, and also create an enabling environment that facilitates and fuels abuse from others. In turn, online abuse emboldens the authorities in their persecution of her. In my view, there is a symbiotic relationship between the abuse Maria experiences online and the progress of the legal harassment offline.*

This case study combines big data analysis - using Natural Language Processing (NLP) techniques (merging linguistics, computer science and Artificial Intelligence) and network analysis - with deepdive interviews and extensive desk research. The analysis of the online violence Ressa experiences aids understanding about how and why she is targeted, and how the threats spread. It also provides insights into the role of high-ranking political actors in manufacturing consensus by fuelling the behaviour of the social media users who target her. "If I wanted to see what the government was going to do, I only needed to look at social media because the attacks to arrest me and shut down Rappler were seeded as meta-narratives in 2017," Ressa said. "And now here we are."

The function of the social media companies in facilitating the abuse is also spotlighted by Ressa: "The only way it will stop is when the platforms are held to account, because they allow it...They have enabled these attacks; they should not be allowing this to happen."

---

128 as this study went to press, ressa was preparing to appeal to the supreme court of the philippines, after losing an initial appeal against her conviction on a trumped-up criminal cyber libel charge.

We have identified 12 key findings from an examination of online violence against Ressa over a five-year period (2016 to 2021).

12 Key Findings

Attacks designed to discredit Ressa's journalism:

- **Almost 60%** of the attacks on Ressa extracted from the combined datasets were designed to undermine her professional credibility and public trust in her journalism.

- Credibility or reputation-based attacks frequently deployed disinformation tactics and abuse maligning Ressa and her journalism as "fake news".

Attacks designed to undermine Ressa's personal dignity:

- **Over 40%** of the attacks in the combined datasets targeted Ressa at the personal level - often viscerally.

- The most prevalent type (**34%**) of personal abuse against Ressa is sexist, misogynistic and explicit. Racist and homophobic abuse accounted for **5%** of personal abuse.

- **14%** of all of the abuse studied is sexist, misogynistic and explicit.

Networked and technically advanced abuse:

- There is direct evidence that the online violence targeting Ressa has offline consequences. It has created an enabling environment for her persecution, prosecution and conviction. It also subjects her to very real physical danger (Occeñola, 2019; Posetti 2020a).

- The use of abusive memes and manipulated images, which 'fly under the radar' of detection, is commonplace.

- There is evidence that the worst of the attacks on Ressa involve coordination or orchestration - often associated with disinformation campaigns led by State actors.

Sources, triggers and impacts of online violence against Ressa:

- Much of the abuse is fuelled by statements and messages from political leaders and partisan bloggers/social media influencers who demonise Ressa and Rappler as unpatriotic criminals. It is also amplified by pro-government news media.

- Lightning rods for attacks include Rappler's investigative journalism; Ressa's reporting and commentary on State-linked disinformation; Ressa's high-profile media appearances; her industry accolades and other awards; and her court appearances (Evangelista, 2016; Ressa, 2016a; Vick, 2018; Reuters, 2020; Sackur, 2020).

The role of the platforms:

- Facebook is the main vector for the online violence Ressa faces. It is also the most used social media site in the Philippines.

- Both Facebook and Twitter have promised to address the attacks on Ressa but during the period under review, Facebook failed "woefully", she said, to effectively stem the tide of hate against her. However, she feels "significantly safer" on Twitter.

i. A climate of impunity

The Philippines is routinely ranked as the most dangerous country in Southeast Asia for journalists, and UNESCO's Observatory of Killed Journalists identifies 110 fatalities since 2006, with 83% of cases currently unresolved (UNESCO, 2021c). Targeted online violence attacks against journalists like Maria Ressa thus need to be examined in the light of the impunity situation. In 2009, the country was the site of the deadliest attack on journalists ever recorded by the Committee to Protect Journalists (CPJ) - the Maguindanao massacre, which killed at least 30 journalists and media workers[143] in an "orgy of political violence" (CPJ, 2009; Crispin, 2020). Immediately before President Rodrigo Duterte took office in 2016, he issued threats against journalists (Lewis, 2016), and a year later he criticised Ressa in a State of the Nation Address (Ranada, 2017). Since 2017, political influencers have been seeding meta-narratives on social media, painting Ressa as a criminal and calling for her arrest. They were also calling for her to be sexually assaulted, killed and even "raped repeatedly to death" (Posetti, 2017b).

In a country like the Philippines, where extrajudicial killings continue to be condemned by UN actors, the potential is high for online violence against women journalists to reap deadly results (OHCHR, 2020d). As Rappler's Executive Editor Glenda Gloria noted:

83%
OF KILLED
JOURNALISTS' CASES
ARE UNRESOLVED
IN THE PHILIPPINES

We never doubted that those online threats would translate to physical threats. That's why we doubled not just the security of Maria, but of the newsroom, because a lot of the online threats against activists turned into reality. There was this female activist who was first blasted online and shot while on her way home. It's real. Especially against women.

ii. "Shark tanks" full of data

In this case study, nearly 400,000 tweets directed at Maria Ressa during a 13-month period from December 2019 to February 2021 are examined, along with more than 56,000 public posts and comments published on Facebook between 2016 and 2021. The tweets were gathered and analysed using Natural Language Processing (NLP) techniques. From the large dataset, a sample of 1,128 tweets demonstrating highly explicit abuse - predominantly expressed in English - was extracted for detailed analysis.

In parallel, data analysts from the digital research firm Graphika conducted a network analysis on a subset of tweets from a spike of abuse in June 2020 that coincided with Ressa's conviction on the first 'cyber libel' charge (Nimmo, Shawn Eib and Ronzaud, 2020). The aim of this analytical process was to map the types and methods of attack, along with the trajectory of the abuse and the interconnectedness of those attacking Ressa.

The Facebook data represents over 9,400 public comments in response to Ressa's posts that were gathered from her professional Facebook page with her explicit permission, and over 47,000 other public Facebook posts mentioning Ressa extracted from a massive database that Rappler maintains, called 'Sharktank.' According to Ressa, 'Sharktank' maps the information ecosystem of the Philippines on Facebook (which is virtually synonymous with the internet in the country, reaching 96% penetration in 2021) (Statista, 2021). As of January 2021, when the data for this case study was extracted, the database had captured 471,364,939 public posts and 444,788,994 public comments made by 4,176,326 users, 68,000 public pages and 26,000 public groups on Facebook. The same NLP techniques were applied to both sets of Facebook data with the assistance of Rappler's research team.

The NLP analysis is based on a 'high accuracy detection' model and it is largely restricted to English language posts, or those which blend English with Tagalog (the other national language of the Philippines). Consequently, the samples of online abuse extracted using this method are considered to be severely underreported, capturing only around 50% of all English-language abusive messages present in the target's social media stream, according to previous studies.

It is also important to note that this analysis excludes the most brutal online violence Ressa has experienced, which she says came via Facebook Messenger. Such content is not only harder to detect and save for big data analysis, it is also more difficult to report because automated abuse reporting systems have very

400K
TWEETS
56K
FACEBOOK POSTS
WERE STUDIED IN MARIA RESSA'S CASE

IS USED BY
96%
OF FILIPINOS ON THE INTERNET
(Statista, 2021)

limited multimedia capability and they do not easily accommodate nuance. Similarly, public abuse is often subtle and potentially designed to 'slip under the radar.'

A classic example of online violence targeting Ressa that evades detection by social media companies is a death threat sent to Ressa on Twitter in February 2021. This is subtle partly because it carries text in an image which is not easily processed by automated tools, and partly because the message itself contains no abusive terminology, though the underlying meaning is clear. This message was finally deleted in mid March 2021 after being reported to the platform by the lead author and others.

Figure 7: A death threat sent to Maria Ressa via Twitter on February 12, 2021. The threat was finally removed by Twitter in mid-March 2021.

iii. Methods, themes and tropes deployed by the attackers

While the vast bulk of online violence that Ressa experiences occurs via Facebook, the most common themes and methods of attack were relatively consistent across both platforms. This data was contextualised with online research and in-depth interviews with Ressa and her colleagues. An examination of the combined datasets led to the following findings:

Dominant themes and tropes deployed by Ressa's online attackers:

- Abuse damaging to Ressa's professional reputation or credibility includes disinformation designed to discredit her as a journalist and erode trust in her journalism. It features calls for her to be charged, tried, raped, imprisoned etc. for her work. This abuse frequently involves false claims that she is a purveyor of "fake news" and includes the pernicious hashtag #presstitute.

- Misogynistic, sexist and explicit abuse includes abuse targeting Ressa's physical appearance (emphasising her skin condition) and manipulated photographs depicting her head associated with male genitalia.

- Racist abuse and memes constitute 3% of the personal abuse Ressa receives.

- Homophobic slurs designed to question her sexuality and increase her vulnerability were determined to represent 2% of personal attacks.

- Threats of physical violence, including death threats embedded in images, and threats of sexual violence (e.g., being "publicly raped to death") were associated with the worst attacks (Go, 2018).

Typical methods of attack:

- Key significant attacks appear to be orchestrated (with the detectable use of fake and bot accounts), and on occasion this has led Facebook to remove networks of accounts identified as participating in what they call 'coordinated inauthentic behavior' (Gleicher, 2020; Nimmo, Shawn Eib and Ronzaud, 2020). However, the company's response to the attacks on Ressa has been inconsistent and in her words "woefully inadequate";
- Hashtags designed to encourage swarms of attackers and fuel 'patriotic trolling' are frequently used, and sometimes include abuse within them e.g., #ArrestMariaRessa (Posetti 2017b; Nyst and Monaco, 2018).
- Memes and manipulated images are deployed to increase engagement with the attacks on Ressa and avoid automated abuse detection tools
- Doxxing is used to motivate Ressa's online attackers to attack her offline as well (Rappler, 2019b).

iv. Scrotums, monkeys, swarms and lies

As indicated above, the dominant theme of the online violence waged against Maria Ressa involves damaging her professional credibility, by extension her reportage, and by association Rappler's. This is clearly evident within a smaller subset of highly explicit abusive comments (mostly in English or English-Tagalog) gathered from her professional Facebook page. Attacks against Ressa and Rappler dominated the comments on that page, which she established in 2015. According to Ressa, Facebook had recommended that she start the page to help better manage her comments. But the harassment she experienced on the page soon became overwhelming. She has not posted to the page since early 2019, and it now lies effectively dormant.

Of the 9,433 comments from Ressa's professional Facebook page - spanning the period 2015 to 2018 - 54% fall under 'attack clusters', while supportive comments represented only 4% of the data. This means that for every one comment supportive of Ressa, there were about 14 comments attacking her. And a more granular analysis revealed that approximately four of these 14 abusive comments would constitute personal attacks, focusing on her appearance, nationality, gender and sexuality. Nearly half of the abuse falls into the category of attacks on her professional reputation. Among these comments were disinformation-laced attacks, including accusations that she was a 'fake news' peddler, like these:

> **You are the Queen of Fake News fucking Bitch (sic)**

> **Stop spreading Lies you Piece of SHITS (sic) I wish you Rotten (sic) in Jail ..**

> **Maria Ressa get the fuck out of our country Philippines! dont (sic) mislead the people with your fake news...**
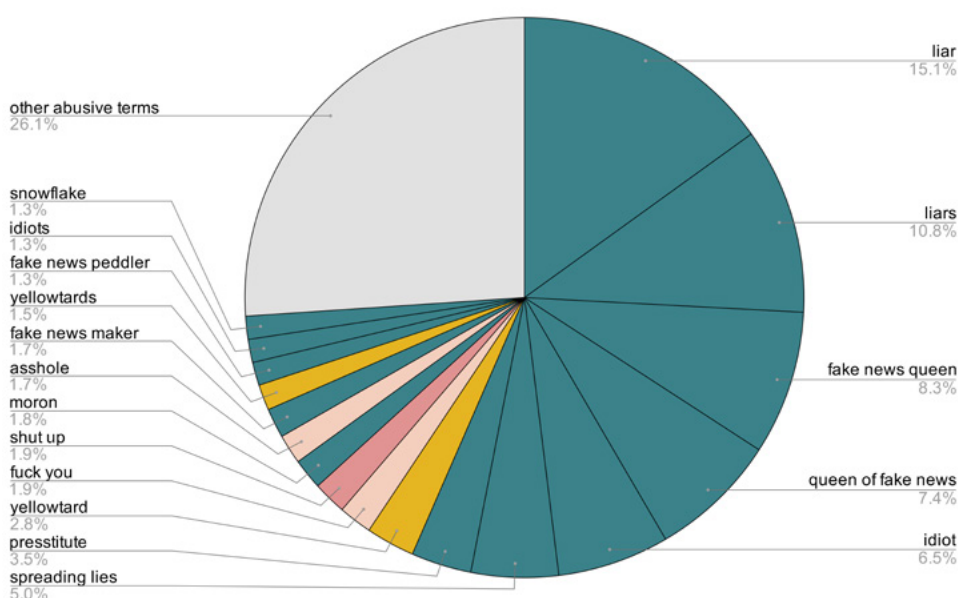
Figure 8: Frequency breakdown of abusive terms that appear more than once in Facebook posts mentioning Maria Ressa ("other abusive terms" includes those which appear between 1 and 15 times in the data). Personal attacks consist of sexist, misogynistic and explicit sexual terms (light pink); and other kinds of personal insult (dark pink). Political attacks (yellow) use terminology associated with (real or imagined) political affiliation.

The most frequently used abusive terms in this data subset were words designed to ridicule, silence and discredit Ressa while simultaneously undermining public trust in her critical journalism. The top ranked words were "idiot," "shut up," "presstitute," and variations on "liar." Around 20% of the attacks on her credibility were related to disinformation - either equating her with it, or falsely accusing her of peddling it.

Disinformation narratives deployed against Ressa were also prevalent in the larger multilingual dataset extracted from her professional Facebook page. These include repeats of claims that she is a "liar"; the "Queen of Fake News"; "Bayaran" (a

Tagalog term for a corrupt journalist who takes payment for favourable coverage); a "presstitute" and a "national security threat" or terrorism supporter, echoing narratives from a prominent Philippine blogger, known as Thinking Pinoy, who tried to get the hashtag #ArrestMariaRessa to trend in May 2017 while in Russia during a mission by former President Duterte (Thinking Pinoy, 2017; Rappler, 2018b). Two years later Ressa was in fact arrested, and within three years she would be convicted on a criminal cyberlibel charge prosecuted by the State.

COMMENTS ON MARIA RESSA'S PROFESSIONAL FACEBOOK PAGE

54% ATTACK CLUSTERS

4% SUPPORTIVE COMMENTS

Figure 9: A screengrab of a Facebook post published by the pro-government blogger 'Thinking Pinoy'.[129]

The same patterns of assaulting Ressa's credibility using disinformation tactics - including falsely accusing her of being a liar and a disinformation purveyor - were evident in the detailed analysis of the Twitter data. On Twitter, the overwhelmingly dominant use of reputational slurs such as "liar" and variants such as "spreading lies" and "lying bitch," as well as terms relating to the spreading of disinformation such as "fake news queen" and "fake news peddler". As Ressa has assessed: "Lies spread faster than facts. And lies laced with anger and hate spread faster and further than facts" (Vosoughi, Roy and Aral, 2018).

Figure 10: This tweet demonstrates the ways in which Ressa is demonized as a 'fake news' peddler in disinformation campaigns that also leveraged a minor error she made in an international TV interview.

129 In line with academic research ethics protocols, we have obscured the identities of most social media users. In this instance, however, we have not redacted the user's identity as a matter of public interest. 'Thinking Pinoy' is identified in the data and in media coverage linked throughout this study as one of the key pro-government influencers targeting Maria Ressa and Rappler.

The word cloud below (figure 5.) shows the 100 most frequently occurring abusive terms directed at Ressa on Twitter. Terms such as "liar" are often entirely capitalized in the messages, indicating strength of emotion - these slurs are being 'shouted.'
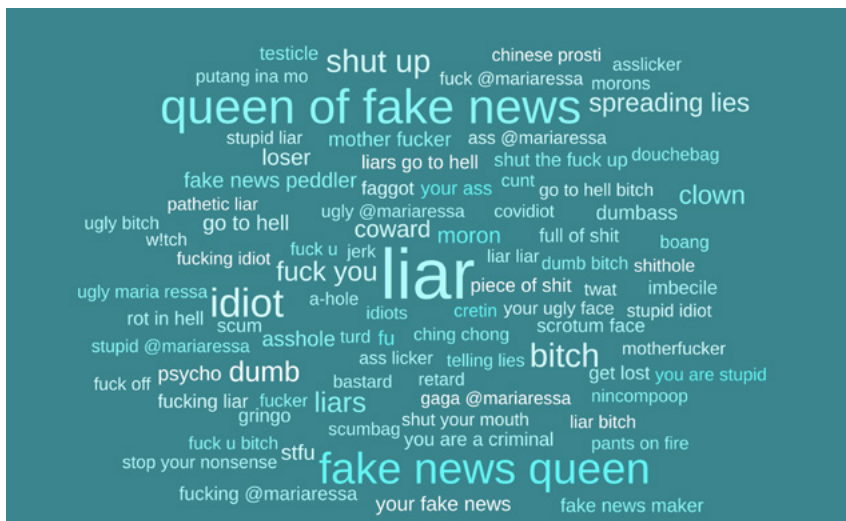
Figure 11: The 100 most prevalent abuse terms in the analysed twitter data. Term size reflects frequency of occurrence.

Looking at breakdown by frequency, the term "liar" alone represents more than 15% of all abusive terms found among the tweets analysed, with "liars" accounting for another 3%, not to mention variants on this theme. The terms "queen of fake news" and "fake news queen" together make up 17% of abusive terms (again, with other similar terms relating to 'fake news' accounting for a similarly high proportion. Also evident are sexist terms (e.g., "bitch," "slut"); sexual terms ("go fuck yourself," "pussy," "scrotum face," "asshole"); homophobic terms (e.g., "lesbo"); and racist terms (e.g., "gringo").

The 'lawfare' effect

Online violence against Ressa undoubtedly impacts on decisions in the progress of the cases and charges she faces according to her international lawyer, Caoilfhionn Gallagher, who said that "when the Duterte administration is making decisions in that environment, the fact that Maria is a hate figure online is enabling those decisions to be taken". Referring to the decision to charge Ressa with a third count of criminal cyberlibel in January 2021, Gallagher said:

> *It's impossible not to be aware of the fact that Maria was the subject of a very large amount of viral abuse throughout that time, when the decision was being made. There was a coincidence in time between when the prosecutor was making that decision about the third charge, and a spike in abuse linked to the arrest warrant for Maria in respect of the second cyberlibel charge.*

In Ressa's view, the State "propaganda machine" has accomplished its goal: "They pounded opponents and journalists to silence to create a bandwagon effect for

seeded meta-narratives of bias, incompetence, criminality, and corruption to be levelled against them."

According to Rappler Executive Editor Glenda Gloria, Ressa transformed into a "warrior" in response: "...trying to think both of strategy and tactics, and the soldier being hit and being machine-gunned." But the global support Ressa mustered through investigative journalism and press freedom advocacy had an international impact and that was a source of tremendous psychological benefit for Ressa, because "It showed her that there's hope," Gloria said.

Foreign actors

Foreign states feature extensively in the tweets identified as containing highly explicit abuse, as well as a conspiracy theory that Ressa was in league with some of them who were allegedly paying her to spread lies. This was also a theme evident in the Facebook data. On that platform, Ressa was targeted in influence operations originating in China designed to foment popular support for political campaigns in the Philippines (Nimmo, Shawn Eib and Ronzaud, 2020). That disinformation network was removed by Facebook based on evidence of 'coordinated inauthentic behavior' (Gleicher, 2020).

Networked gaslighting

Another feature of credibility-based attacks is the orchestration of disinformation-laden 'pile-ons' aimed at Ressa, which are designed to discount her investigations into disinformation campaigns associated with political interests (Morrison, 2020). These practices are features of what can be called 'networked gaslighting'. In Ressa's case, the target of the attack is falsely accused of practising the behaviour of her attackers.

One early example involves Ressa's 2016 investigation *Propaganda War: Weaponizing the Internet* examining 'astroturfing'[130] and 'sock puppet networks' of fake accounts linked to President Duterte's election campaign, practices also later associated with the extrajudicial killings connected to the so-called 'drug war' (Ressa, 2016a). In response to Ressa's and Rappler's reporting on what they describe as "government-sponsored information operations," 'patriotic trolls' swarmed Ressa's social media zone, often prompted by partisan bloggers who encourage their followers to prove that they are not paid trolls.

"Even if some of the people who send Maria horrendous, abusive images have no known direct link to the State," Caoilfhionn Gallagher KC said, "they are likely to be doing this, or feel emboldened in doing this, because the State is sending the message that this person is fair game." These campaigns have spilled offline, increasing the physical threats Ressa is facing.

In one instance, involving the doxxing of Ressa with her email and office address published online, pro-government social media activists came to the Rappler newsroom in person (Posetti, Simon and Shabbir 2019a; Navallo, 2021). They

---

130 The act of manufacturing consent through influence operations designed to create the false impression of a groundswell of support within online communities.

bypassed security and protested outside the glass walls of the newsroom, while holding up signs replicating some of the offensive hashtags and narratives swirling on social media.

Figure 12: The online abuse targeting Ressa and Rappler spills offline in February 2019. Thesemen came to the Rappler newsroom in Manila and held their 'posts' up to the glass windowsof the office when they were barred entry by security. Image: Rappler

Personal attacks designed to shame, humiliate and silence

A total of 40% of all the abusive posts studied could be categorised as 'personal attacks' and these are frequently sexist, misogynistic, racist and vulgar in combination, focusing on Ressa's physical appearance. The biggest cluster of personal attacks were comments related to her face ("*mukha*" in Tagalog). She was often compared to animals like monkeys and dogs (classic racist and sexist tropes) and in several instances, her eczema was compared to a scrotum - a form of abuse which has more recently grown into a viral meme that jumps platforms.

Terms like "scrotum face," "idiot," "imbecile," "bobo" (dumb or stupid), "moron," and "psycho" were prevalent. Additionally, she was condescendingly referred to as "ang babae", or "this woman," demonstrating the sexist undertones of much of the abuse that targets her. Comments about Ressa's sexuality, including homophobic slurs (e.g., 'Tomboy' is slang for 'lesbian' in the Philippines) and outright profanities like "fuck you" and "*putang ina*" (son of a bitch) were also prevalent within these clusters.

The risk of physical violence associated with online threats against Ressa

Outright threats to Ressa were also detected on Facebook, with commenters saying she should be sexually assaulted, die, or be killed. "I've always known online violence leads to real world violence," Ressa said. This is a view shared by Caoilfhionn Gallagher KC, who points to chilling similarities between Maria Ressa's case and that of the murdered Maltese journalist Daphne Caruana

Galizia, who was brutally attacked online - with misogynistic references to 'witch burning' - before she was killed with a car bomb in 2017:

> *"There are shocking similarities between Maria and Daphne's cases, including a long period of time in which they both experienced a combination of attacks, from multiple different sources, online and offline – State-facilitated and State-fuelled".*

Gallagher represents Caruana Galizia's bereaved family, who issued a statement highlighting the parallels between the cases when Ressa was convicted in 2020 (Daphne Caruana Galizia Foundation, 2020). "I'm aware of where this can go," Ressa said.

Hoy **Tomboy** Maria Ressa! Ur such an idiot Prestitute! Bias Media na Bayaran! Like · Reply · Message · 3y 60 ↳ 6 Replies	Hey Maria Ressa you tomboy! You're such an idiot *presstitute*! Bias paid media!
ugliest monkey bitch. mukha mo sunog na sunog. virgin ka pang hayop ka. walang papatol sayo Like · Reply · Message · 3y	Ugliest monkey bitch. Your face looks burnt. You are still a virgin, you beast. Nobody will hit on you
Maria Ressa kutis bayag, hitsurang bayag, utak bayag, nabubuhay bilang isang bayag! Ambobo pa sa math! Tulad ng amo nyang si leni! 😂😂😂 Like · Reply · Message · 41w	Maria Ressa scrotum-skinned, scrotum-looking, scrotum-minded, lives like a scrotum! You don't know math! Like her master, Leni [Robredo]!
Ipatapon sa mindanao ang **babaeng** galisin nato at may sakit sa bato. Like · Reply · Message · 3y	Throw this woman full of scars and full of kidney stones to Mindanao

Figure 13: Samples of Facebook comments with personal attacks against Maria Ressa. In thecomments above, Ressa is called a "lesbian" and "monkey," and her skin was compared to ascrotum. She is also often condescendingly referred to as "this woman".

vi. Spikes and triggers

The attacks on Ressa frequently spike in association with:

- **Rappler's investigative journalism focused on the 'drug war', and the erosion of democracy in the Philippines (Evangelista, 2016; Smith, 2020).**
- **Ressa's reporting and commentary on disinformation and the government (Ressa, 2016a);**
- **Ressa's high-profile media appearances (Sackur, 2020);**
- **Ressa's international awards and civil society statements of support (Vick, 2018; ICFJ, 2020b);**
- **Ressa's court appearances (Reuters, 2020).**

"Every time a complaint reaches the court, every time a statement is made in support of Maria, there's a troll army that really is commanded to respond," Rappler's Executive Editor Glenda Gloria said. This phenomenon was again

underscored in December 2021 while the researchers were monitoring the surge of social media attacks targeting Ressa in the aftermath of the Nobel Peace Prize ceremony in Oslo which focused unprecedented global attention on the laureate (BBC, 2021d). In a related episode, five days after the ceremony, Rappler came under what Ressa described as a "massive DDOS144 attack", with more than 6 billion access requests during the attack.

The largest surge in online violence identified against Ressa on her Facebook page occured in October 2016, when Rappler published a three-part investigative series into State-linked disinformation networks - two of which were written by Ressa herself (Ressa, 2016a).

The Facebook data covers a five-year timespan and clearly demonstrates other attack spikes connected to Rappler's critical coverage of the 'drug war' and the extrajudicial killings associated with it, along with international media attention (e.g., associated with the award-winning film A Thousand Cuts which documents Ressa's struggles), her high-profile awards, and her arrests, detentions, and trials are also associated with increased attack spikes in the Facebook data (ICFJ, 2018; Taddonio, 2021).

An examination of the Twitter data detailing frequency of abuse reveals three major spikes, each of which have more than 50 tweets per week identifiable as highly explicit abuse (largely in English or hybrid English-Tagalog). The largest peak was in early May 2020, and it was triggered by an interview with ABC Australia in which Ressa misspoke, and this was used to attack her credibility (Rappler, 2020).

The next biggest spike came in June 2020, when a Manila court delivered a guilty verdict against Ressa in a criminal cyberlibel charge prosecuted by the government. Her criticism of the government due to the Philippines being recorded as having the highest proportion of COVID-19 cases in Asia in August 2020 represented the third-biggest spike.

vii. Deep dive: Network analysis of Twitter reactions to Ressa's June 2020 conviction

According to Rappler's Executive Editor Glenda Gloria, Maria Ressa's 2020 conviction on a criminal cyberlibel charge "...really provided the trolls [with] a powerful hashtag, because this was like a court already saying what the troll army believed to be true. And so that gave them an editorial agenda - 'it's not just us saying that she's a criminal, it's the court!'"

In parallel with this research, data analytics company Graphika conducted a detailed analysis of 196,000 tweets from 80,886 distinct users featuring citations of @mariaressa and the term "Maria Ressa" posted between June 9th and June 17th 2020. Activity began accelerating in the days leading up to the court's verdict, peaking on June 15th 2020, the day of the decision. While 'mentions' expressing shock at the verdict and support for Maria Ressa flooded the zone on the day of her conviction - including solidarity messages from international journalists and

civil society organisations - pro-government accounts provided evidence of an orchestrated response. These accounts largely celebrated her conviction while attacking Ressa based on her dual Filipino-American citizenship.

Here, there are strong pointers to organised trolling, with over 40 accounts from the pro-government segment constantly mentioning @mariaressa or citing the term "Ressa" over 30 times each within the period. Besides directly targeting her (@mariaressa), these accounts were also predominantly retweeting anti-Ressa messaging pushed by a select few accounts.

Top false narratives deployed by the pro-government 'troll army' in the immediate aftermath of Ressa's conviction were:

1. She is now proven to be the criminal we said she was.

2. Disinformation about the role of the State in her prosecution (i.e. they falsely argue that the case was prosecuted by a private citizen, but it was a criminal prosecution waged by the State).

3. She is a foreigner (she is a dual American-Filipino national), subject to foreign masters (this theme helps prosecute the false argument that Rappler is foreign-owned, which is attached to a string of cases designed to shut down the news publisher).
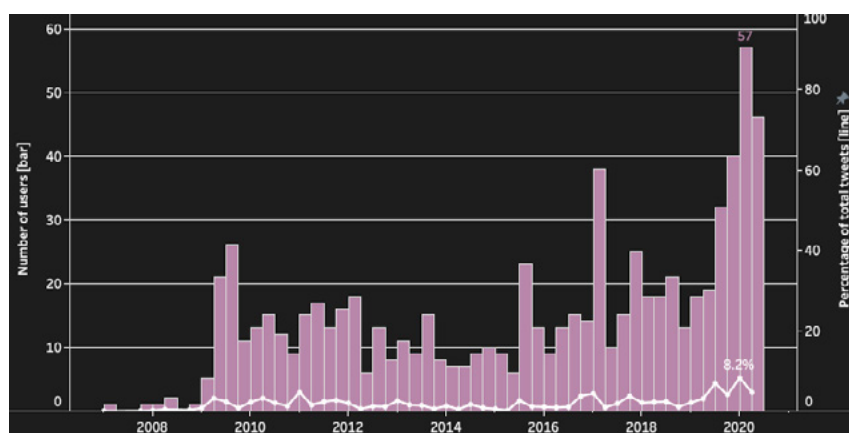
Figure 14: Distribution of pro-government account creation dates for Twitter accounts targeting Ressa (source: Graphika).

Pro-government accounts showed the strongest tendency to mention other users in regard to this narrative. Thus, almost 60% of these accounts referenced at least four other users within the dataset in this time period. This level of interactivity is uncommon and indicative of a possibly aligned/coordinated harassment campaign in which users aim to amplify attacks.

A large number of the pro-government accounts were also fairly recently created. Accounts created after 2019 also tended to be more active. And accounts created during or after April 2020 were the most active within this set - producing about 27% of the total activity (662 tweets).

Finally, a significant number of these accounts have low follower counts with 5% having zero followers and a little over 25% having 10 or fewer followers. This

combination of high activity, low follower count, and recent creation date are together possible indicators of accounts created in bulk to amplify pro-government messaging and target government critics.

viii. The enabling role of the platforms in Maria Ressa's abuse

As this big data analysis demonstrates, while a coordinated and vitriolic mob of digital aggressors instigate and fuel the attacks on Maria Ressa, Facebook - which is equated with the internet in the Philippines - is the major vector for the disinformation-laced online violence she experiences. Rappler was 'born' on Facebook and former President Duterte made use of the platform to win victory in 2016 (Posetti, Simon and Shabbir, 2019a), as did his successor, President Bongbong Marcos in the 2022 election.[131]

Ressa and the Rappler team have consistently flagged the online attacks with Facebook, which has largely done little to deal with the problem, they say. Ressa's Facebook page was overwhelmed by attacks for years, and posts that incite violence, despite violating Facebook's community standards, remained visible on the page at the time of writing. Rappler responded to the attacks with stricter comment moderation policies, but the abuse on Ressa's page only dropped when she wound down posting in the second half of 2018.

In September 2018, Ressa told Facebook executives: "If you don't change what you're doing, I could go to jail" (Posetti, 2020a). After she was convicted in June 2020, she apportioned some blame to the company for her conviction, and she continues to argue that Facebook has enabled the destruction of democracy in the country (ibid). Ressa is very critical of Facebook's moderation policies and practices, saying that their automated reporting systems just do not work when it comes to dealing with online violence against women journalists.

"I have very rarely had anything taken down when trying to report attacks using Facebook's standard online reporting system," she said. "The only times Facebook has done something about the attacks against me is when I have gone directly to people I know inside the company, many of whom have nothing to do with content moderation! Their systems need to be clearer, faster and more responsive to the impacts of their inaction."

"The design of social media turned 'wisdom of the crowds' into the mob. It's the chaos of a mob. And beyond that, it's actually pumping hate into the system," Ressa said. She believes the platforms need radical renovation - of business models and technical design - to stop the toxicity that overruns them. "I don't think anything is possible until we clean up the information ecosystem, until you stop the virus of lies," she said. "It's a perfect comparison to the COVID-19 virus, because it is very contagious. And once you're infected, you become impervious to facts."

Twitter is also a significant distributor of abuse against Ressa, although she says she feels "much safer" on that platform - especially since the company appears, in her view, to have begun working harder over the past few years to protect women

journalists and human rights defenders among its users. "Their reporting tool aggregates similar tweets, takes less time, and is far more effective in takedowns," Ressa said.

She continues to work at the intersection of investigative journalism, research and policy to respond proactively to the information ecosystem crisis (RSF, 2020h). And Rappler is still as a Facebook fact-checking partner in the Philippines, and a collaborator on investigations into disinformation networks which, in some cases, have resulted in the company removing clusters of inauthentic accounts from the platform.

Ressa has long spoken about being the 'canary in the coalmine,' warning that the Philippines' information ecosystem represents the West's dystopian future (GIJN, 2019; Rappler, 2019c). She is extremely pessimistic about the prospect of the platforms responding quickly and effectively enough without changes to accountability and liability. "The only way it will stop is when the platforms are held to account…because they allow it. It's kind of like if you slip on the icy sidewalk of a house in the US, you can sue the owner of the house. Well, this is the same thing. They have enabled these attacks. They've certainly changed my life in many ways."
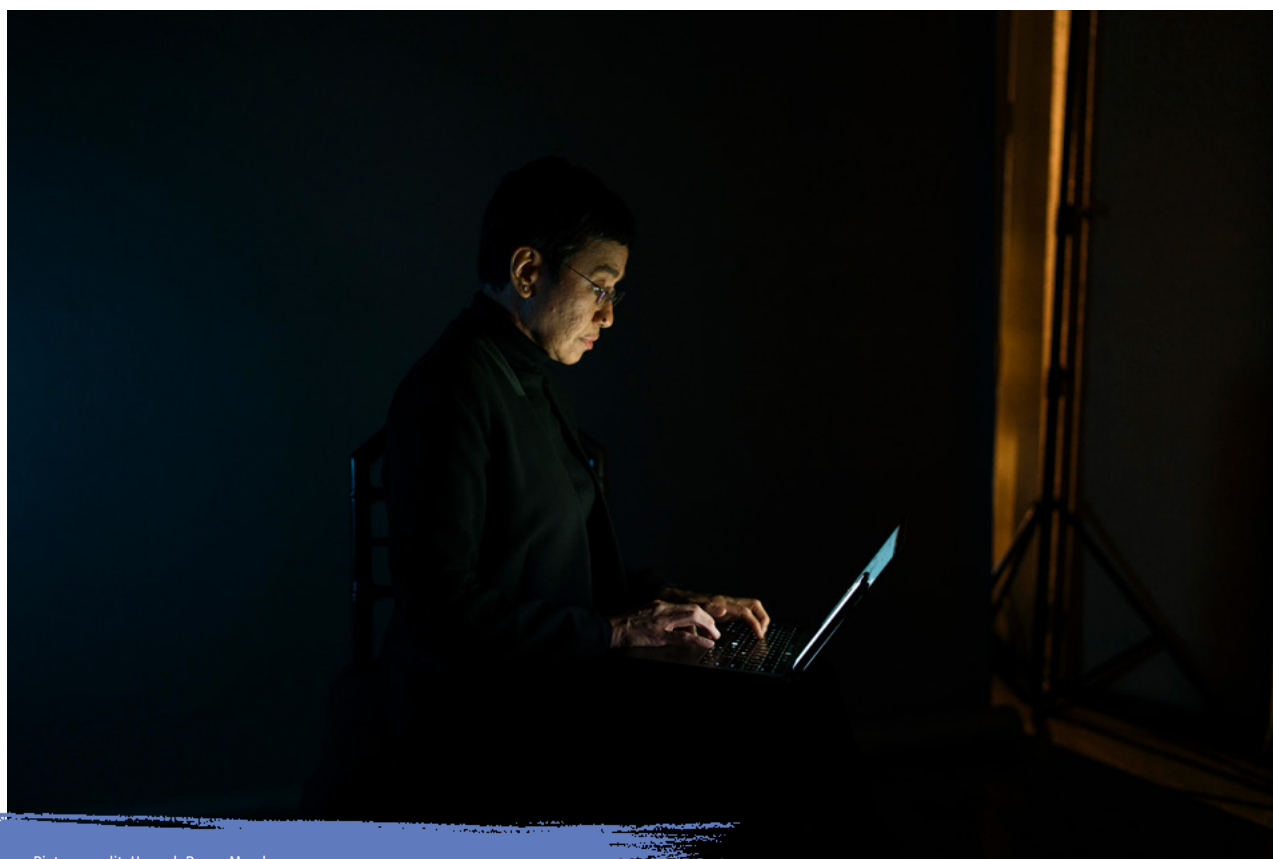
Picture credit: Hannah Reyes Morales

3.2

Big Data Case Studies

Carole Cadwalladr: The networked gaslighting of a high-impact investigative reporter

Julie Posetti, Diana Maynard, Kalina Bontcheva, and Nabeelah Shabbir
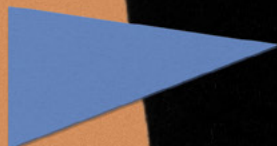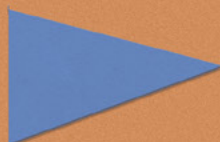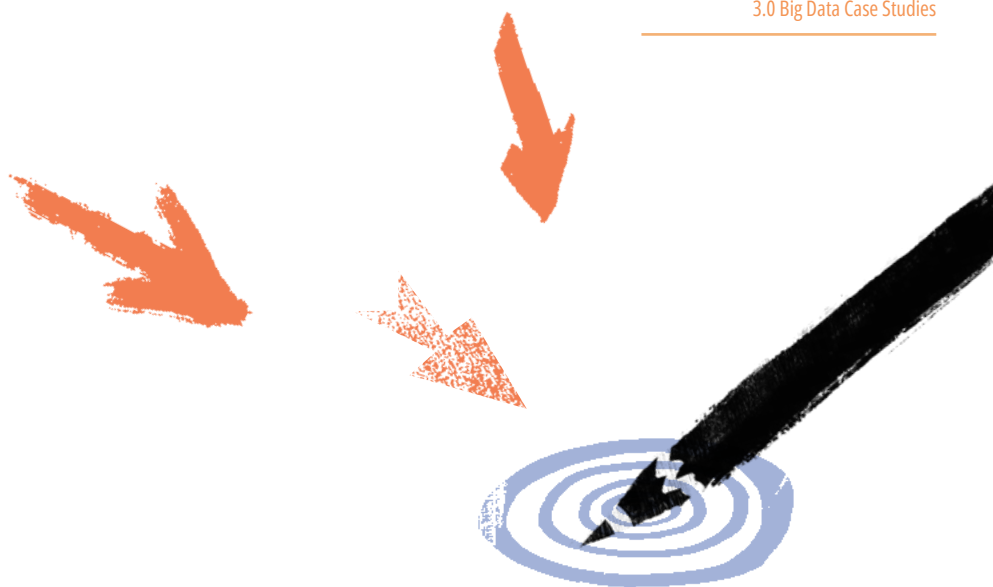
"I have become, through evidence-based reporting, a cultural war hate figure and a 'valid' target."

Carole Cadwalladr

Carole Cadwalladr is a multi-award winning British journalist whose investigative work exposed the Facebook-Cambridge Analytica scandal that ultimately led to the biggest fines in history being imposed on the social media giant by privacy regulators on both sides of the Atlantic. It also resulted in the collapse of the "sketchy" data analytics company closely associated with former US President Donald Trump's successful 2016 election campaign (Lomas, 2019).

But from the moment *The Guardian* and *Observer*[132] newspapers published the first story in Cadwalladr's investigative series on the scandal that compromised the Facebook accounts of up to 87 million users, the journalist became the target of a malign, misogynistic, disinformation-laced campaign of online violence which grew increasingly threatening over time. This campaign has also created the enabling environment for her legal harassment by political actors.

Cadwalladr's reportage linked the Cambridge Analytica scandal to both the election of former US President Donald Trump, and Brexit - the referendum which led to the withdrawal of the UK from the European Union. It suggested the widespread manipulation of Facebook users by political actors using microtargeting techniques, which fed into highly divisive politics in the US and the UK between 2016 and 2020. And it continues to resonate. Her journalism has led to criminal investigations, and parliamentary inquiries in multiple countries.

Despite reporting extensively on the global impacts of Facebook privacy breaches and disinformation on the platform, Cadwalladr barely uses Facebook, she said. Twitter has been the main vector for the cascading gendered abuse, threats and harassment that she endures. But the fuel is provided by pro-Brexit political actors and donors, whose abuse has been amplified by right-wing media, and even a prominent (now former) BBC Politics presenter. The objective, she said, is to hold her up to ridicule, discredit her, and thereby erode trust in her accountability journalism:

---

132 The *Observer* is The Guardian's companion Sunday national newspaper. Both papers are published by Guardian News Media, along with the Guardian Weekly.

One of the most important things is the way that credentialled people become part of it. For me, it's right-wing journalists and 'trolling' MPs. That's when it becomes really hard to counteract. From the professional point of view, you're perceived as a sort of divisive, controversial figure [whereas] your male colleagues would not have attracted comment. And I think this is very effective.

In this interdisciplinary case study, the researchers conducted a big data analysis of nearly 2.1 million English language tweets directed at Carole Cadwalladr from December 1st, 2019 to January 14th, 2021. Applying the same Natural Language Processing (NLP) techniques that were used to analyse the data in the companion Maria Ressa big data case study[133], 10,400 tweets directed at Cadwalladr were identified by Artificial Intelligence (AI) tools as clearly abusive, and subjected to granular analysis.

This quantitative research was conducted in parallel with longform qualitative interviews conducted with Cadwalladr, *Observer* editor Paul Webster, and Guardian Media Group's Director of Editorial Legal Services Gill Phillips to understand the impacts of the online violence Cadwalladr experiences, and the institutional responses to it. Then, a timeline of events was created to help us map her experience of sustained online violence over the past four years.

The nature of the abuse levelled at Cadwalladr is highly gendered, and at times misogynistic. Most typically, she is maligned and dismissed as a "mad cat lady," mocked for being middle-aged, childless and single; labelled "crazy", "hysterical" and a "conspiracy theorist". All this is "really humiliating", she said. The main goal of the abuse levelled at Cadwalladr is to discredit her professionally, thereby undermining trust in her critical reporting of the Cambridge Analytica scandal and its aftershocks, including questions about the accountability of the platforms regarding disinformation and hate speech.

10,400 TWEETS DIRECTED AT CAROLE CADWALLADR WERE **"CLEARLY ABUSIVE"**

The *Observer's* editor Paul Webster has witnessed the impacts on Cadwalladr: "Carole has been subjected to a fairly sustained and aggressive fusillade of online attacks in her reporting. She's a controversial writer who's chosen a series of very contested areas to report on, and she has been subjected to a great deal of online abuse." Cadwalladr's free-wheeling social media style and her refusal to be silent make her a bigger target. And when she posts controversial comments or corrects occasional errors on Twitter, the abuse becomes torrential. To Cadwalladr, this feels like the digital equivalent of a mob attack on an "unrespectable" woman. **"A few hundred years ago I would have been burned at the stake,"** she said.

The abuse has not been contained to the online realm. Cadwalladr was physically stalked in 2018 by a "shady" man with a military and cyber espionage background who first tried to befriend her, and then began sending threatening text messages (Cadwalladr, 2019a).

She has also been hit with multiple defamation claims by one of the subjects of her investigative reporting - a wealthy businessman who is also a prime instigator of the online harassment she experiences. That man is Leave.EU (the unofficial

133 See the detailed methodology laid out in the full length Maria Ressa big data case study here: https://www.icfj.org/our-work/maria-ressa-big-data-analysis

pro-Brexit campaign) founder Arron Banks. Some of the tweets he directs at Cadwalladr have been received by her as menacing. The defamation cases filed by Banks against Cadwalladr pertain to her commentary on his relationship with the Russian Government. They have been labelled as Strategic Lawsuits Against Public Participation (SLAPPs) by a coalition of eight freedom of expression groups led by Reporters Without Borders (RSF) (UK Anti-SLAPP Coalition, 2021). They described the defamation actions as: "vexatious in nature and intended to silence Cadwalladr's courageous investigative journalism" (RSF, 2019e).[134]

The Guardian and *Observer* have also been threatened with legal action in the course of publishing Cadwalladr's investigative reporting into Cambridge Analytica - including by Google and Facebook (Zerofsky, 2019). However Banks' defamation claims target her as an individual - a typical characteristic of SLAPP suits. They respond to statements she made during a 2019 TED Talk and a tweet in which she shared the talk, rather than articles published by her employer on the same themes (Halton, 2019).

These factors, combined with her status as a freelance columnist at *The Observer* rather than a staff reporter, have left her personally exposed and responsible for her own legal defence. "The thing which I think is important, is that the lawsuit that I have is within this context of a four year campaign of harassment, intimidation, threats of violence... It was kind of like falling down the 'rabbit hole'." In her view, she is considered a threat to these very powerful actors.

Figure 15: Screengrab of a tweet from Arron Banks quotes a tweet from the communications director (Andy Wigmore) of the unofficial Brexit campaign Leave.EU. The Wigmore Twitter account was suspended at the time of writing.

While the Guardian Media Group (which owns *The Guardian* and *Observer*) is not funding Cadwalladr's defence in the SLAPP cases, her editors have publicly defended her reporting through statements like this (Bowcott and Hern, 2018):

Carole's brave reporting has made waves around the world, and given the public much more insight into the secretive ways some powerful people and organisations have sought to influence our democracies... This case is a very worrying example of a wealthy person singling out an individual journalist and using the law to stifle legitimate debate and silence public interest journalism.

134 In June 2022, Cadwalladr successfully defeated the remaining libel action brought against her by Banks in the UK High Court, relying on a public interest defence. See: https://www.judiciary.uk/judgments/banks-v-cadwalladr/ and https://www.theguardian.com/uk-news/commentisfree/2022/jun/19/arron-banks-set-out-to-crush-me-in-court-instead-my-quest-for-the-facts-was-vindicated. However, Banks was subsequently granted the right to appeal the decision and the matter was ongoing at the time of writing.

In the past three years, Cadwalladr's experience of online abuse and harassment has also been influenced by the coronavirus pandemic. She said she has become a "national punching bag". Her fightback strategy in the face of sustained online violence has involved countering disinformation through investigative journalism, forming support networks with other women journalists experiencing online violence, and launching civil society collaborations designed to hold social media companies and political actors to account. It is a "solutions focus," Cadwalladr said.

These collaborations include a network of scientists critical of the UK's response to COVID-19 and the "Real Facebook Oversight Board" - both initiatives of grassroots advocacy group, The Citizens.[135]

> *I'm just trying to process what I've gone through for the last four years... I'm trying to look structurally at the news and information ecosystem and work out how we can counteract disinformation [and] help others who are experiencing this. I really have had enough of being the national punching bag.*

Cadwalladr recruited Maria Ressa - the subject of the other big data case study presented as part of this study - to sit on the Real Facebook Oversight Board (Halpern, 2020). And while their cases are not directly comparable, there are some noteworthy similarities in the patterns of attack, and the role of gendered online violence against a journalist in creating the enabling environment for legal harassment.

In both cases, 'patriotic trolling', niche bloggers/influencers and partisan media figures fuel the attacks, but domestic political actors, not foreign States, are the biggest source of the online violence they experience. As Cadwalladr points out, there is a "misconception" that online attacks come from "faceless foreign bots". She points instead to a "right wing information system that is all powerful in Britain, working at all of these different levels" - which she described as "pernicious and troublesome".

40%

OF THE ABUSE LEVELLED AT CADWALLADR WAS DESIGNED TO UNDERMINE HER **CREDIBILITY & TRUST** IN HER JOURNALISM

12 Key Findings

Types of abuse levelled against Cadwalladr:

- **55%** of obvious abuse detected by the researchers targeted Cadwalladr at the personal level. It was highly gendered and designed to hold her up to ridicule, humiliate, belittle and discredit.

---

135 In October 2021 the non-profit activist group The Citizens (a group of journalists, filmmakers, advertising creatives, data scientists, artists, students and lawyers) [https://the-citizens.com/about-us/] tweeted about various successes, including winning a legal battle to challenge government transparency in the UK High Court https://twitter.com/allthecitizens/status/1448039316010250241

- **40%** of the abuse was categorised as harassment designed to undermine Cadwalladr's professional credibility and trust in her journalism.

- **5%** of the abuse was politically-based.

- **21%** of all obvious abuse (across the personal and professional categories) levelled at her was sexist, misogynistic or sexually explicit.

Characteristics of abuse against Cadwalladr:

- The online violence Carole Cadwalladr experiences is a feature of the enabling environment for her offline legal harassment.

- The abuse was constant and sustained, with several peaks per month delivering intense abuse.

- The cumulative impacts of the sustained online abuse, harassment and attacks over a four year period have created a gaslighting effect, chilling Cadwalladr's investigations and delivering deep personal impacts.

Abuse tactics and triggers:

- Disinformation tactics are clearly in evidence.

- There is obvious evidence of trolling behaviours among abusers, with key identifiable political instigators leading the pack.

- The online violence experienced by Cadwalladr has been instigated by subjects of her investigative reporting, amplified by fringe right-wing media, worsened by conservative columnists, and inflamed by the sexist behaviour of high profile establishment journalists.

- Cadwalladr's internationally consequential, multi-award winning journalism focused upon powerful individuals and entities are what made her a prime target for online violence.

- Cadwalladr is disproportionately attacked for making minor corrections and contentious comments, as the spikes of abuse in the Twitter data demonstrate.

OF ALL OBVIOUS ABUSE AGAINST CAROLE CADWALLADR
21%
WAS SEXIST, MISOGYNISTIC OR SEXUALLY EXPLICIT

The gaslighting effect

Cadwalladr has been subjected to five years of deeply sexist and misogynistic online violence that takes the form of a constant wave of abuse, with several crests each month. This pattern of abuse - building gradually over time - appears designed to destroy the confidence of the target and undermine their credibility without appearing overtly offensive on a tweet-by-tweet basis - a classic case of networked gaslighting. This method typically operates via an organised, or semi-organised, network of abusers for greatest effect (the 'pile-on', 'dogpiling', or 'brigading' approach), and can lead to echo chambers of abuse, where the same abusive message is retweeted many times. These methods are clearly in play in the case of Cadwalladr. Her abusers are frequently interlinked - organisationally, or through association, with pro-Brexit rhetoric.

i. Analysis: Hysterical hag, stupid bitch, crazy cat lady, shut up!

10,400 separate instances of obvious abuse against Cadwalladr were identified by the researchers in a Twitter dataset consisting of 2.1 million English language tweets collected between December 2019 and January 2021. Sexist, misogynistic and sexually explicit abuse represented 40% of obvious abuse in the 'personal attacks' category and 21% of abuse detected overall. It is notable, too, that the pile-on of abuse Cadwalladr experienced included significant elements of anti-journalism and anti-mainstream news media rhetoric, reflecting the demonisation of the press on the global stage, and the weaponisation of terms like "fake news" to chill critical reporting.

Figure 16: The word cloud above shows the most frequently occurring abusive terms tweeted at Carole Cadwalladr (occurring at least 20 times), normalised by case.

The dominant abusive phrases and terms illustrate the gendered nature of the online violence Cadwalladr experienced during the period studied (e.g., variations on "stupid woman", "crazy cat lady", and "witch"). The abuse is clearly designed

to belittle and humiliate her, while also eroding trust in her reporting. The main objective appears to be to intimidate her into silence.

Next, after analysing the clearly identifiable abuse in the 'personal' category (55% of all abuse), the researchers classified it into two subcategories:

- **Sexist, misogynistic or explicit abuse (40%)**
- **Other types of personal abuse (60%)**

40% of the abuse identified is aimed at undermining Cadwalladr's journalistic credibility, especially her alleged "stupidity" as well as her integrity. Typical slurs of this kind include labelling her a "liar" or claiming she is talking "crap", "bullshit", or "bollocks". The objective of this abuse is to scare her off critical reporting. She is frequently told to "fuck off". She is often referred to as "Carole Codswallop" or "Codswallop Cadwalladr", again insinuating that she 'talks rubbish' and her journalism is not to be trusted. But it is the sustained, low-intensity, high-frequency, high-volume nature of the attacks Carole Cadwalladr experiences, not only the content of the abuse, that can be understood to be so cumulatively damaging.

Looking at breakdown by frequency, we established that over 25% of the credibility-related abuse involves abusive terms implying that Cadwalladr is stupid or mentally ill in some form (e.g., "idiot", "twat", "moron", "cretin", "silly", "crazy", "fool", "nutter"), with the term "liar" also appearing over 300 times, alongside terms like "shut up", "STFU" and "fuck off". This language is specifically designed to undermine Cadwalladr's journalistic reputation and professional credibility while also seeking to silence her. 'Sexist, misogynistic, and explicit' terms include examples such as "dickhead", "fuck off", "cunt", "witch", and "hag", along with terms involving sexual acts and intimate body parts.

Instances of political abuse frequently involve anti-Brexit sentiment such as calling her a "remoaner" (a pejorative term for people who supported the campaign for the UK to remain in the EU). Other insults included political abuse such as "Common Purpose globalist whores", "Fuck off you commie twat" and frequent use of "Libtard" (a pejorative term for a liberal thinker) which categorised Cadwalladr as left of centre and lacking objectivity - another way to discredit a journalist in the UK context. Some of the worst abuse levelled against Carole Cadwalladr in this category during the period of this study involved people wishing she would die, as illustrated by this tweet:

> *@carolecadwalla The only failure is scum like you. You [sic] still throwing your toys out the pram that we left the EU. Hope you get COVID and die from it*[136]

Many tweets using such hashtags combined misogynistic language with pejorative descriptions of the mainstream press. Below is an example of a tweet sent to Cadwalladr after she tweeted a link to a story about this UNESCO-supported

136 This tweet, from a now suspended Twitter account, was published November 20th, 2020.

research, ahead of her appearance on a 2020 World Press Freedom Conference panel (Cadwalladr, 2020; Posetti, Harrison and Waisbord, 2020):

> *@carolecadwalla @julieposetti Female? Male? All #scummedia should be treated accordingly - there should be no discrimination when it comes to destroying all you Common Purpose globalist whores. Is there anything lower than you and your ilk? #presstitute #evil*

This now deleted tweet, which also tagged this study's lead author (i.e. it radiated to capture affiliated targets) features the familiar hashtag associated internationally with gender-based online violence against women journalists - #presstitute.

According to the analysis of the Twitter dataset studied, the main themes associated with abuse against Carole Cadwalladr were: democracy; public health and COVID-19; Brexit, Europe and immigration; and foreign affairs. These are also the issues that Cadwalladr most frequently wrote about and commented on during the period in focus.

But one of the most striking features of the abuse against her was not the emphasis on these issues as such, but the deployment of misogynistic tropes. She is frequently labelled "mad" and "hysterical" by critics who emphasised her age, relationship status and childlessness, and questioned her desirability. This sexist 'othering' is a way to treat her as a deviation from the norm. It has also been amplified by some very high-profile figures in the UK press, in addition to popular right-wing blogs. To Cadwalladr, this has all the hallmarks of a witch hunt.

"Codswallop", the "crazy cat lady" and the former BBC presenter

2,921 mentions of the term "cat lady" and its various manifestations (e.g., "cat woman") were detected in the dataset of obvious abuse against Cadwalladr. It is one of the most pernicious and pervasive forms of abuse in evidence, and it was usually used in conjunction with "mad" or "crazy." "I would still say that to this day Andrew Neil is largely responsible for the ongoing misogynistic abuse and threats that I get," Cadwalladr said. She is referring to a now deleted tweet from the then BBC Politics presenter and chair of the parent company of the conservative magazine *The Spectator*, Andrew Neil. In November 2018, Neil tweeted about Cadwalladr, branding her a "mad cat woman" and ridiculing her as "Karol Kodswallop", insinuating that she resembled a character from the satirical cartoon The Simpsons.

According to this research, the Leave.EU donor and founder Arron Banks - who, as described above, pursued SLAPP-style defamation claims against Cadwalladr - was the first Twitter user to disparage the journalist as a "crazy cat lady". He also frequently labelled her "Carole Codswallop". This language was used to denigrate her as she pursued stories probing possible links between the Cambridge Analytica scandal and Leave.EU, the unofficial pro-Brexit campaign that Banks funded. Ten days before Neil called Cadwalladr a "mad cat woman", she wrote in *The Guardian*

that since her investigative reporting began to focus on Leave.EU and Banks, he had called her: "...hysterical, insane, a lunatic, a mad woman, a conspirator, a loony, a mad cat lady, a nasty piece of work, a criminal, a bully, a mad cat lady, a loony, a tinfoil hat nutter, a hacker, a mad cat lady, a loony, a bitter Remoaner, a lone conspiracy theorist, an enemy of the people" (Cadwalladr, 2018c).
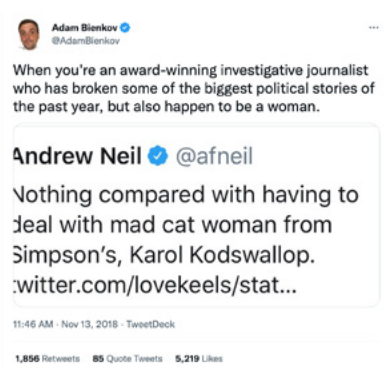
Figure 17: Business Insider's Senior Politics Editor Adam Bienkov[137] defends Cadwalladr after she receives sexist abuse on Twitter from Andrew Neil.

Figure 18: Cadwalladr tweets about the abuse she received following Andrew Neil's 2018 tweet.[138]

This abuse mirrors a pattern of online violence against women journalists that is now recognisable internationally. It involves a process of instigating, amplifying and legitimising gendered online harassment, abuse and attacks. In Cadwalladr's case, the sexist and misogynistic tropes deployed by Banks and his supporters were amplified by the popular right-wing blog Guido Fawkes, and then legitimised by sympathetic voices in the mainstream news media.

Within this right-wing information ecosystem, we see abusive content about Cadwalladr being cross-pollinated through retweeting and quote-tweeting by various actors affiliated with, or sympathetic to, the campaign for Britain's exit from the EU, and/or disparaging towards Cadwalladr's reporting. Notable among them during 2017/2018 are Arron Banks, Leave.EU's Twitter account, Leave.EU's Communications Director Andy Wigmore (who remained suspended from Twitter

137 https://twitter.com/AdamBienkov/status/1062295554946932736?s=20.
138 https://twitter.com/carolecadwalla/status/1064074315828744192?s=20

in mid 2022), the Guido Fawkes blog, a foreign State actor, *The Spectator*, and its holding company chair, Andrew Neil (Murray, 2020; Cadwalladr, 2019b).

Neil deleted his "mad cat woman" tweet following a backlash, but he did not apologise (Mayhew, 2018a). The BBC's corporate communications team did, however, issue a tweet in which they stated that Neil recognised the tweet was "inappropriate" (BBC News Press Team, 2018). The following day, the then BBC Director General Tony Hall told a conference (BBC Media Centre, 2018):

> *On Twitter there are constant anonymous threats to journalists simply reporting on opinions that some people might not want to hear. Some of the material that journalists have had to face is quite frankly disgraceful. It is an attempt to intimidate people and stop them doing their jobs. For the sake of all journalists - we need to defend our role - seeking out the facts, no matter how inconvenient they may be for others. Because journalism matters - whether you're in broadcasting, in the press or working online.*

The next week, *The Guardian* reported that a number of senior women journalists at the BBC had complained to executives about Neil's tweet, suggesting inequitable treatment regarding social media policy enforcement (Waterson, 2018; BBC, 2018e). A BBC spokesperson was quoted as saying: "The tweet was sent from Andrew's personal account, however the BBC has social media guidelines which it expects all staff to follow and these have been discussed with Andrew."

Meanwhile, the misogynistic insult continues to reverberate: "I am now called a 'mad cat lady' dozens of times a day on social media, every day," Cadwalladr said (Cadwalladr, 2018d). In addition, "Codswallop", and Neil's variant "Kodswallop", are also pejorative nicknames still in regular use in abusive tweets against Cadwalladr according to this research.

ii. Who is sending the abusive tweets targeting Carole Cadwalladr?

Of the 10,400 obviously abusive tweets identified in the dataset, there are 7,744 unique authors. Of the serial abuse senders, the most prolific sent 38 obviously abusive tweets, while 155 tweeters sent five or more that the NLP tool categorised as abusive. Only two of the top eight most prolific abuse senders in the dataset still have live Twitter accounts. Of the other six, three have been deleted and three were suspended at the time of writing. Leave.EU's Director of Communications Andy Wigmore was the Twitter user who abused Cadwalladr most prolifically during the period. He sent 38 highly abusive tweets before his account was suspended on November 8th, 2020. Here is one of the abusive tweets Wigmore sent before his suspension, in which he alludes to SLAPP suits and tags both Andrew Neil (who triggered the "mad cat woman" pile-on of 2018) and Arron Banks, who sued Cadwalladr for libel:

Pointing out facts by @afneil (a proper journalist) must be like a SLAPP round the chops or scraping your nails down a chalk board Codswallop @carolecadwalla get used to it more facts coming your way. @Arron_banks

The second most abusive account was an anonymous Twitter account which overlapped with another anonymous account using the same profile photo and an almost identical Twitter handle. Both accounts were suspended in late 2020. Out of the 7,744 authors of obviously abusive tweets in the dataset, as of March 15th, 2021, 943 of these had deleted their accounts, and 1,021 had their accounts suspended. In total, this means that just over 25% of the accounts are no longer active. This suggests that those accounts abusing Cadwalladr most prolifically are people whose misconduct on the platform has been relatively extreme, generating a high removal response rate from Twitter. Authors of abusive tweets also: have more recently more recently established accounts than tweet authors in the non-abusive set; have fewer followers; follow fewer users; and post slightly fewer tweets. Viewed together, this data pertaining to the behaviour of the Twitter users in the dataset who were obviously abusive towards Cadwalladr during the period under examination indicates clear patterns of trolling behaviour with identifiable instigators and 'ring leaders'.

Figure 19: Cadwalladr says she is a target of a toxic right-wing media ecosystem in the UK.

iii. Abuse spikes and triggers

Cadwalladr is an outspoken and reactive Twitter user, reflecting her status as a high-profile columnist at the *Observer*, and she is relatively unfettered in her social media conduct compared to journalists bound by corporate social media policies because she is not a staff writer. At the same time, she is also a woman and a journalist who feels battered after years of exposure to unrelenting online violence. "It's changed my life," Cadwalladr said.

Early on, she was hacked; she has been stalked; she believes that the trolling has seriously damaged her reputation within the UK news industry; she acknowledges her reporting has been chilled; she has experienced the trend of victim blaming - judged for "having answered back on Twitter" - which is slowly muting her social media use; and two of the whistleblowers at the heart of her Cambridge Analytica investigations were also targeted in coordinated online and news media attacks designed to undermine their credibility, she said.

All of this was going on as Cadwalladr was winning a cache of premier British journalism awards, such as the 2018 Orwell political journalism prize, and she was recognised as a Pulitzer Prize finalist with the *New York Times*. Then, she faced possible bankruptcy, as defamation and libel actions brought by one of the powerful subjects of her award-winning investigative reporting wended their way through the courts. "I'm processing actually what it's like being under attack for four years. At various points, I think I have overreacted to things as well, you know. You're so used to being attacked that you respond like you are on a bit of a trigger."

Her realtime reflections are weaponised against her, as are any realtime errors. Cadwalladr corrects the few mistakes she does make but she is offered no forgiveness, nor shown any grace on Twitter by her dedicated detractors and trolls. The dataset demonstrates a continuous ebb and flow pattern of abuse, frequently associated with a backlash against her commentary, with two to three attack spikes per month. This sort of digital gaslighting - a long-range persistent pattern of abuse - is designed to wear the target down.

As a woman covering the global political, social and human rights ramifications of big technology and viral disinformation, and historically divisive political movements (on both sides of the Atlantic); who dared to break major investigative stories as a features writer and columnist, Cadwalladr could be seen as a natural lightning rod for online abuse in the toxic communications ecosystem of the early 21st century. But while her reporting and tweets are often cast as "controversial", the abuse she receives is an entirely disproportionate response to her occasional errors and her engaging, reflexive, opinionated and consequential writing.

iv. The role of The Guardian and Observer

The online violence Carole Cadwalladr is subjected to is designed to silence her, chill her critical journalism and discredit those who publish it, namely: *The Guardian* and *Observer.* In August 2020, Guardian Media Group took legal action to shutdown a website generating fake *Guardian* headlines and byline profiles which were being shared with the trending Twitter hashtag #TrollingTheGuardian. Carole Cadwalladr was one of the journalists targeted by the now defunct guardianmeme.com website. But the action was triggered by a tweet from a UK member of parliament who was similarly attacked after she wrote a *Guardian* column (Mayhew, 2020b). The MP tweeted that the site put her in danger, along with her family and staff.

Guardian Media Group's decision not to fund Cadwalladr's defence in the ongoing defamation cases - the enabling environment for which includes the online violence she experiences - is a source of grief and frustration for Cadwalladr, who regularly pointed out that she could "lose her house" as a result of the litigation.

The company's position was that because the defamation and libel claims were made against Cadwalladr as an individual, in relation to speeches (reviewed by her employer prior to delivery) and a tweet, not reports carrying similar claims and imputations published by the outlets, they are not a party to the legal action, and they are unwilling to set a precedent by funding Cadwalladr's defence.

"If she had been sued individually over something she wrote for us, we would be all over it. But if we do this for her, where do we stop?" Editorial Legal Director Gill Phillips asked. "So we've had to say, really regrettably, we cannot financially support this, but we will give you whatever support we can morally, ethically, including every time we do anything on SLAPPs. And we will continue to defend the journalism she does for us."

However, Cadwalladr said that her concerns about her employer's response to the online violence and associated legal harassment she experiences, are not purely financial. She also described what she perceived to be a lack of organisational support. Her response has been to crowdfund her legal defence (GoFundMe, 2019). She raised several hundred thousand pounds from supporters. She is also working proactively to advance protocols for dealing with online violence against women journalists at *The Guardian* and *Observer* - emphasising the function of disinformation campaigns in such attacks, along with the experiences of more isolated freelancers, and the need to publicly defend women journalists in the midst of attacks.

v. The role of the platforms

Cadwalladr said she is very frustrated with the big tech companies that act as vectors for online violence, viral disinformation and privacy erosion. Facebook and Google have been the focus of her investigative journalism while the online attacks she experiences are facilitated by Twitter. "I report stuff all the time on Twitter and they never ever do anything. They never take it down," she said. "There's these endless spoof accounts...I'm always reporting stuff." She was particularly aggrieved by Twitter's failure to remove a Leave.EU deep fake video in which she was depicted being repeatedly slapped. The video, which featured one man with a gun and one with a hammer, was clearly "incitement to violence," in Cadwalladr's view. In the end, it was *The Guardian'*s intervention in contacting the source of the post that resulted in its removal, 48 hours later, she said.

Believing that reporting, blocking, deleting and muting are acts of futility against self-regenerating troll armies, Cadwalladr has turned to grassroots campaigning and the development of loose support networks through her work at The Citizens. This has also involved working with data scientists she has met on Twitter to map networks of abuse and develop response mechanisms to support other users under attacks. Concerning Facebook and the impunity with which the company has been able to act, Cadwalladr said representatives repeatedly lied to her in the course of the Cambridge Analytica investigation, and she believes that the company is just too big to hold accountable:

The FTC fined them a record US$5billion but it had no impact whatsoever. Their share price actually went up...because there is no mechanism to hold them to account. The only accountability structure we have has no effect because the platforms are too big. You have got to find other ways, and more creative ways essentially.

Her alternative approach has been to launch the "Real Facebook Oversight Board" - a collection of academics, civil society experts, social media critics and prominent journalists who do not have faith that the Facebook-funded official Oversight Board and its limited remit will do the kind of urgent accountability work needed, with appropriate transparency, and at the scale required.

One thing that needs to change within the social media companies is resourcing to deal with online abuse and harassment, according to *Observer* Editor Paul Webster. "They need far more people," he said. Adding: "They need to recognise their roles and responsibility as publishers and act accordingly. So curate that material in the way that we do - take legal responsibility for it. So if somebody is libelled in tweets and in online posts, then you have recourse to the people who carry the material as well as the people who make the libels."

While *The Guardian* has had some limited success in getting the platforms to remove abusive content in certain cases, and deplatform individual offenders in others, the scale of the problem is not able to be managed by present systems, according to Guardian Media Group's Director of Editorial Legal Services Gill Phillips: "Look, of course, we can write letters and ask people to stop it, and point out it's horrible. [But] the only thing that can stop it is the people who are giving them the platforms that allow them to make these attacks."

4

What more can news organisations do?

Julie Posetti and Nabeelah Shabbir

Given the gravity of online violence targeting women journalists demonstrated by this study, and in the face of new threats such as viral disinformation, far-right social networks, and digital conspiracy communities, it is essential to assess news organisations' responses. Research underpinning this chapter indicates that many newsroom reactions to gender-based online violence appear to have been non-existent, ad hoc, or inadequate. At times, they have even damaged the women journalists targeted.[139]

Large global news organisations sometimes identified as "best practice" exemplars by expert responders interviewed for this study were nevertheless criticised by the journalists interviewed in the course of the research with regard to their responses to the crisis. They were accused of failing to fully understand the gendered nature of the attacks, appreciate the serious psychological impacts, adapt to emerging and increasingly sophisticated threats, and provide effective and holistic support that recognises intersectional risks and hybrid security threats. A number of outlets were also criticised for insensitive and counterproductive victim-blaming and/or speech-restrictive behaviours.

Many of the journalists interviewed for this study expressed exasperation and a sense of abandonment by their employers when they were in the midst of an online violence storm, even when there were credible threats of offline violence associated with these attacks. This was linked to gender-unaware policies, or those that had stagnated as a result of a failure to take account of increasing online toxicity and hostility towards journalists - especially on social media platforms - in the context of escalating disinformation, along with political polarisation and populism. The COVID-19 pandemic further heightened the risks, as news organisations and journalists became more reliant on digital communications platforms for socially distanced reporting and audience engagement (Posetti, Bell, and Brown, 2020).[140]

This chapter explores the systemic reasons for these failures, identifies impediments to better practice, and develops evidence-based recommendations for more effective responses by news organisations.

There are wider obstacles to more effective newsroom responses (e.g., structural failures associated with Big Tech's management of online violence, and the need for legal and legislative reform). These aspects of the problem have been well-documented as it has evolved (See: Citron, 2014; Posetti, 2014; Löfgren, Nilsson, and Örnebring, 2016; Posetti, 2016; Gardiner, 2016; Posetti, 2017a; West, 2017; Adams, 2018; Gardiner 2018; Chen et al., 2018; Jane, 2018; Posetti, 2018a). Such

---

139 An earlier version of this chapter was published by UNESCO and ICFJ on 3 May 2022. It is available here: https://en.unesco.org/publications/thechilling
140 A comprehensive bibliography accompanying this study is published separately by ICFJ *here*

considerations signal that newsrooms not only have to ensure effective comment moderation on their own websites but, among other things, develop preventive and responsive strategies that recognise that the primary site of online attacks against women journalists is now social media. Accordingly, there is a need to shift the onus for managing gendered online violence from the individual journalists under attack to the news organisations that hire them, the political actors who frequently instigate and fuel attacks, and the digital services that act as vectors for abuse. News organisations at least have to navigate all these dimensions.
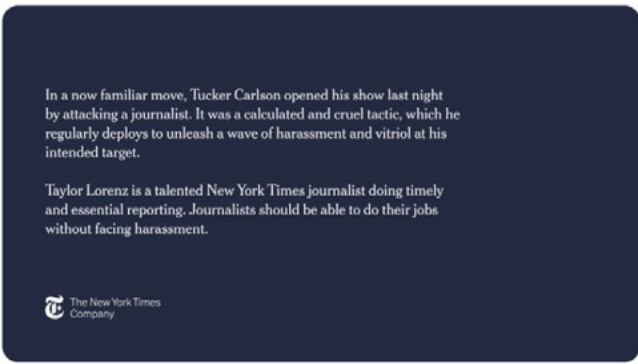
Emerging signs of a positive shift in news organisations' responses to gendered online violence

As this research was concluding in late 2021, there were fledgling indications that the proforma newsroom response to online attacks against women journalists - "don't feed the trolls" - was being recognised as inadequate. This approach, which has been subject to increasing criticism from women journalists, has had the effect of constraining those who might seek to defend themselves publicly in the midst of an attack, and leaving them feeling more exposed. It has also served to diminish the seriousness of online violence by suggesting that it was possible or necessary to ignore online attacks.

Four indications of more empowering and effective responses from large news organisations and one notable industry collaboration were identified during the final phase of this research:

- **Employers publicly defending their journalists:**
  Major outlets like the *The New York Times*[141] and *The Washington Post*[142] have used their editorial power to issue public statements of support for women journalists in the midst of targeted online attacks. NBC News[143] and HuffPost UK[144] also issued public statements in defence of their journalists while they were experiencing significant online violence, criticising the perpetrators.

In a now familiar move, Tucker Carlson opened his show last night by attacking a journalist. It was a calculated and cruel tactic, which he regularly deploys to unleash a wave of harassment and vitriol at his intended target.

Taylor Lorenz is a talented New York Times journalist doing timely and essential reporting. Journalists should be able to do their jobs without facing harassment.

The New York Times Company

Figure 20: A statement released by The New York Times in March 2021, in defence of then NYT technology journalist Taylor Lorenz, after she was targeted on television by Fox News' Tucker Carlson in an attack that fuelled online abuse and harassment.

141 https://twitter.com/NYTimesPR/status/1369747504565256193?s=20
142 https://twitter.com/kriscoratti/status/1365040361572339714?s=20; See also this full page advertisement defending Indian investigative journalist and Washington Post columnist Rana Ayyub in February 2022: https://twitter.com/wppressfreedom/status/1495395937820282884?s=20&t=QalGRxlMjKBKs9ysoyjO2Q
143 https://www.poynter.org/ethics-trust/2020/nbc-news-defends-a-reporter-after-a-bizarre-tucker-carlson-criticism/
144 https://www.huffingtonpost.co.uk/entry/kemi-badenoch-twitter-huffpost_uk_603780a1c5b69ac3d35ca3e5

The Washington Post

On Wednesday morning, an old tweet sent by Neera Tanden became an issue in her effort to be confirmed as the head of the Office of Management and Budget. That afternoon, White House reporter Seung Min Kim bumped into Murkowski in the Capitol. She asked the senator general questions about outreach from the White House regarding Tanden's nomination and asked about Tanden's tweet. Murkowski had not seen the tweet, so Seung Min showed it to her and Murkowski's reaction was included in a Post story about Tanden's nomination.

A picture of Seung Min showing the tweet to Murkowski was posted to Twitter, which led to a flood of racist, sexist and ill-informed attacks aimed at Seung Min. The racist and sexist attacks have been vicious -- and typical. She and other minority women endure vile, baseless attacks on a daily basis, no matter what story they are working on or tweeting about. The attacks on her journalistic integrity were wildly misguided and a bad faith effort at intimidation.

What she did was basic journalism. In reporting about Tanden's tweet, she asked Murkowski for comment. Murkowski had not seen the tweet, so Seung Min showed it to her. This is standard practice. If a subject of a story is not aware of the information they are being asked to comment on, reporters share it with them. This only makes sense and is the fair and responsible thing to do.

No one should have to deal with the hate that has been directed at Seung Min. She did her job, and she did it well, like she always does. We could not be prouder that she is our colleague and a reporter for The Washington Post.

-Steven Ginsberg, National Editor

Figure 21: A statement released by The Washington Post in February 2021 in support of White House correspondent Seung Min Kim, who received racist abuse online for what the Post said was practising "basic journalism".[145]

- **Increasing attention on intersectional threats:**
  News organisations like the BBC, CNN, CBC (Canada), and Al Jazeera acknowledged the need to increase security and consider intersectional threats in the context of escalating online violence associated with COVID-19-related digital conspiracy networks: "We know these attacks are more often aimed at women and journalists of colour, so we want to make sure we have particular support for those groups and are looking at what this could be," BBC Director of News and Current Affairs Fran Unsworth told staff in mid-2021 (Townsend, 2021).

- **New roles created in newsrooms to more effectively address the crisis:**
  For example, the UK's biggest commercial news publisher Reach PLC advertised for an Online Safety Editor in June 2021. The Group Editor-in-Chief said: "We will not allow our journalism to be silenced or our colleagues to live in fear"[146] (Tobitt 2021b).

In October 2021, former *Yorkshire Post* head of news Dr Rebecca Whittington was hired by Reach PLC in 'the industry-first role' of online safety editor (Tobitt, 2021d). Six months into her role, she described how she was creating an abuse database of sorts to improve understanding and prevent prevalent abuse types, ensuring protocols and procedures were "up to scratch", and seeking to update abuse reporting and newsroom awareness systems, while also taking the audience's safety into account (Tameez, 2022).

---

145 Tweet from Seung Min Kim in February 2021, who was told in an email to 'Go back to China, b*tch: https://twitter.com/seungminkim/status/1364762782554587141

146 In October 2021, former *Yorkshire Post* Newspapers head of news Dr Rebecca Whittington was hired in 'the industry-first role' of Online Safety Editor: https://pressgazette.co.uk/reach-hire-industry-first-online-safety-editor-tackle-abuse/
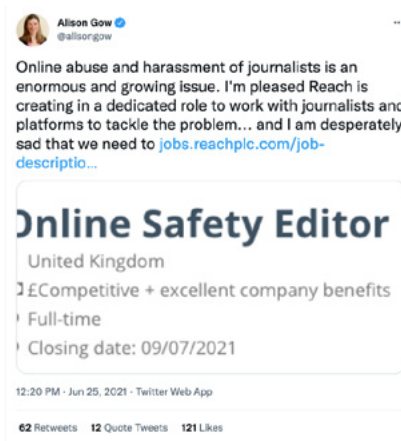
Figure 22: Tweet from Alison Gow, Regional Editor-in-Chief (Digital) for Reach PLC in June 2021.

- **Reporting on the crisis:**
  Various outlets (detailed later in this chapter) have published editorials and reportage reflecting on the first findings from this study, decrying gendered online violence, critiquing the platforms' role in the crisis, and calling for new approaches to combat it. This was an important demonstration of advocacy journalism focused on a press freedom and journalism safety threat directly affecting their own staff.

- **Canadian news industry collaborative campaign:**
  In late 2021, Canadian news organisations collaborated on an industry initiative designed to respond to gender-based online violence against journalists. Coordinated by the Canadian Broadcasting Corporation and Radio Canada, the effort involved a conference designed to facilitate a collaborative industry response to the crisis in Canada (CBC/Radio-Canada, 2021), and specially commissioned research (Ipsos, 2021).

i. Methodology and research context

In order to evaluate news organisations' responses to gendered online violence against journalists and develop recommendations for better newsroom practice, a large data corpus associated with this study has been synthesised and analysed for this chapter. It consists of:

- Responses to a UNESCO-ICFJ survey (conducted in late 2020)[147] focused on questions regarding newsroom practices and journalists' experience of them;

- 15 detailed country case studies[148] produced to underpin this study which scanned news organisations' responses in diverse national contexts;

- Longform interviews with 113 international journalists and editors exposed to online violence or tasked with managing it,[149] and;

---

147 https://en.unesco.org/themes/safety-journalists/women-journalists/global-survey-results
148 Produced by regional research collaborators identified on the inside cover.
149 These interviews were conducted by chapter co-author Julie Posetti, co-editor Nabeelah Shabbir and several of the international researchers identified in the attributions at the start of the chapter.

- 20 additional interviews (conducted specifically for this chapter)[150] with key news organisation responders, and external digital safety experts who deliver online violence defence training to newsrooms. These interviewees - 12 women and 8 men - are senior editors and managers from newsrooms and investigative journalism organisations in six countries, as well as senior members of civil society and investigative journalism organisations based in Europe and North America.[151]

ii. Hits and misses: Assessing newsroom responses to gendered online violence

Although 73% of women survey respondents to the UNESCO-ICFJ global survey underpinning this study said they had experienced online violence in the course of their work, only a quarter (25%; n=179) had reported this to their employers. This indicates that many women journalists, who are frequently structurally disempowered and disadvantaged in newsrooms, only escalate the more extreme instances of online violence with their employers. The reluctance to report and escalate when attacks occur can be linked to systemic failings such as unsympathetic, misogynistic, patriarchal or otherwise hostile workplace cultures, poor leadership, a lack of clear and established reporting procedures, and/or a lack of a formal protocol to deal with the problem.

ONLY 1/4 *OF WOMEN JOURNALISTS SURVEYED* REPORTED ONLINE VIOLENCE TO THEIR EMPLOYERS

The survey respondents highlighted a double impediment to effective action from media employers: low levels of access to systems and support mechanisms for targeted journalists, and low levels of awareness about the existence of measures, policies and guidelines addressing the problem (Posetti *et al*., 2020a). If women under attack do not know how, nor to whom, they should report such attacks, or have reason to believe that their alerts will not be responded to sensitively or effectively - possibly even impacting negatively on their careers - they might resign themselves to silence, and silence can effectively perpetuate the failure to develop adequate systems.

The responses that the women survey participants reported receiving when they did report online violence to their employers were, on the whole, very unhelpful. 10% (n=71) of the respondents said their employers did nothing at all. Very few women who reported online violence were offered support by their employer: only seven respondents were offered counselling, time off work to recover, or physical security. It is concerning that so few participants who reported online violence were offered psychological support, considering the top ranked impact of online harassment and abuse indicated by survey respondents was mental health issues (26%; n=186). Meanwhile, only 21 respondents who had reported online violence incidents said they were given any digital security support. This highlights a significant gap in newsroom responses considering the risks to their journalists and their sources posed by digital security threats, along with the propensity for doxxing to lead to physical violence. One in five women journalists surveyed (n=119) said they had experienced offline attacks, harassment and abuse associated with online violence.

---

150 Chapter co-author Becky Gardiner conducted these interviews.
151 901 self-selecting international journalists, editors and experts responded to the survey fielded by UNESCO and ICFJ between September and November 2021.
714 of them identified as women and it is that cohort whose responses are reflected in the statistics presented in this publication.

Victim-blaming, speech-restriction and punitive social media policies

According to interviewees and survey respondents participating in this study, there remains a pattern of victim-blaming associated with online violence against women journalists. This in effect makes the targets responsible for the violent behaviour of others and puts the onus on them to respond to the problem by modifying their behaviour to protect themselves. It is a parallel to treatment women still often encounter in the context of sexual harassment, assault or domestic violence. Ten per cent (10%) of survey participants said they were told to "toughen up" or "grow a thicker skin" (Chen *et al.*, 2018) when they reported online violence to their employers. This reflects a failure within many news organisations to take the crisis of gendered online violence seriously as a workplace safety threat (Jane, 2018).

A Kenyan journalist who chose to remain anonymous was clear on this point: "Having a thick skin does not protect you from a personal attack that leads to your data being shared and someone promising that they will rape you". Northern Irish freelance crime and investigative journalist Patricia Devlin, who was a staff reporter at *Sunday World* when she was interviewed in mid 2021, also spoke with clarity on the futility of the standard "grow a thicker skin" advice:

> *The last two years, I have lived every week of my life getting abused and then getting death threats. And [you] just start to say to yourself, is this ever going to end, is it ever going to stop? I just feel hopeless. ...Can I continue on like this? I can't do my job without being on social media, so I can't come off social media. That's not the answer. ...And now the death threats. I just can't see how I can sustain that... . You try [to] brush it off. And I'm very thick skinned. But wouldn't it be good if I just didn't have to go through that?*

Fourteen of the women survey respondents and several interviewees said their employers had asked them what they did to provoke the attacks, highlighting the perpetuation of victim-blaming, along with women's ongoing subjugation.

Corporate social media policies that problematically restrict journalists' reactions to attacks and incorporate punitive measures including disciplinary action and dismissal (Tobbitt, 2020e), are becoming more commonplace globally. These policies tend to focus on the obligations of the journalist rather than preserving their safety, freedom of expression, or dignity. Sometimes they require journalists to avoid engaging with their attackers, expressing opinions, and discussing controversial matters. For example, the BBC's social media rules for journalists,[152] include instructions to: "always treat others with respect, even in the face of abuse" and tell journalists not to "be drawn into ill-tempered exchanges" and "avoid 'virtue signalling'". They were highlighted as problematic by a number of interviewees. These constraints may be misconstrued by news executives as appropriate

---

152 https://www.bbc.co.uk/editorialguidelines/guidance/individual-use-of-social-media

measures to avoid escalating episodes of online violence. However, such rules can undermine the agency of women who are under attack, and contribute to a damaging feeling of powerlessness, potentially exacerbating mental health impacts suffered due to online violence.

Five BBC, Washington Post and New York Times journalists interviewed expressed concern about their freedom of expression being chilled by their own news organisations, and the sense of disempowerment that resulted. One said the stringent guidelines which discouraged engagement with, or criticism of, those who attack journalists left her feeling almost like her employer was targeting her; another said she felt she was being 'tone-policed', "so I cannot be too aggressive with somebody else, even if I'm just defending myself". A third said she had been disciplined for her tweets, even in the context of ongoing online attacks which involved partisan journalists from competing news outlets campaigning for her dismissal.

Research for this study surfaced several examples of women journalists being sidelined, disciplined, suspended and even fired in the midst of pile-ons, which they believe was a result of such policies, or the insensitive implementation of them. Among them are the cases of Lauren Wolfe at *The New York Times* (Gabbatt, 2021), Amy Wilder at AP (Place, 2021), and Felicia Sonmez at *The Washington Post* (Holpuch, 2020). These episodes, along with anonymous accounts gathered from women journalists in the course of research for this study, suggest that it might be appropriate for online violence response teams to sensitise editorial leadership to the need to support women under fire in such circumstances, and to view patterns of behaviour that might otherwise result in social media policy breaches in the context of devastating work-related online violence.

Additionally, there is an element of inconsistency in the application of these policies. This is a point that *Observer* investigative journalist Carole Cadwalladr has raised with her editors: "The news organisation's response to attacks on journalists has been ad hoc and inconsistent. It has publicly supported some individuals on some subjects, but not others". *The Washington Post's* Felicia Sonmez also underlined this point.
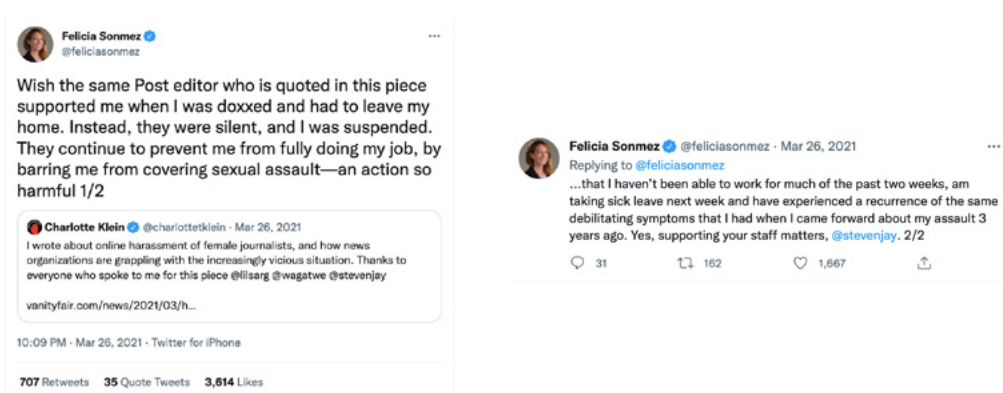
Figure 23:[153] Tweet from Felicia Sonmez, sent in March 2021, critical of what she describes as the double standards which led to her being put on administrative leave from The Washington Post *over her tweets regarding a news story about sexual assault allegations.[154]*

---

153 https://twitter.com/feliciasonmez/status/1375555534879866894
154 Sonmez, a sexual assault survivor, was placed on administrative leave for tweeting about historic sexual assault allegations against Kobe Bryant at the time of his death in early 2020. Sonmez said she received thousands of abusive social media messages, including death threats, and had to flee her home after she was doxxed. See: https://www.cbsnews.com/news/washington-post-guild-defends-reporter-placed-on-leave-after-tweeting-about-rape-allegation-against-kobe-bryant/

There are also cultural factors which need to be considered when assessing action taken against women journalists in response to their online abuse. In Pakistan, for example, Gharidah Farooqi - a respected TV anchor - lost her job after online slurs about her sexual morality, which were followed by accusations of bias in her coverage. It took her 18 months to find another job in the news media (Butler and Iftikhar, 2020). Nigerian documentary reporter Ruona Meyer, who has worked for a range of news organisations, said she had chosen to remain freelance so that she could respond to racist and misogynistic online violence on her terms, including by using counter-speech "without fear of breaching employer guidelines regarding online conduct".

This highlights how newsroom protocols designed to address online abuse and harassment can be undercut by guidelines on journalists' use of social media. Michelle Stanistreet, Secretary General of the UK's National Union of Journalists (NUJ), said that too many employers were ignoring gendered online violence as an aspect of their duty of care for their employees, while taking punitive action against women under attack. The pendulum needs resetting, she said, relaying instances where being abused and harassed had landed journalists "in hot water when they snap or retort back on social media". This involved employers stepping in "from the perspective of: you're somehow damaging our reputation and we'll resort to disciplinary processes".

The Observer's Carole Cadwalladr called for a "two-way relationship with news organisations" and said a "rapid response unit" could ensure that when a journalist is attacked, the "obligation" comes back to employers as well.[155] Guidelines to this effect could therefore be embedded in online violence response protocols. US sports journalist Julie DiCaro, who received rape and death threats for reporting on the case of an ice hockey player's sexual assault court case, also referred to the need for a two-way social media policy: "What can employees expect from their bosses when online harassment begins? Will you help protect me? Will you back me up if I go to law enforcement? Will you provide security for me if I need it? What will you do if you find one of your employees is using social media to troll people? Will you use your relationship with platforms like Twitter to advocate for me? Can I take a mental health day [leave] if I'm in the middle of a storm of abuse?" (DiCaro, 2021)

Globally, 38% (n=271) of the survey respondents described reacting to online violence in ways that made them less visible. When responses are triggered by employers - however well intentioned - the situation can be a double-edged sword. After reporting instances of online violence to their employers, some survey respondents were told to take time off social media, others were taken off air or had their bylines removed. In two cases, a woman journalist who had come under attack had her byline handed to another journalist on an exclusive story as a 'preventive' measure. While these measures are sometimes recommended to help mitigate the trauma, they can also serve to render women journalists invisible, and undermine their ability to do their jobs, while also limiting career development.

RETREATING FROM VISIBILITY WAS AN ONLINE VIOLENCE SURVIVAL TACTIC IDENTIFIED BY

38% *OF WOMEN JOURNALISTS SURVEYED*

---

155 See the big data case study on Carole Cadwalladr published in Chapter 3.2.

Concerns about freedom of expression can lead to institutional reluctance to confront online violence

Traditionally, news organisations are reluctant to take action (including legal action) associated with limiting speech because of the propensity for powerful actors to use legal mechanisms (e.g., defamation law) to chill critical independent journalism. This has led to a false binary between 'free speech' on the one hand, and the protection of press freedom and the UN-mandated right to safely practice journalism (online and offline) on the other hand.[156] However, freedom of expression rights enshrined in international law do not uphold the right of a person to use hate speech and threats of violence to limit the speech of another person. This is particularly relevant in the case of journalists, who also frequently benefit from press freedom protections enshrined in international law.

Nevertheless, there is a reluctance for news organisations to call out abusers. Particularly in the US where 'First Amendment absolutism' makes women journalists reluctant to report online violence, according to those interviewed. *The New York Times'* Jason Reich articulated this problematic tension from an organisational perspective:

> *As First Amendment organisations, we're obviously incredibly hesitant to [support anything] that can be used to restrict speech. [...] we are very hesitant even to try and get... a Twitter account banned or something [...] The New York Times does not want to be in the business of banning voices. So that's a major obstacle for us. But it's one that's, you know, self-imposed.*

This "hesitancy" is heightened by consciousness of power asymmetries between news organisations and many individuals, Reich said.

Sexual harassment and structural sexism discourage reporting of online violence incidents

Patriarchal norms still permeate newsrooms internationally. While women dominate journalism schools and entry level positions in many newsrooms, they remain seriously underrepresented in top management positions where editorial and workplace policies are devised (Global Media Monitoring Project, 2020). They also continue to be subjected to physical manifestations of sexual harassment and assault in the course of their work - including at the hands of colleagues and supervisors. Some interviewees described groping and sexual harassment as being commonplace in their newsrooms. This makes reporting online violence both intimidating and futile for many women journalists around the world. Where women's physical safety within the workplace is not taken seriously, it is unlikely

---

156 See detailed discussion in Chapter 5.0.

that online violence perpetrated by outside actors will be addressed. "In some countries, workplace harassment is such a problem; that would need to be tackled first," Pamela Morinière of the International Federation of Journalists said.

However, even where overt sexual harassment from colleagues within the workplace is not tolerated, sexism in newsroom cultures may still prevail, making women journalists reluctant to report online violence and doubtful that it will be taken seriously if they do so. Many interviewees spoke of prevailing structural workplace inequality, saying they struggled to be given 'hard news' roles, for example. Others referred to the ongoing power of the newsroom 'boys club' which works against the promotion of women journalists to leadership roles. For example, Lebanese women journalists reported having to constantly fight being further marginalised within their male-dominated profession in a context where stigmatisation and sexualisation of women is entrenched (Melki and Farah, 2014; El Hajj, 2019). One of the Lebanese journalists interviewed, Youmna Fawaz found through her editorial supervision work that harassment often begins within the news organisation itself, coming from people in managerial positions.

Barriers to addressing intersectional abuse

One very clear point of agreement among the women journalists interviewed who experienced intersectional online abuse[157] came back to the role of newsrooms. Sri Lanka's Methmalie Dissanayake, Deputy News Editor at Ceylon Today, who has been accused online of wanting "to ruin Buddhism and Sri Lankan culture", suggested one reason Sri Lankan newsrooms were not addressing online abuse against women reporters was because journalism was not highly professionalised across the country and it was easier for them "to find new people rather than keeping…the [journalist] who faced this kind of violence".

In the US, a reporter who asked to be anonymous, was labelled a "terrorist… promoting Islam or Sharia law" and suffered other online abuse after interviewing local Muslim groups. Fatima Hussein, president of the Washington Baltimore News Guild at the time of her interview for this study, said that those "traumas" have to be "absorbed" by women journalists from marginalised ethnic communities: "Since there is this lack of diversity in editorial ranks and among reporters, even the notion of complaining about ill treatment is frightening because I want to keep doing my job and I don't want something to be taken from me. I don't want my beat to be taken from me. And so I think there's an expectation that you have to deal with it."

At the BBC, despite the existence of a well-established protocol for responding to the mental health impacts of online abuse, the researchers identified gaps regarding intersectional abuse, reflecting a failure to properly account for the additional risks and impacts faced by women journalists from diverse backgrounds in an organisation lacking in diversity (James, 2021). On Twitter and via Gmail, Black journalist Rianna Croxford was called a "monkey": "the N-word… mocking

---

157 Intersectional online violence refers to attacks that target women who experience multilayered forms of discrimination e.g., racism and religious bigotry in addition to misogyny.

my appearance in different ways or criticising my ability… 'That's crap'. 'Shouldn't be on air', 'isn't qualified'…". The online violence experienced by Croxford came amidst intimidation from political actors and targeting by partisan news media and fringe blogs, which led to a political campaign for her dismissal. "I felt like I had to be silent because it's the BBC. You don't want to bring it into disrepute," she said. "As a journalist of colour, I sometimes feel you've got to work harder, that you can't afford to make mistakes, and this feeling suddenly felt amplified."

Fran Unsworth, BBC Director of News and Current Affairs, wrote to staff in June 2021, after a white male journalist was physically attacked by anti-lockdown protesters, which radiated to other reporters who were doxxed and harassed online, including on channels such as Telegram. Unsworth acknowledged the additional exposure of women and journalists from minority ethnic groups to online violence, and she said efforts would be made to adapt support mechanisms to recognise this (Townsend, 2021).

When old tweets from Guardian US Senior Reporter Julia Carrie Wong were surfaced by reporters at Breitbart, her editors put out a statement supporting her. "[Which] I was also appreciative of, because they made it clear that they were on my side and that they weren't going to bow down to Breitbart calling me racist," Wong said. At the BBC, Rianna Croxford later had tweets from her time as a student unearthed by trolls (and used by prominent politicians) as 'proof' that she was not politically impartial following an article she had written. Croxford said she had never been found to have breached editorial guidelines in her time as a BBC journalist. The BBC social media guidelines provide for this scenario; stating that tweets from a journalist's life before they worked for the corporation are not taken into account when enforcing social media policies that demand impartiality. A spokesperson for the BBC publicly defended her "invaluable reporting", noting her tweets before joining the company were "completely irrelevant" (Revoir, 2020). Nevertheless, at the time Croxford was advised to stay off social media and deprive the trolls of oxygen.

In another example highlighting the need for greater attention on intersectional threats in existing protocols, a reporter from a minority racial background in the UK who wished to remain anonymous, told a researcher that she was threatened by far-right racist Tommy Robinson. He had wrongly identified the reporter as another journalist from a minority background who had reported on him, and he had sent a threat which she said warned that he would "come to my house and wake up my kids". The woman said her news editor had responded to the threat she had received from him by emailing senior leadership, including the managing editor, who contacted Robinson's legal representative to demand that he desist.

On her own initiative, the reporter contacted the police, and was advised by both them and her newsroom to scrub her private details from the internet, the onus being on her to do this. Robinson went on to be convicted of stalking another British journalist and her partner[158] (PA Media, 2021). Speaking with hindsight, the journalist told the researchers that while it was "helpful" that her employer showed her online abuse policy documents at the time, the company should have escalated the case to the police on her behalf, recognising the compound risks

---
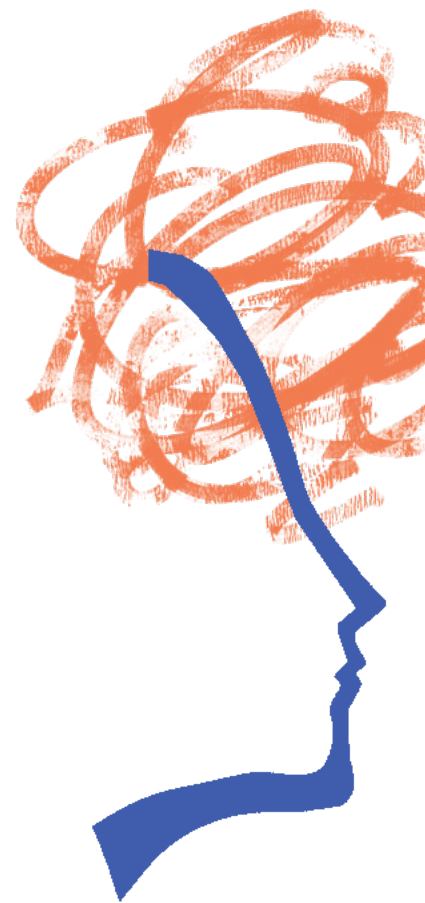
158 See Chapter 2.5 (When online violence spills offline) and Chapter 5.0 (Legal and Normative Frameworks for Combatting Online Violence Against Women Journalists) to learn more about UK reporter Lizzie Dearden.

she faced. The targeted woman suffered additional exposure to psychological injury through intersectional abuse. She accepted counselling offered by her employer at a later stage, unconnected to this event, and only then realised how badly affected she had been by the experience.

Failure to adequately address mental health impacts

The UNESCO-ICFJ survey revealed very low levels of access to practical or psychological support within news organisations for targeted journalists. Globally, only 14% (n=100) of respondents said they knew that their employer had a policy about online violence or had been issued guidelines about how to deal with it. The same small percentage said they were able to contact a digital security expert who could assist them. Only 17% (n=121) knew that they could access legal support and just 11% (n=79) said that their employers provided access to a counselling service, while 20% (n=143) said they had access to a gender-sensitive peer support network. "From all the social media guides I've seen, there's nowhere where it prioritises the mental health of each individual journalist in dealing with social media trolls, or the gendered nature of the harassment," Nigerian documentarian Ruona Meyer said.

ONLY 14% *OF WOMEN JOURNALISTS SURVEYED SAID THEY HAD* ACCESS TO A DIGITAL SECURITY EXPERT WHEN THEY WERE UNDER ATTACK ONLINE

Only 50 survey respondents had been offered any training in their workplace to deal with online violence. Many interviewees and survey respondents either did not find the training that was on offer particularly valuable or gender-sensitive, or they had difficulty accessing it. Gladys Rodríguez Navarro, freelance journalist in Mexico, said when employers do give permission for training, the caveat is that it is on unpaid days off. On the other hand, a British digital safety consultant who asked not to be identified said journalists she trains in the US and UK can be blasé and question "why they need to know this".

Even in those newsrooms where a formal protocol for responding to online violence was in place, this study revealed a lack of understanding of the psychological harm caused by gendered and intersectional online violence, and a corresponding failure to address the mental health consequences for those targeted. Hannah Storm, former International News Safety Institute Director, highlighted the need for an emotional 'flak jacket' to address gendered online violence: "We wouldn't send somebody to war without a physical flak jacket. We shouldn't be sending anybody anywhere where their mental health might be compromised without an emotional flak jacket. That means talking to them about how their work is going to impact them and thinking about resources and risk mitigation in advance" (Tobitt, 2021a).

At the BBC there are around 100 mental health "first aiders" – members of staff who have undergone trauma risk management (TRiM) training, and can offer peer-to-peer support. First aiders are identifiable by their different coloured lanyards, and their names are publicised on the intranet. The aim is to have one first aider in every team, and in every bureau internationally. Like many large news organisations, the BBC also has a well-established employee assistance programme, run by a third party. Through this programme, journalists (and their families) can access professional help of various kinds, including counselling.

Thomson Reuters also has a long-running global trauma programme for journalists. Originally set up 15 years ago to support journalists who worked in war zones, this service was expanded in 2015 with the addition of a peer network, which is independent of management, and supervised by a third party provider of trauma support. In 2021, there were 60 people in the network, offering peer-to-peer support to journalists.

Suspicion of authority, corporations and Human Resources departments make peer support networks an important bridge to professional psychological support for women journalists enduring online violence, according to Michael Christie, General Manager, Global Logistics & Security at Thomson Reuters. Kristen Neubauer, also from Thomson Reuters, described "a macho swagger". That she said it fed a stigma associated with mental health issues. This is one of the industry impediments to more effective responses to the problem.of gender-based online violence against women journalists.

Another kind of newsroom 'swagger' is the type that insists that online violence is not "real", or that it is incomparable to physical violence. Finnish journalist Jessikka Aro has said women journalists are "targeted in cyberwars the same way they are in kinetic wars" (The Economist, 2019). One of the Serbian journalists interviewed - Jovana Gligorijević of *Vreme* - said that the lived experience of women journalists exposed to online violence is frequently diminished by those who covered conflicts in the former Yugoslavia in the 1990s:

> *[W]hen it comes to taking online threats seriously, that will never happen. I talked to [managers] very openly... to let them know that... I don't feel safe and that [online violence] happens. The answer I got comes down to one sentence and that is: 'You know how we were all war reporters, ... so to us all this is like nothing. [It is] nonsense compared to what we went through'.*

In an industry in which the idea that journalists can develop PTSD after being exposed to traumatic events has only recently been acknowledged (Smith et al., 2015), online violence (most often experienced by women journalists) may be dismissed as trivial compared to the "real" violence that hard news reporters (still mainly men) face. However, "real" and "virtual" harms cannot be separated and psychological injury needs to be understood as "real" harm (Lumsden and Morgan, 2017; Global Fund for Women, 2019; Chocarro *et al.*, 2020; Posetti et al., 2020b, 2021). Former CNN Executive Editor, Digital Inga Thordar said: "There has to be, in my view, a total review of how we are handling these cases. Are we understanding that the psychological impact is just as great as you kicking me in the shin?" The fact that online violence is disproportionately experienced by women journalists, and that the most frequently reported impacts are psychological, may help explain the low priority that is often given to tackling the problem by media employers, even when its existence is acknowledged.

41% BELIEVE ONLINE ATTACKS ARE CONNECTED TO **ORCHESTRATED DISINFORMATION CAMPAIGNS**

Recognising and responding to disinformation and political extremism related to online violence

According to the UNESCO-ICFJ survey, 41% (n=293) of female respondents said they had been targeted in online attacks that they believed were connected to orchestrated disinformation campaigns. Such attacks often involve attempts to discredit the professionalism of the journalists targeted, and this damage radiates to news organisations, contributing to the erosion of trust in journalism and facts in general (Posetti 2017b; Posetti 2018a; Posetti, Maynard and Bontcheva, 2021). These are the sorts of challenges that newsrooms on the frontlines of the disinformation war, such as Rappler in the Philippines, The Quint in India, and Daily Maverick in South Africa, have responded effectively to through editorial and strategic advocacy efforts (Posetti, Simon and Shabbir, 2019a; 2019b). Fundamental to their responses has been an understanding about how viral disinformation and political extremism are bound up with gendered online violence.

Julia Carrie Wong's reporting for Guardian US has led to the de-platforming of several white nationalist groups on Facebook. After a year of constant online violence, including rape and death threats, combined with isolation due to COVID-19 lockdown, Wong experienced an escalation of psychological trauma after receiving graphic images associated with one of her investigations. "That also just really triggered some bad mental health consequences for me and a return of the panic attacks. And basically some conversations with my editors that soon I am not going to report on tech anymore." She said *The Guardian* has been very supportive and understanding about her need to take time off.

Julia Carrie Wong ✔ @juliacarriew · Mar 10, 2021 ···
one of the biggest threats to journalists who get targeted with harassments is that they'll lose their jobs bc management gets spooked. I've been very lucky to have a very supportive employer, this has happened many times at other outlets. Publications need to back their writers.

💬 8   ↻ 119   ♡ 1,301   ⬆

Julia Carrie Wong ✔ @juliacarriew · Mar 10, 2021 ···
My main point however is that this is a labor issue. It's a condition of our work. And I will continue to argue and advocate for solidarity among journalists to mitigate the harms and to hold our employers and industry accountable to supporting the journalists affected.

💬 6   ↻ 41   ♡ 401   ⬆

Figure 24: Twitter thread from March 2021.[159]

According to Wong, who was a freelancer prior to joining Guardian US, employment insecurity, weak labour laws, and the absence of universal health care in the US can exacerbate the impacts of gendered online violence, something employers should be aware of: "So much of the way that this stuff works is [about] attacking your sense of safety as far as your livelihood goes, trying to get people fired. You know, losing your job in this country obviously means losing your health care. It can just be catastrophic".

In August 2020, recognising the real damage caused by disinformation campaigns targeting news organisations and those who write and report for them, Guardian Media Group took legal action to shut down a website generating fake *Guardian* headlines and byline profiles which were being shared with the trending Twitter hashtag #TrollingTheGuardian. Carole Cadwalladr was one of the journalists

---

159 Twitter thread by Julia Carrie Wong about her experience of online harassment: https://twitter.com/juliacarriew/status/1369504775545397252

targeted by the now defunct guardianmeme.com website. At the time of her interview for this study in early 2021, Cadwalladr said that her employer fails to adequately recognise, understand or respond to disinformation as a factor in the attacks she endures. However, Gill Phillips, director of editorial legal services for Guardian Media Group, pointed out that when it comes to the disinformation beat or reporting on political extremism and being targeted in an orchestrated far-right attack, there are limits to what a news organisation can practically do, unless a threat can be deemed as "serious" and imminent: "If things look like they're getting someone outside your door physically or the equivalent of that online, then definitely you feel you can move that into the criminal sphere."

At the BBC, while the Infosec (information security) team can flag threats in disinformation reporter Marianna Spring's public Twitter timeline, initially it fell to her to sift through the torrents of abuse and threats that often come via her personal devices and private messaging channels. The inability of the BBC to monitor the direct messages Spring receives on social media platforms due to privacy concerns also hampers their responses. This is a challenge for media employers: some of the worst online violence directed at women journalists comes via closed channels like Facebook Messenger and direct messages on Twitter and Instagram, but legitimate source protection and privacy protocols can prevent routine monitoring of journalists' private messaging spaces by employers. Spring said, however, that it is "problematic" that the onus was on her to deal with this.

While Marianna Spring has been the subject of offline stalking connected to online attacks, she was only recently provided with a BBC phone to allow her to better separate her personal life from her work and help protect her from doxxing. Another BBC employee, who wished to remain anonymous, explained that in her newsroom, she and her colleagues had been increasingly exposed to online violence and doxxing because they were required to use their own devices for socially distanced reporting and social media-based investigations during the pandemic. In a number of instances, she said, women journalists on her team had been harassed on their personal mobile devices by sources they had contacted which also exposed them to a real risk of doxxing.

A freelance journalist, based in Europe, told researchers she personally absorbed the costs associated with replacing her digital devices after being hacked by a suspected foreign State actor, despite being continuously employed by a major news organisation for over 15 years.

Formal online violence protocols are a good start, but no guarantee that journalists will be supported

The expert trainers and consultants interviewed for this chapter observed that all newsrooms should have a written protocol for dealing with online attacks, but the majority did not yet have one in place. The journalists and editors interviewed and surveyed for the broader study echoed this perspective, indicating that many newsrooms do not have any clear protocols, policies, or guidelines in place to tackle online violence against women journalists. Some journalists mentioned
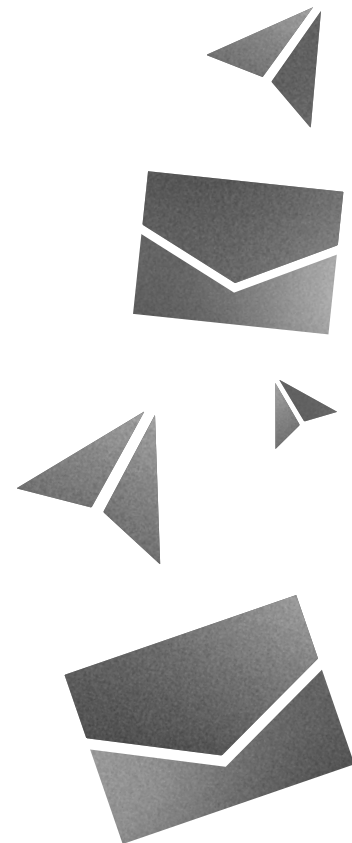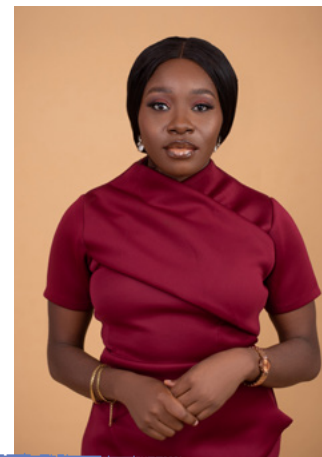
ad hoc institutional support, but most said that the newsrooms they worked for did not see online violence as a workplace safety issue (Jane, 2014). There was little institutional understanding of either the mechanisms or the consequences of the threats, abuse and harassment that women journalists experience. And where the targets did get support, it was most often informal or came from external networks.

A number of larger international news organisations do have formal, written policies, guidelines and procedures in place to deal with online violence, particularly in the US and the UK. But a protocol for responding to online violence does not equate to an effective prevention or protection mechanism. Michelle Stanistreet from the UK's National Union of Journalists (NUJ) said that there is often a "gap between written policies and effective practice". When asked for examples of good practice, the expert trainers and consultants interviewed for this chapter all identified the same few (Western) news organisations.

However, as discussed, several journalists interviewed for this study who worked for these newsrooms said they had not received effective support. They saw gaps in policies such as the absence of an intersectional lens and/or a failure to understand the function of disinformation in connection with such attacks, while others described feeling unsupported, and isolated, or being punished while under attack. Carole Cadwalladr's experiences at *The Observer* reflect this perspective; so too do the experiences of some journalists at the BBC, *The New York Times,* and *The Washington Post,* who also told the researchers (on condition of anonymity) that they had not felt adequately supported by their employers, despite there being well-established protocols and support systems in place.

Kiki Mordi
Picture credit: Adebakin Fifoluwa

While this study has accumulated evidence of policy gaps, leadership failures and problematic protocols in some of the world's most influential news organisations, there are also demonstrations of strong, empathetic and effective leadership at these outlets. When she embarked on a high-risk investigation about the exchange of sexual favours for grades, the BBC took freelance Nigerian journalist Kiki Mordi, a founding member of the Feminist Coalition,[160] through its safety policies, including how to deal with online harassment, to improve her digital security and protect her personal details online. Other employers, she said, had not informed her of any policies regarding online abuse. The online violence that Mordi has experienced, and the lack of clear, accessible legal mechanisms and employer policies to support Nigerian women journalists who are victims of online violence means Mordi now feels that she has to take a deep breath before tackling stories that champion women's rights. "No matter what, I will always stand up and fight for women. But it may be harder...because of what I went through."

Many of these leadership-related parallel weaknesses and strengths are also evident internationally in smaller news organisations, as highlighted below.

---

160 A feminist collective in Nigeria which campaigns on a range of issues.

iii. Newsroom leadership: failings, awakenings, and exemplars

At *Sunday World* in Northern Ireland, management is dealing with extreme risks to reporters, including blanket death threats issued against all their staff. When she was a staff reporter on the paper, Patricia Devlin was offered counselling by her employer and the legal department pursued one paramilitary figure who threatened her. But she recalls that she was still a freelance journalist for the paper in 2020 when the stress and fear associated with attacks on her and threats against her family caused her to take a month off work: "I just said, I can't do this anymore. Work was great that way, they allowed me to take the time off. But that was actually before I was [on] staff. So I didn't get paid for those four weeks. That's hard when you have three children at home and you have to take that time out, and you're not getting paid for it."[161]

According to Michelle Stanistreet of the UK's National Union of Journalists (NUJ), strong newsroom leadership - which publicly practises empathy - is an important feature of a robust response to gendered online violence. "[You need to know] that your employer cares; that your employer takes your safety seriously and actually has measures that they put in place pretty swiftly." Stanistreet also explained that practical assistance should be provided to protect women journalists under attack and facing financial disadvantage:

> *A member of ours in England was forced to flee her house and was told by the police that there was a credible risk of attack against her. She had to leave with her kids. This was during one of the lockdowns. The company paid for the first night she stayed in the hotel and then pretty much she was on her own. … She's a local reporter, so she's seriously out of pocket as well as under enormous stress and strain. And yet the company did the bare minimum really.*

In the male-dominated Pakistani media industry, women journalists say they have to struggle even harder to convince their male counterparts and management that online violence is real, and that it affects their personal and professional lives. According to a 2016 Solidarity Center report on the working conditions of women journalists in Pakistan, 82% of the 214 survey respondents (n=176) received no physical or digital security training or resources, and 75% (n=161) had had no digital safety training. Only 17 respondents said their media house had provided 'stress counselling' to help them cope with negative experiences (Solidarity Center, 2016). In the view of freelance journalist Yusra Jabeen, most women journalists are still told by editors and producers to "go silent" in response to online violence. "They've always said 'this is Pakistan, this is how society is, you have to deal with it, or learn to deal with it'. Such responses are not helpful and leave us feeling even more worn down than before." Sabahat Zakariya, who reports on gender

---

161 Devlin left *Sunday World* in December 2021 due to concern about a lack of management support in the context of ongoing online violence. She is now a freelance crime and investigative journalist.

issues, said complaining about online violence to Pakistani media employers can lead to them positioning women as "nuisances", or "trouble makers", and result in lost job opportunities. These experiences are compounded by the "many cases of sexual harassment and discrimination from male members of journalism unions" that the Coalition For Women in Journalism (CFWIJ) responds to, said founding director Kiran Nazish.

In Lebanon, instead of fighting back against extreme online attacks, however, the journalists' own news organisations are accused of disregarding the problem or denying responsibility for what the journalists experience. The effect is a reinforcement and further perpetuation of stereotypes that re-entrench the women's secondary status in society (Melki and Hitti, 2020). Lebanese interviewees said that news organisations where they worked seldom invested in protecting their journalists' safety. Journalist Youmna Fawaz said that although they may pay for hospitalisation in the event of injuries sustained on the job, media employers are unlikely to pay for any costs incurred by women journalists dealing with online abuse. Luna Safwan, a freelance journalist and journalism safety trainer in Lebanon, observed that: "Male employers would always say you're a journalist, you have to be tougher than that, while female employers are more compassionate and would give me time off." She considers that there are two things missing in many Arab news organisations: clear policies against sexual harassment; and policies for mental health and safety.

Modelling good practice in newsroom leadership in news organisations small and large

In South Africa, Branko Brkic, Daily Maverick's Editor-in-Chief, may provide a model for (mostly male) editors to follow in responding to gendered online violence against their staff. Investigative journalist Pauli Van Wyk, who has received brutal threats of sexual assault and murder, including being "shot in the pussy", credited Brkic's leadership and support for her survival as a journalist. "[He was] incredibly supportive.[162] He really was quite amazing. He shielded me from attacks and he helped me to move around the country as well. Without my editor, I can tell you now, I probably wouldn't be a journalist anymore," she said.

When US disinformation reporter Brandy Zadrozny was targeted in a TV segment presented by Fox News' Tucker Carlson, after first being abused online, she was doxxed. She received hundreds of threatening voicemail messages and emails, including threats radiating to her children. But she said her employer, NBC/MSNBC, was "really good about taking care of me and... rallying around me", and she described their "thoughtfulness" in putting out a statement, which she recognises a lot of companies do not provide: "You have to loudly stand up for your people when these sort of bad faith campaigns happen... I had armed guards outside my home, which is great. My company offered to move me... they were very interested in my physical security". But Zadrozny pointed to what she sees as a missing element in her managers' response: the need to protect her

---

162 Under Brkic's leadership, Daily Maverick commissioned a film about the online violence experienced by its women journalists who are routinely targeted in the course of their work, including Pauli van Wyk and Ferial Haffajee. The film is called Section 16 in reference to the South African constitution, which protects freedom of expression with the exception of hate speech based on gender, race, ethnicity or religion. Completed in April 2022, the film was reviewed by this study's editors prior to its public release. It represents valuable example of editorial and investigative responses to the problem of gender-based online violence.

from ongoing exposure to the threats as they rolled in. She said the onus on her to monitor the threats she receives has to shift: "I needed someone to take my accounts, take my email, take my Twitter. I don't want to have to read through all these and listen to all the voicemails and then say, 'Oh, this is a credible threat'".

Tara John, a digital news reporter with CNN based in the UK, was subjected to severe online abuse and harassment when she reported on transgender law reform efforts in Scotland in 2020. She was shaken at the time, and felt uncomfortable leaving her home, but she told this study that she felt safe to continue reporting on the theme of trans rights because she had two "incredible", "progressive" and supportive women editors to rely on – former CNN Digital Executive Editor, Inga Thordar,[163] and Blathnaid Healey, Senior Director for Europe, the Middle East and Africa - who understand the threats. Tara John has not always felt so well supported in other newsrooms, however. "Honestly, the change in management is just amazing. You know, finally, I feel like I'm functional. They really get it".

Whilst she was Executive Editor, Digital, Inga Thordar said she tried to steer CNN's protocols to ensure her staff under attack are protected as well as possible through legal and safety interventions. "And we have experts deep diving where those threats are credible. And overall, their response to this has been to my satisfaction." Thordar said she currently had more journalists coming to her in the past two years to report online abuse than in the past 15 years. Although the taboo of discussing gendered online violence has been removed and gender-focused support networks exist in many newsrooms, Thordar said: "We still have a long way to go. And even if these processes are being put in place, they have been slow to materialise. There's probably still a lot more work that can be done". In her view, news organisations must tackle digital misogyny as vigilantly as they are responding to disinformation.

At the BBC, Marianna Spring said she is empowered by having an editor who himself has experienced pandemic-related online conspiracy-based abuse. While saying that media employers in general "could do more to be supportive", *The New York Times'* Apoorva Mandavili acknowledged that she was in a comparatively privileged position, having access to structured support. "I've been lucky that we have a security team." But there also needs to be an emphasis on holistic care, she said: "I think it would be very helpful for everyone to have not just that kind of practical help, but also sort of emotional support, moral support and just validation that this is not OK."

In October 2021, US TV reporter Betty Yu was targeted by a conservative YouTube host for her "aggressively Asian face". Her employer, CBS Television and its local San Francisco station KPIX 5 put out a joint statement of their support for their "colleague" and said that they condemned "the horrific, racist comments directed at Betty, as well as the other demeaning Asian stereotypes spread during [the] programme". The statement added: "We stand in solidarity with Betty, an accomplished journalist and valued member of our CBS family. These hateful and offensive remarks are outrageous and destructive and reaffirm the importance of our work as journalists to shine a light on anti-Asian violence and hate speech when it occurs" (Patten, 2021).

---

163 Inga Thordar left CNN in early 2022.

In 2018, as an editor managing a Canadian newsroom with a dozen reporters - mainly people of colour - Joanna Chiu noticed that no matter the topic or shared bylines, trolls would come after the women journalists. The abuse ranged from condescending and ill-informed criticism, to rape and death threats. As national correspondent at *Toronto Star*, she spent a lot of time compiling evidence of attacks on her which she screenshotted for her editor. "The *Toronto Star* was trying to learn more about what's happening," and she came across old death threats that she had forgotten about amidst of torrents of abuse: "They said I should get my neck ready because they were coming over to my house to behead me - horrifying".[164] The police told her that judges were hesitant to allow investigating officers access to identifying IP addresses unless there was an imminent threat to someone's safety (she had received death threats), and that the cyber crime unit was mainly focused on economic crimes like fraud. She wrote a front-page story about how common online hate is, in which the Minister for Public Safety said Canadians who experienced online abuse should go to the police (Chiu, 2021). Chiu described the process as "very circular" (CBC/Radio-Canada, 2021).

Journalists working for Rappler in the Philippines described feeling well-protected and supported by the women in charge of their newsroom, despite a constant bombardment of online violence. Central to this trust was the fact that the newsroom leaders joined them 'in the trenches', fighting back against their online attackers. At a conference, political reporter Pia Ranada told an audience that Rappler's support included defending her in the Philippines Supreme Court (Rappler, 2019), while also fighting back online: "I'm just so proud of my company… It takes a certain kind of boss to stand up against a greater power, when they are already themselves on the receiving end of so many attacks. So thank you Maria Ressa[165] and the rest of Rappler for standing up for us."
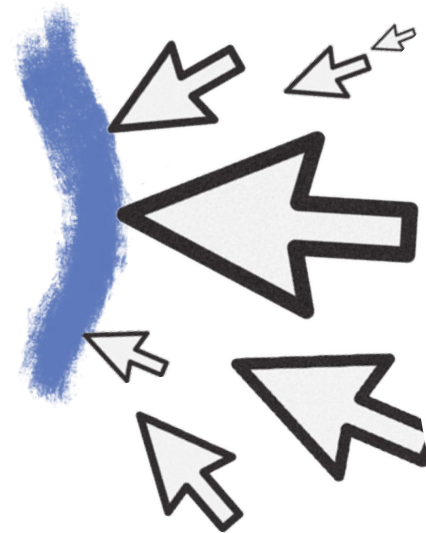
Increasing awareness of the need for greater gender sensitivity

Al Jazeera English's White House Correspondent Kimberly Halkett was sexually assaulted in 2008 while reporting on International Women's Day: "It just about destroyed my life and it almost took away my career… My middle managers blamed me…so I kept trying to navigate [it] with these completely inadequate and ill-equipped middle managers who saw it as a problem and tried to cover it up to protect their own jobs. Because the fact that this had happened to me meant they weren't doing their job."

Halkett was reluctant to escalate the case to senior management because she was afraid of hurting her career. But when her immediate supervisors fired her to cover up the attack and the lack of security that helped to cause it, she did complain and the managers were sacked instead. She said she would be "forever grateful" for the interventions of these senior managers, whom she said

---

164 Apoorva Mandavilli at *The New York Times* told us she had received a similar death threat, that she should have her head cut off in public, via email - and also that she was similarly abused online disproportionately compared to her male colleagues: "the emails are actually worse, because they're more private… that I should have my head cut off in public…I am a liar. I should be ashamed of myself. I don't deserve to live. One said, I hope you get the virus and choke…very nasty and vile emails".
165 Ressa, founder and CEO of Rappler, is a 2021 Nobel Peace Prize laureate and the 2021 UNESCO World Press Freedom Prize winner.

immediately recognised the magnitude of the attack and its effects, and supported her as she rebuilt both her confidence and her career.

When Halkett found herself at the centre of an online violence storm in mid-2020 - which included death threats, doxxing, and the targeting of her teenage daughter, she realised that: "The internet and technology has moved so rapidly that we have a new avenue for violence against women, and no protections in place in the way we do with sexual assault or any of these other crimes against women because of their gender". In this instance, though, her current employer's response was the opposite to what she experienced in 2008. She said additional security was deployed immediately by her bosses - a product of superior management training and policies, she said. They called her at midnight to tell her exactly what they were doing to protect her. But there was still room for improvement, she said: "… they were doing everything they could within the parameters of what's in place, and I just don't think that any company is really fully equipped because it's just such a new frontier. There just aren't the level of protections because the policy just hasn't been formulated yet".

Al Jazeera English Managing Director Giles Trendle acknowledged that there may be a gender blind-spot in terms of responses to online attacks that the company needs to consider:

> *We need to be more gender sensitive. We'll address online violence against our journalists but we don't necessarily specify whether it's male or female. We'll first see it as an attack on an Al Jazeera journalist. An attempt by someone or some people to try to stop us from doing a story that they don't like by one of our journalists who just happens to be one gender or another. But I do think it is critically important to be more gender aware.*

Noting this study's emphasis on the propensity for gendered online violence to lead to offline attacks, Giles Trendle said he now felt compelled to read more research on the theme. "It's a bit of a paradigm shift because it's opened up the door where you completely reconsider things."

When former HuffPost UK reporter Nadine White was harassed online after being targeted by a UK government senior official, her then Editor-in-Chief Jess Brammar, who has held senior roles at both ITN and at the BBC, mounted a strong public defence of White, demonstrating strong leadership and meaningful support. Brammar explained that her first response was a feeling of shock and a protective instinct. There is a conflict, she said, of wanting "to minimise the situation and minimise the impact on the journalist. But equally, you want to take an important ethical, robust stance against what's happening". In defence of White, Brammar tweeted publicly at the senior official and asked for an apology, continuing to follow up weeks later. As deputy editor of Newsnight at the BBC, Brammar experienced a slew of online attacks herself, triggered by a story about a political leader. She said that experience – "48 hours of the most horrendous abuse" – helped shape her reaction as an editor.

The risks of requiring journalists to participate in real time audience engagement

Having tied business model development to new modes of digital newsgathering, audience engagement and content distribution that are partly dependent on journalists' online brands and their direct relationships with audiences, many newsroom leaders are left with a conundrum. A combination of editorial and commercial pressure (Gardiner, 2018) can result in newsrooms trying to 'balance' the benefits of personalised social media engagement for audience development, distribution and amplification, with the rights of journalists to be protected from online violence in the course of their work. The shift from open to more closed social media systems of audience engagement - where the actions of perpetrators are less visible and even harder to combat - has created a whole new set of problems for news organisations trying to respond to gendered online violence.

As well as often being expected by their employers to help develop audiences on social media, individual journalists sometimes feel the need to share personal information online in order to build their own 'brand'. Viktorya Vilk from PEN America said that "journalists sometimes feel pressured to grow a strong following on various social media platforms, including Instagram, and to do that they feel they have to share personal stuff to appear authentic and help engage followers". Although it is not necessarily a binary decision, some women journalists feel that they face a stark choice: either build an online brand for the economic and professional advantages (Finneman *et al.*, 2019), or operate as anonymously as possible in their work to avoid online violence, and suffer the economic consequences (Adams, 2018), along with likely limitations to their career progression. This 'invisibility cloak' approach also has serious implications for press freedom and gender inequality in and through the news media.

The bottom line: Managing digital spaces effectively is both challenging and expensive

While the costs associated with mitigating the threats posed by online violence are no excuse for abrogating the responsibility of news media employers to protect women journalists, they are an ongoing burden for news organisations faced with existential economic challenges worsened by the pandemic (Posetti, Bell and Brown, 2020). In this context, the quickly evolving nature of online attacks, along with the mushrooming and interconnectedness of the platforms on which they take place, can feel overwhelming. This is true even for large organisations, like Thomson Reuters, where Michael Christie is General Manager, Global Logistics & Security. "Like many media organisations, [...] we struggle with the basics of finding out that [harassment is] occurring. We don't have the tools to measure the vitriol coming through direct messaging, and all that kind of stuff. We're by no means where we need to be."

The scenario is similar at Vice Media, where Senior Information Security Engineer Destiny Montague said: "We are constantly playing whac-a-mole". In the UK, *Observer* editor Paul Webster also acknowledged the challenges of trying to address online violence 'at scale': "We will provide support where we can to people who are impacted, although that's probably limited because we're hard pressed and because... it happens across the whole waterfront." Nevertheless, Webster acknowledged that there were potential gaps that needed addressing, starting with supporting writers more: "We need to think about psychological support for people who face online abuse...because that's the level at which it operates." Webster also recognised that journalism "must not be closed down" due to these attacks, proposing a solution of "better mechanisms for monitoring and responding to the disparities". *Observer* investigative reporter Carole Cadwalladr said that in addition to routinely providing psychological support to women journalists under attack, news organisations should also consider a strategic communications plan to counter serious attacks on a journalist.

Collaborative leadership

Beyond the need for internal collaboration between key responders to online violence, it is evident that the scale of the problem, and the increasingly sophisticated methods of the perpetrators, require newsroom leaders should cooperate with their competitors on solutions to gendered online violence. Examples are collaborating on knowledge sharing, developing resources and training to help mitigate online violence against women journalists. Many interviewees underscored this point.

Canadian news industry leaders have begun developing a formal collaborative approach to combating online violence against women journalists in that country, which could serve as a model. The partnership has adopted a three-pronged approach: a) A collaboratively organised industry virtual conference designed to facilitate a cooperative industry response to the online violence crisis held in November 2021 (CBC/Radio-Canada, 2021b), Commissioned research into the scale and nature of the problem in Canadian context (Ipsos 2021), and c) Facilitated small group conversations designed to deepen understanding and extend awareness.

The need to hold social media companies accountable for the gendered online violence crisis

The social media platforms have become enablers or facilitators. ...We need to better understand how we can push back against these big tech companies. ...I do think they have a very, very big role and responsibility in combating online violence and abuse and threats... How much do they see as their responsibility, and how much are they willing to take action?

Giles Trendle, Managing Director, Al-Jazeera English

Many of the interviewees and survey respondents, including newsroom leaders, expressed frustration about the limited, unclear and ineffective routes for reporting online violence incidents to the platforms. The fact that the individual journalists receiving the abuse and threats are required by social media companies to report incidents independently through automated systems remains a major barrier to newsrooms' efforts to manage the problem. While some outlets (e.g., Rappler in the Philippines), actively campaign for reform of social media companies to better address the problem, many news organisations do not rise to challenging these companies to address the crisis at a systemic level.

The symbiotic relationship between news organisations and the platforms - which involves commercial contracts and funding - can make news organisations reluctant to demand action from the platforms. Nevertheless, numerous research participants – both journalists and editors – described using informal back channels to report incidents to the major companies because the official automated reporting mechanisms were not fit for purpose. At *The Guardian*, while journalists are often encouraged to make their own representations to the platforms and follow official reporting procedures, Director of Editorial Legal Services Gill Phillips said that in particular cases the media organisations can "make representations to the big guys [at Facebook and Twitter] about things. If we've got a back door route that we can use, we will use it and deploy it".

Former HuffPost UK editor Jess Brammar used her own contacts at Twitter LA's press office when her journalist Nadine White was doxxed in the middle of the night, but she acknowledged hesitation about using back channels to get action:

> *That's the only time I've ever actually got in contact with them...you don't want to use your journalistic privilege to be like, 'I'm going to the press office of Twitter rather than by means of address that any other person would have'. But in that case, it was because she was in danger and it was to do with the fact that she was a journalist. So I felt like [it] was the right thing to do and they just immediately responded and took it down.*

This example - triggering action via a social media company's press office in preference to relying on standard (and largely ineffective) automated procedures - highlights the problem of opaque, disjointed policies and personnel disconnects within these companies. One of the key challenges for editors and other newsroom responders is finding the right individuals within these companies with the expertise, capability and power to respond to incidents. The research suggests that some newsroom managers do lobby the major platforms for better tools and features, as well as more support (including specialised human contact points for urgent response and escalation), and more capability for responding to complaints in languages other than English. But they are less likely to challenge "the gendered assumptions and worldviews embedded within, and reproduced by, technology" (Salter, 2018), or the constructed nature of algorithms that shape online discourse in which misogyny and abuse thrive (Gillespie, 2010; Noble, 2018; Van Dijck, 2013; Seymour, 2019).

Former CNN Executive Director Inga Thordar said news organisations need to explicitly challenge the platforms on the misogyny that underpins online violence against women journalists:

> *Your bar is always the white man. And that just cannot, should not, be the bar for anything anymore. And I think the online harassment bar that the platforms set themselves has got to be put in different places for men and women. But in order to get that, you have to understand the hill that women have to climb and the differences in the abuse that women get online.*

Several editors and a number of civil society actors promoted the idea of regulation as a solution. However, *Observer* editor Paul Webster highlighted the reluctance on the part of many news outlets to openly call for legislation to force the platforms to take more effective action against online violence. While Webster wants "quicker responses" to complaints and "confidence" that they will be "taken seriously", he also noted the "fine balance" to be struck "between protecting journalism, and in particular the journalists who are subject to abuse, but not curtailing the debate, and not preventing people from having a chance to take part in it".

In 2019, the International News Safety Institute (INSI) announced a year-long project in partnership with Google, Facebook and Twitter to convene meetings of journalists from 40 news organisations[166] to try to work collectively to combat the

---

166 The 40 news organisations that took part in these discussions were all members of INSI, and there was very little representation of the Global South.

online harassment of journalists (INSI, 2019). However, the details of the ongoing discussions have not been made public. While collaborations with the platforms on potential solutions are often suggested as a way forward, several newsroom leaders interviewed for this study expressed a lack of trust in such processes, as well as a sense of resignation that they are unlikely to change anything.

iv. Guidelines for building on current good practice

This section describes elements of effective responses to online violence identified within the overarching data corpus of nearly 1,100 interviews and survey responses associated with this study. The insights gleaned were bolstered by the 19 in-depth interviews conducted specifically for this chapter with news organisation staff responsible for responding to online violence at their outlets, and a sample of trainers and experts who consult to international newsrooms on digital safety, and online violence. What follows is a distillation of their combined insights and advice, together with examples of good practice - from newsrooms large and small, from print to broadcast and digital-born news media - to illustrate practical implementation. The aim is to provide a roadmap for outlets looking to take more effective steps to counter online violence.

The time to be prescriptive is now

A number of larger news organisations referenced here do have formal, written policies and procedures in place to combat online violence, particularly in the US and the UK. Jason Reich at *The New York Times,* said he is now comfortable being prescriptive about the need for every news organisation to have a formal response mechanism to deal with online violence: "At this point, [an effective protocol] should be common practice everywhere... Every institution needs to have a transparent, reactive and quick process in place for responding to the harassment of its staff. It's not an easy problem to solve but it's long past the time to start trying." It should be noted, however, that the nature of digital threats changes quickly, and so any measures adopted in response to online violence need to be regularly reviewed and updated.

Some smaller news organisations studied, such as Rappler and Vera Files (Philippines), the Center for Investigative Reporting (Sri Lanka) and the Premium Times' Centre for Investigative Journalism (Nigeria) did not have a written protocol by mid-2021, but nonetheless had effective practices to counter gendered online violence. Rappler, for example, adopted pioneering strategies to combat the problem when the women-led digital news outlet came under concerted attack in 2016. Under the leadership of then Editor-in-Chief and CEO Maria Ressa,[167] they introduced the following holistic measures (Posetti, 2017b; Posetti *et al*., 2019b):

1. Escalating digital and physical security in tandem.

2. Providing tailored psychological support for all staff coming under attack.

---

167 Ressa stepped back as Rappler Editor-in-Chief in 2021 but she continues as CEO.
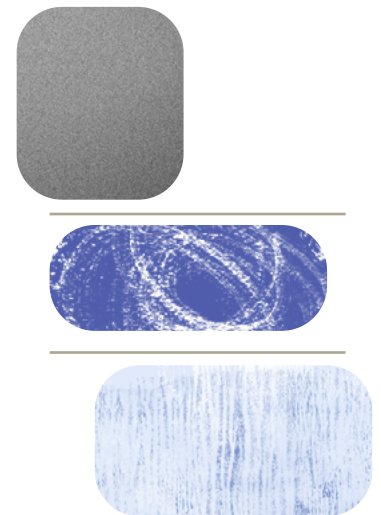
3. Deploying advanced digital investigative journalism techniques to expose orchestrated online violence campaigns.

4. Allowing targeted reporters to take a break from beats that attract intense abuse, or pairing them with a reporting partner, or sending them on international fellowships for their own safety.

5. Harnessing loyal audiences and deploying 'movement journalism' techniques to mobilise community support online.

6. Holding the platforms and State actors to account for their facilitation and enablement of online violence.

7. Speaking up and raising awareness rather than staying mute and avoiding exposure.

Developing an effective formal response protocol

So, what should a robust online violence response protocol for news organisations look like in 2022? Firstly, while incorporating gender-sensitive and digital-security aware guidance and procedures for comment moderation on the outlet's own website/s, it would recognise that social media platforms and apps are now the primary sites of online violence against women journalists, in particular orchestrated and cross-platform attacks. It would also acknowledge that some of the most severe or high risk attacks involve perpetrators associated with State and political actors, powerful individuals and networks operating at the nexus of disinformation.

Secondly, it would focus on preventive measures, such as threat assessment procedures tuned to identify 'lightning rod' beats and stories (e.g., those focused on gender issues, politics/elections, human rights issues, disinformation and investigative reporting) which this study has highlighted. Such procedures would also seek to predict the type of attacker, and mode of attack potentially triggered by particular beats and stories - especially important when State actors or far-right extremist networks are implicated. Risks at the intersection of misogyny and other forms of discrimination - such as racism, religious bigotry and homophobia - should be factored in to threat assessment matrices too.

Also on the preventive side, this ideal protocol would seek to embed holistic training for all staff (tailored according to their roles and levels of exposure), along with provision of base level digital security support which incorporates online violence defence tactics, such as anti-doxxing procedures, at the point of a journalist's 'onboarding', and periodically thereafter.

Then, the protocol would include a number of informal and formal response measures to be enacted in the case of attacks, to support the journalist while she is at the core of a storm. Making sure the woman's digital security is tight when she is under attack is a vital first step. Many of the journalists interviewed for this study also said they wanted a newsroom "point person" whose job it would be to take over their email and social media accounts in the event of a major online violence attack. They need their accounts monitored to minimise further exposure to abuse and threats, preferably by someone who can also assess risk, catalogue threats and screenshot abuse in case of future legal action. One recommendation

comes from a report with guidelines for West African newsrooms, imagining the role of an "Online Safety Coordinator" who could "regularly attend editorial meetings to become aware of upcoming content that may trigger online abuse", and liaise between management and legal teams (Sarpong, 2021). As described earlier, Reach PLC in the UK has recently appointed a Online Safety Editor.

Those newsrooms that have recognised gendered online violence as a workplace issue, and have taken their responsibilities to their women journalists seriously, have ended up introducing remarkably similar protocols, regardless of their size or location. Analysis of a range of international approaches has allowed identification of the following six key features of a gender-sensitive online violence response protocol for newsrooms.

1. Awareness of leadership on, and communications about, the problem

The first step in any protocol is to acknowledge the existence of online violence, and the harm it can cause in order to create an environment where employees feel supported enough to report their experiences. Some larger organisations survey staff as part of this process, such as *The Guardian* (Gardiner, 2016), the *Toronto Star* (Campion-Smith, 2020) and the UK's biggest commercial news publisher Reach PLC (Tobitt, 2021b). Acknowledging the intersectional nature of much digital violence is also crucial, because those targeted are often already marginalised and therefore may have legitimate concerns about reporting attacks.

Once the nature of the threat is understood, policies and protocols must be developed and communicated to all staff and freelancers. They also must be championed by top editorial leadership. An example is Oluwatosin Alagbe of the Premium Times Centre for Investigative Journalism in Nigeria, who said she regularly discusses the threat of online violence in the newsroom, encourages journalists to report it to her, and organises regular training sessions as part of a holistic approach to journalists' physical, psychosocial and digital security.

Strong leadership can also help bridge gender divides in news organisations to avoid gendered online violence being relegated because it is a 'women's issue' - a position which, apart from being sexist, underestimates potential male solidarity and also fails to recognise the chilling effects on audience development (and therefore business models), editorial diversity (the absence of which affects trust) and press freedom. A number of the male editors interviewed expressed shock and alarm at the scale and ferocity of online attacks described by their female colleagues who participated in the research for this study, demonstrating the need for cross-gender recognition of the impacts of online violence on women journalists. Editorial leadership-facilitated conversations that empower women journalists - particularly those from additionally marginalised groups - to speak about their experiences and needs, while requiring male journalists to actively participate, could be a valuable practice. Menaka Indrakumar, former parliamentary correspondent for the *Daily News* and *Sunday Observer* newspapers[168] in Sri Lanka, agreed discussions on the issue of online violence should involve men and

---

168 Menaka Indrakumar is now a freelance journalist.

women, so that: "Irrespective of their social standing, their sexual orientation, their...position or whatever, they should come together and come to a common... emotional understanding [of the problem]".

2. A prevention strategy

An effective protocol focuses on prevention. At minimum, all journalists should be required to undertake digital security training aimed at protecting their devices, communications and personal data, all of which could potentially be compromised as part of an online attack. Digital security experts agree that this training should include the secure use of passwords, two-step authentication, data encryption, general protection of personal data, protection from malware and the use of secure (end-to-end encrypted) email. It should also include advice about scrubbing personal/identifying data from the internet, and subscriptions to services that enable this process (e.g., DeleteMe in the US).

Digital security needs to be led at the organisational level – it is difficult for an individual journalist to make the necessary behavioural and technological changes without institutional support (Posetti, 2017a). Such support should also be extended to freelancers whose journalism for the outlet places them at risk. Also, a protocol that makes the individual journalist responsible for digital security may unintentionally shift the focus to women under attack, or as the ones in the chain primarily responsible for dealing with the problem, thereby normalising gendered online violence. Organisational support should also include setting time aside for training, keeping encryption software up to date, and ensuring that journalists do not have to use their personal phones and laptops for work.

Vice Media Group (USA) has established a formal training programme for its staff and contracted freelancers, including an e-learning programme and in-person small-group training. The in-person sessions target high-risk groups, including women and racial minorities. The InfoSec team undertakes a personalised digital security risk assessment for their most exposed journalists, checking and "cleaning" their digital footprint. The Vice training encourages participants to do their own gender-aware "threat modelling". This is similar to the 'self-doxxing' workshop developed by *The New York Times,* in which staff discover for themselves how much of their personal information is available online, and are shown how to clean up their digital footprint (NYT Open, 2020).[169] *New York Times'* staff have also described hosting 'lox[170] and doxx' parties with colleagues from other outlets during which they use this guide to collaboratively reduce their exposure to doxxing.

Gender-aware digital security should also form part of a risk assessment protocol: rather than only considering a journalist's physical safety before sending them out on an assignment, organisations should also do a digital risk assessment, and consider addressing the need for an "emotional flak jacket" (Tobitt, 2021a). In other words, responses must integrate physical and digital security measures, along with psychological support, and editorial and workflow strategies. Training offered as a preventive measure must be holistic to reflect these different but interrelated elements. Unless the defensive strategies are joined up, they will not succeed.

---

169 Consumer Reports' online guide. "Keep your data secure with a personalised plan" is also useful: https://securityplanner.consumerreports.org/
170 Yiddish term for smoked salmon commonly used in the US.

As this study demonstrates, there is an increasing risk of online violence spilling offline in the form of physical attacks and harassment. The trend is also clear in reverse: offline abuse can trigger online pile-ons, creating a vicious circle. So, it is now more important than ever to ensure responses to gendered online violence are holistic from the outset, and this integrated approach is understood and championed by senior newsroom leadership. Heightened risks at the intersection of misogyny and other forms of discrimination should be factored in to threat assessment matrices too.

At Rappler (the Philippines), the likelihood of a journalist being targeted online is considered before the publication of every story. If it is deemed to be "sensitive", the social media team is alerted; care is taken over the time of publication, to ensure that staff are available to deal with any hostile response, and headlines are reviewed to ensure they do not invite pile-ons. Allies within Rappler's grassroots community networks (Posetti, Simon and Shabbir, 2019b) are also alerted to help support and defend the journalists if and when they do come under attack.

Prevention also encompasses editorial approaches such as comment moderation and clear community standards. Despite a widespread flight from hosting on-site comment platforms (Reagle, 2015; Ellis, 2015; WAN-IFRA, 2016), many media organisations do still offer such facilities. If poorly managed, these can be a source of online violence. The Coral Project, originally an initiative of the Mozilla Foundation, *The New York Times* and *The Washington Post* and now part of Vox Media (US), designed open source software to enable newsrooms to host more civil discussions. This software is translated into several languages and is used by nearly 200 newsrooms in 21 countries.

Key to the Coral approach is the notion that online violence is not just a technological problem, but also "a strategy and culture problem that technology can help or hinder", according to Head of Coral Andrew Losowsky. But the role of unmediated social media platforms as vectors of online violence makes hosting civil discussion within those communities extremely difficult.

Online violence is also a design and business model problem in the social media realm. The inability to pre-moderate comments on news organisations' social media assets (e.g., Facebook), for example, remains a substantial impediment for news outlets trying to improve the standard of conversation through community guidelines. It also poses significant legal risks (Karp, 2021). As a result, the Canadian Broadcasting Corporation (CBC) decided to keep comments closed on most of its Facebook posts and within its Facebook groups to help develop "safer online spaces" (Fenlon, 2021). Following this shift, the CBC said readers' comments on their website surged.

Effective comment moderation on news outlets' websites can help reduce journalists' exposure to gendered online violence in certain contexts. Some of the smaller newsrooms represented in this study, as well as some larger newsrooms with relatively small numbers of comments (e.g., Sveriges Radio), took active steps to engage with the comments on their websites to help keep conversations constructive. At larger outlets, like *The Guardian*, a very high volume of comments makes such engagement impossible – instead, they limit the number of articles open to comments. Comments below "sensitive" articles are either turned off,

or are pre-moderated, and the community standards have been re-written to make them more robust.[171] *The New York Times* pre-moderates all the comments on its site to ensure no published comments are overly threatening or abusive. Other outlets, like Daily Maverick in South Africa and De Correspondent in the Netherlands, have adopted a membership model, which limits commenters to those paying to join the community on their website - a measure that can help curtail abuse. Again, though, these practices and standards cannot simply be transposed onto social media platforms, and these are the sites of much of the most brutal and dangerous online attacks against women journalists.

3. Clear reporting lines, and a means of tracing and tracking cases

Targeted journalists need to know who to tell, and how to describe what they are enduring. They also need to feel confident that they will not be penalised if they do tell – and the people within the organisation to whom they report need to have the necessary training to respond effectively, in a gender-sensitive manner. Under-reporting will occur if any of these needs is not met.

In some newsrooms – particularly those where physical sexual harassment is a significant problem – there are strong disincentives to report any kind of gendered violence, as discussed above. In some contexts, women may be disrespected and shamed if they are seen to be targeted sexually, or subjected to sexualised misrepresentation, discouraging reporting and public discussion of their experiences. Additionally, as mentioned, there is sometimes a stigma attached to mental illness, making it difficult to admit to being harmed by such attacks. This makes setting up clear reporting lines in newsrooms both more difficult, and more important.

Following a report of sexual harassment in the workplace, Tayyeb Afridi, managing editor of the Tribal News Network in Pakistan, introduced a staff training programme and set up an 'inquiry committee' to encourage reporting and to respond to incidents that did occur. Afridi said that the increased understanding of workplace sexual harassment and its impacts led in turn to more awareness of digital violence. Digital security training is now also given to staff, and the committee devised a protocol to deal with online violence as well as workplace harassment.

At larger organisations in less conservative contexts, under-reporting may be a result of a lack of clear reporting lines as well as a fear of being perceived as overly sensitive - a product of sexist attitudes that peg women journalists as 'weak' because they may express emotion. Some have tackled this by encouraging all staff to report online violence that they experience, or that they see others experiencing – even if they themselves do not feel threatened by it. *The New York Times*' Jason Reich said threat assessments are best conducted by people with expertise, not the receiver of the abuse: "Everyone has a different threshold… we've got reporters that get death threats every day, and they're like, 'whatever'… But [we tell them] that's for us to determine, not for you".

---

171 These can be seen at: https://www.theguardian.com/community-standards

In a large organisation, it is easy for reports to get lost. Reich advocates a "single contact culture" – rather than put the responsibility on the individual to decide if what they are experiencing is a digital security problem, a physical security problem, or an HR problem, and all reports are made to a single email address. This simplifies the process, and lowers the bar to reporting, Reich said. Importantly, freelance journalists also have access to this email. The individual being targeted receives a response that is both rapid and human – an email or call from a member of the team who reassures them that the matter is in hand, and offers any immediate support that may be needed.

The New York Times, as a large organisation, has a team that responds to these reports – members include a senior editorial member of staff who provides context and editorial judgement, an expert in physical and digital security, a legal expert to evaluate whether the online behaviour meets the threshold for legal action, and someone from human resources, to offer appropriate support for the targeted individual. In smaller organisations, the primary responder may be an individual, but if so they may still need access to legal and technical support for the victim. It is important that the same individual or team responds, so that responses can be monitored and evaluated over time. In this way an organisation can develop deep, contextual expertise. Additionally, the response team should have no responsibility for disciplinary matters.

Once online violence is reported, it should also be traced (to the source, if possible), and the case should be monitored. These steps are important to the pursuit of legal remedies, should they become necessary. But they also allow news organisations to learn from incidents in order to predict the types of attacker and mode of attack triggered by particular beats and stories to help prevent their recurrence and to improve their responses as the patterns and methods of perpetrators evolve. Further, tracing and tracking perpetrators can aid investigative reporting and research as a response to the problem.

There have been various attempts to track cases on a national or international basis, though these have mainly been under-resourced and/or time-limited, and often not disaggregated for gender. The International Press Institute (IPI) hosted one such database in partnership with Jigsaw[172] in 2016, enabling the reporting and tracking of online harassment of a number of journalists in Turkey and Austria over a period of time. This was the year of a failed coup attempt in Turkey, and a presidential revote in Austria, and journalists in both countries were under intense attack online. This work forms one of seven case studies in the report 'State Sponsored Trolling: How Governments Are Deploying Disinformation as Part of Broader Digital Harassment Campaigns' (Nyst and Monaco, 2018).[173] There are clear potential benefits of large-scale tracking, if maintained over a period of time, Javier Luque from IPI observed:

"What we learned was that it would be better to track individual cases [...] analysing the way the attacks [are] disseminated on various platforms and propaganda sites, using content analysis, social network analysis, and so on – this would be a better way of identifying both trends and perpetrators. A database of this sort,

172 A unit within Google which the company describes as building solutions technology for social problems such as disinformation and toxicity, defined as "rude, disrespectful, or unreasonable language that is likely to make someone leave a discussion". Perspective API was launched in 2017.
173 See also: https://www.bloomberg.com/features/2018-government-sponsored-cyber-militia-cookbook/

properly funded and which newsrooms around the world could feed cases into, would be invaluable."

A more sophisticated version of this approach was adopted for the big data case studies focused on journalists Maria Ressa (the Philippines) and Carole Cadwalladr (UK) (Posetti et al., 2021).[174] In those cases, network analysis was combined with Natural Language Processing (NLP) techniques to describe patterns, methods, and types of attack within 2.5 million social media posts before being synthesised with qualitative research to create timelines and contextualise escalation pathways. Case studies of this kind could ultimately help inform newsroom protocols for referring cases to the police and to the platforms through partially automated identification of threat signals, for example (ICFJ, 2021).[175]

4. Risk assessment guidelines

A prevention protocol will include a gender-sensitive assessment of the risk of digital violence before a story is published. But what about when an attack does occur? A trustworthy assessment of risk not only helps protect journalists from potential physical harm, it can also minimise psychological harm. When the journalist knows that steps have already been taken to mitigate risks of physical violence, and that the risk of data breaches have been minimised through defensive digital security tactics, then any subsequent attack that does take place may have reduced psychological impact.

Risk assessment guidelines must be based on the best available knowledge, and should be regularly reviewed, but they need to evaluate the physical, psychological and digital security aspects, as well as the professional risks to the journalist who has been targeted. The guidelines will be dependent on local contextual variables and on the ever-evolving nature of online attacks.

To a list of potential threats could be added State actors, political leaders, conspiracy networks, disinformation purveyors, anti-vaxxers, anti-lockdown protesters, online networks of misogynists and racists, and so-called 'stan armies'. In certain countries, security forces, law enforcement officers, local government officials and dark PR operatives would rise up the list, as would the threat of offline violence connected to online attacks.

Various threat modelling tools may be used,[176] but a robust risk assessment is alert to the context in which the journalists work, as well as to the journalists' gender, ethnicity and other risk factors such as religion and sexual orientation. Additionally, monitoring threats in real-time is important to ensure risk assessments respond to changing threat environments. This could involve collaborations with experts - in academia and civil society - to apply monitoring tools that can function as early warning systems for offline violence. Such a system is currently being built by ICFJ and the University of Sheffield as part of a research commission from the UK Government.

---

174 See snapshots from these big data case studies in an earlier extract from The Chilling https://en.unesco.org/sites/default/files/the-chilling.pdf
175 This is the basis of a major new study commissioned by the UK's Foreign, Commonwealth and Development Office which is being undertaken by ICFJ in partnership with University of Sheffield computer scientists and linguists.
176 Jason Reich uses the FBI's strategies as outlined in "Making Prevention a Reality: Identifying, Assessing, and Managing the Threat of Targeted Attack": https://www.fbi.gov/file-repository/making-prevention-a-reality.pdf/view

5. A holistic and adaptable recovery plan

An effective protocol would include a number of informal and formal response measures to be enacted in the case of attacks. Making sure the woman's digital security is tight when she is under attack is a vital first step. This should apply to both her professional and personal accounts and devices. The journalist will likely need support to lock down her accounts (i.e. restrict privacy settings and filter replies) to deflect further attacks or monitor ongoing threats, to relieve her of those burdens while she is under attack. Sveriges Radio's development editors Robert Jakobsson and Niklas Malmberg call this process 'storm management': "We get an overview of the [situation]: what's the storm about; where is the storm; how big is it; how big can it get?… We always remove [the targeted individual] from dealing with it on his or her own…the security department [takes charge]".

The individual "point person" (e.g., responsible managing editor) suggested by many of the journalists interviewed for this study should contact social media platforms on the journalist's behalf to escalate reports and aid efforts to trace users behind the most serious or prolonged attacks. This approach, rather than locking or temporarily deleting accounts, enables real-time monitoring of threats of physical or sexual violence to determine the offline risk and trigger appropriate physical security measures. It can also aid investigations when cases are escalated to law enforcement agencies. So, a robust protocol should also include such a process.

Physical security and temporary relocation may need to be considered in serious cases of online harassment involving credible threats of physical or sexual violence. A robust protocol will also recognise the impacts on the woman's personal life. Recovery plans should therefore factor in the additional stress and trauma created and consider ways to support affected family members.

At least seven journalists in six countries reported to researchers attached to this study that they had been made to either pay for emergency hotel accommodation after a single night, or absorb the costs for children travelling with them. The costs of these measures should be absorbed by the news organisation, with the assistance of NGOs or state mechanisms where justified.

Mitigating solutions as a response include: changes to working conditions or duties for a woman journalist who has come under extreme or long-term attack online. The opportunity to take time off work can at least be a chance to withdraw from social media temporarily. It might also be helpful to offer employees a temporary break from a particular beat or role, or to work with a partner on assignments. Opportunities to recover away from the newsroom can allow for professional development (e.g., on a fellowship or on a novel assignment with a long deadline) but the targeted journalist must be fully involved in these decisions, which should be clearly communicated as temporary (unless the journalist requires otherwise) to avoid perceptions that they are being sidelined because of the "drama", or removed from duties because they are judged to be "weak".

Peer support is essential. This should include opportunities for supportive conversations and mentoring in gender-sensitive environments. In some newsrooms this involves formal networks, but in smaller newsrooms, peer support may be

offered informally. Informal networks formed between targeted women from different outlets and in different countries were also highlighted as valuable interventions by many women journalists interviewed. One example recommended by Real Amazônia reporter Kátia Brasil was the Zero Hora[177] unit providing specific support for journalists experiencing violence. Psychological injuries can develop over time, therefore newsroom leaders should remain appropriately attentive regarding the mental health and well-being of women recovering from online violence episodes, especially when these have been cumulatively damaging. Long-range mental health support should be made available for this reason.

Legal support should be provided wherever it is requested by women under attack. And newsrooms should not hesitate to call in law enforcement officers to protect journalists and prosecute perpetrators where necessary (with the consent of the targeted journalist). In several cases uncovered in this study, women journalists were required to independently seek access to (often costly) legal support, and initiate police complaints. Both of these processes can be exhausting and potentially re-traumatising, compounding the mental health impacts. However, it should also be noted that some interviewees expressed their reluctance to pursue legal avenues because they did not trust the police or the courts in their countries. Vulnerable internet users (such as marginalised women or racial minorities) may be particularly reluctant to call the police.

Reporting cases of online threats, harassment and abuse against women journalists to the police and the platforms, or initiating defensive legal action, should not be left to the journalist herself but spearheaded by her employer. The news organisations' contact point should then follow up on these reports to track progress until there is a resolution. In the case of the platforms, reporting should not be limited to automated on-platform reporting mechanisms. Newsroom leaders should insist on the social media companies providing a human contact point for the rapid escalation of high risk, high volume or long-range attacks as part of any commercial arrangement connected to the use of their tools or the production of content, and not hesitate to use these conduits. Although involving police or the judiciary can be a vital component of a holistic response, newsrooms should be sensitive to intersectionality and local realities.

6. An editorial and advocacy strategy for newsrooms dealing with online violence

Rappler in the Philippines is perhaps the best known example of how powerful a hybrid editorial and advocacy response can be as a function of an online violence protocol. They approached the online attacks they experienced as they would any other important story: they investigated them (Posetti, 2017b; Posetti, Simon and Shabbir, 2019b; Posetti, Maynard and Bontcheva, 2021). Rappler journalists were able to show that the attacks they experienced were systematic and organised, and published evidence that they were State-linked. As this example shows, the

---

177   Brazilian newspaper based in the city of Porto Alegre, the sixth largest in the country.

core skills that should be found in every newsroom – the ability to investigate and expose wrongdoing, to explain complex ideas to the public, and to report in the public interest – can become powerful tools in the fight against gendered online violence, if newsrooms are willing to use them.

Rappler's Maria Ressa (CEO), Glenda Gloria (Editor-in-Chief) and Chay Hofileña (Managing Editor) also highlighted the importance of public advocacy about the issue in national and international fora to underscore their refusal to be silenced by online violence. In 2016, Rappler launched a campaign called #NoPlaceForHate which focused on countering social media toxicity. The outlet's community engagement team published stories about what was happening online, held webinars and organised in the community. Their message was "this is how social media is being misused and abused, and you have to know that this is being done, and you have to fight back!" Hofileña said. However, Maria Ressa later declared that a "naive" mission because it did not account fully for the viral nature of online abuse and hate speech, or the platforms' unwillingness or inability to effectively counter it (Posetti, Simon and Shabbir, 2019b; Posetti, 2020b). However, since then, Ressa and Rappler have prioritised investigative journalism responses, public advocacy focused on the need for accountability from the platforms, and media literacy campaigns.

The Guardian and *Observer* newspapers also pioneered editorial leadership efforts to address online abuse by publishing an analysis of 70 million comments on their own website, together with a series of articles under the banner, 'The Web We Want' (Gardiner, 2016). They also routinely publish editorials and statements advocating for their journalists when they have come under attack. More recently, they published a call to action triggered by the preliminary findings of this study, advocating action against gendered online violence: "The chilling effect of mob censorship cannot be overestimated and must be urgently addressed, with a view to stopping and rolling back its baleful spread" (*The Guardian*, 2021).

The Toronto Star's Public Editor, Bruce Campion-Smith adopted a similar approach, conducting a survey of staff in 2020 (Campion Smith, 2020) and writing an editorial in 2021 in response to this study's earlier outputs titled "Online abuse of women journalists is a crisis we can no longer ignore" (Campion-Smith, 2021). Similarly, the Director General of Sveriges Radio (Sweden) Cilla Benkö wrote a clarion editorial for International Women's Day in 2021: "Women journalists are subjected to hatred and threats online to a much greater extent than men, and above all, this is coming from a sexualised and threatening climate on digital platforms. If the situation does not improve, journalism and the community at large run the risk of losing important voices and perspectives" (Benkö, 2021).

v. The role of intergovernmental and civil society organisations in facilitating newsroom responses

Many interviewees and survey respondents mentioned the valuable role played by intergovernmental and civil society organisations in funding, developing and facilitating research, training, advocacy campaigns, and resources to help news

organisations combat gender-based online violence. Prominent among them are trade unions and professional bodies, press freedom and media development organisations, women-centred media support organisations, international and regional level intergovernmental organisations such as UNESCO, UN Women, and the OSCE.

This research surfaced dozens of initiatives designed to respond to online violence against women journalists. However, one problem identified in the course of the research is the proliferation of outdated and dormant resources and tools developed to respond to online harassment, such as phone-based support services which are no longer operational (Gober and Posetti, 2020). This makes the job of surfacing and accessing reliable, trustworthy and current resources for responding to online violence when a journalist comes under attack even more challenging for news organisations. This research-identified gap led to the launch of the Online Violence Response Hub in 2021, as a partnership between the International Women's Media Foundation (IWMF) and the International Center for Journalists (ICFJ). The Hub is a project of the Coalition Against Online Violence, a network which aims to connect journalists, experts and civil society organisations working collaboratively to respond to the problem as it evolves.

Several other initiatives are worth noting for the benefit of news organisations seeking access to knowledge, resources and advice. Reporters Without Borders (RSF), the Committee to Protect Journalists (CPJ), PEN America and ARTICLE 19 provide training in a number of countries, and several journalists interviewed said they had benefited from these courses. CPJ has also produced a range of resources,[178] and PEN America has published a "Field Manual", offering guidance for navigating online abuse, which is available in multiple languages.

The International Press Institute (IPI) published a report in 2019 identifying good newsroom practices (Trionfi and Luque, 2019), which was accompanied by a platform[179] hosting resources for news organisations and detailed guidelines for how to set up a gender-aware online harassment protocol (IPI, 2017). The IPI protocol guidelines were credited by several interviewees in the US, Nigeria, Serbia, Sri Lanka and Brazil as having provided the blueprint for their own protocols. The International Association of Women in Radio and Television (IAWRT) has also produced resources for women journalists.[180] The Dart Centre for Journalism and Trauma has developed a trauma risk management (TRiM) approach to gendered online violence that has been adopted by some larger newsrooms; they have also made factsheets and other resources available online.[181]

vi. Conclusion

Firstly, it needs to be acknowledged that news organisations have as much responsibility to ensure the safety of women journalists online - including those who work in a freelance capacity - as they do offline. This means working to

---

178 These include blog posts with guidance on self defence strategies online (CPJ, 2019b), and safety notes focused on digital security in reference to harassment (CPJ, 2019a).
179 https://ipi.media/programmes/ontheline/
180 These include the video, How to Counter Gender-based Online Harassment of Journalists (IAWRT, n.d.b) and the 2017 handbook; What if ...?: safety handbook for women journalists : practical advice based on personal experiences (Saady, 2017).
181 Dart publishes some free online resources: https://dartcenter.org/resources?page=1&featured%5B0%5D=750

both mitigate attacks and respond sensitively to the personal and professional impacts of gendered online violence. Recognising the interplay of physical, digital, psychological, professional, personal, and press freedom-related as well as editorial risks and impacts is key. This is particularly important where misogynistic attacks intersect with other forms of discrimination, and related threats such as viral disinformation, hate speech and far-right networks.

Secondly, news organisations seeking to respond more effectively to gendered online violence need to recognise the widespread and entrenched structural sexism and misogyny that enables sexual harassment of women journalists by colleagues and superiors, as highlighted by the #MeToo movement, and documented by researchers over decades (Global Media Monitoring Project, 2020). As one woman journalist under fire online put it: "Even the most open-minded media organisations are still run by men who don't fundamentally understand the misogynistic nature of these attacks" (Klein, 2021). Beyond the need for awareness, there is a clear requirement for employers to provide practical, holistic, research-informed training to both help prevent attacks (e.g., enhancing individual digital security and safety), and improving the capacity of individual journalists to respond to gender-based online violence when they are targeted.

Thirdly, it is important to note that news organisations are not homogeneous - they range in size, focus, culture, capability, independence, and viability.[182] Many are also subject to varying degrees of press freedom violations and institutional capture. Additionally, newsrooms have become more 'dispersed' and technology-dependent in the context of the pandemic. These factors mean that there is no 'one size fits all' approach for newsrooms seeking to respond to gendered online violence around the world, and any policies or protocols developed to help manage the problem need to be localisable, adaptable to deal with emerging challenges, and regularly reviewed.

News media employers, often already struggling with issues connected to the lack of diversity among staff, especially in senior ranks, and the associated lack of diversity in the representation of communities through their journalism, need to be mindful of these impacts (Kassova, 2020; Cherubini et al., 2021). It is important that they find other, more effective ways of protecting women journalists under attack - especially those made more vulnerable due to intersectional risks such as racism - to ensure that their journalism can be seen and heard.

Finally, it must be acknowledged that news organisations can only do so much in an environment where exposure to organised online attacks and audience toxicity primarily comes via the social media platforms, not their own websites, and perpetrators include powerful political forces. Yet, while wealthy technology companies are the main vectors of gendered online violence, and State-actors and political leaders stand accused of instigating and fuelling some of the worst attacks on women journalists, this is not a justification for news media employers to sit still, stay silent, or fail to listen to their female staff.

Because of the chilling effect on women journalists' practice and the increasing involvement of State actors and political forces in coordinated attacks, there are also implications for press freedom and the public's right to know. The compounding problem of pandemic-era conspiracy networks that increasingly

182 https://unesdoc.unesco.org/ark:/48223/pf0000379826eng

meld with disinformation and political extremism online, also heightens and extends the risks facing women journalists. All of these factors should cause news organisations to redouble their efforts to combat online violence against women journalists, and significantly upgrade their response protocols. They have a responsibility to do everything they can to prevent these attacks, protect and defend targets when they are under attack, and in the aftermath; and report and ensure that perpetrators face consequences from the state and internet companies where appropriate.

5

Platforms and vectors: Assessing internet company responses to online violence

Julie Posetti, Kalina Bontcheva and Nabeelah Shabbir

The role of internet communications companies in online attacks against women journalists cannot be underestimated. They operate in an era of digital journalism, networked disinformation, online conspiracy communities, and political actors weaponising social media and misogyny as tools to attack women journalists. Their claim that they are simply operating as passive 'platforms' for third party use distracts from their role as vectors and enablers of gendered online violence. Firstly, they have an obligation to provide services that are safe to use, and to act against users who perpetrate online violence against others. Secondly, these companies should address their content recommendation algorithms, which are aimed at maximising user engagement and serve to escalate abuse through the promotion of misogynistic content and groups engaged in online harassment and abuse (Spring, 2021).[183]

For many women journalists around the world, Facebook (along with the company's other assets WhatsApp, Messenger and Instagram, which are now grouped under the new brand Meta), Twitter, YouTube and other services are essential tools for newsgathering, content distribution and audience engagement.[184] But the necessity to work in these spaces has resulted in a double bind: women journalists are heavily reliant on the very same services which are most likely to expose them to online violence. This tension is a feature of news organisations' dependent integration with big tech, a feature of what has been termed 'platform capture' (Posetti, Simon, and Shabbir, 2019), and it has been exacerbated by the COVID-19 crisis, which has made journalists even more reliant upon these technologies. This development may help explain why many journalists, including those interviewed for this study, said they had experienced "much worse" online violence in the context of the pandemic (Posetti, Bell and Brown, 2020).

The failure of these companies to ensure safe environments for many users is widely recognised. For women journalists, this curbs their ability to research stories, share journalism, and engage safely with audiences. But it also reinforces a climate of impunity for crimes against them - online and offline. For example, Al Jazeera principal Arabic presenter Ghada Oueiss raised concerns with the researchers about threats to her life which have been made with impunity on social media platforms. One person posted on Facebook that he would give US $50,000 to anyone who would kidnap or kill her. Oueiss called the police and the perpetrator was arrested, however the menacing threat remained on Facebook, she said, increasing the physical danger she faced.

---

183 An earlier version of this chapter was published by UNESCO and ICFJ on 3 May 2022. It is available here: https://en.unesco.org/publications/thechilling
184 Journalists interviewed shared their experiences of online harassment on other apps including Telegram, Clubhouse, Google Ads/Google Voice/Google Play and Discord servers amongst others.

Twitter's Nick Pickles acknowledged that more needs to be done to deal proactively with serious threats against women journalists on the platform, including those emanating from organised crime figures and cartels. However, a sense of impunity is emboldened by the platforms' failure to take action against the content and perpetrators involved in gendered online violence.

Almost without exception, the women journalists interviewed for this study complained about the companies' unresponsiveness, inaction, ineffective action and convoluted and cumbersome processes for reporting and escalating incidents. Some said that all this compounded the effects of abuse they endured on the platforms themselves. Getting the companies to deal with perpetrators of online violence against women journalists is "like trying to talk to God...pulling out a tooth from a child is easier", Catherine Gicheru from the African Women Journalism Project said. Additionally, the research participants were highly critical of the companies' perceived failure to recognise and adequately respond to the role of misogyny in attacks against women journalists on their platforms, especially at the intersection of racism, religious bigotry, homophobia and transphobia. **Some interviewees described these US-based companies' incapacity to deal with diverse cultures and linguistic variations as particularly problematic.** South African journalist and editor Ferial Haffajee said: "They treat their users in Africa like a colonial outpost".

Arbitrary features of content moderation, opacity of processes and responses, and corporate resistance to scrutiny and accountability for their role in violence against women journalists, were also heavily criticised by the research participants. This underscores the "distinct lack of clarity about what platforms are currently doing to combat abuse" (Dragiewicz *et al.*, 2018), which represents a major impediment to assessing the relative effectiveness of their responses to the problem (Suzor, van Geelen and Myers West, 2019).

Another key dimension that needs to be taken into account is the quickly evolving nature of online abuse tactics. As this study confirms, gender-based online violence against women journalists now occurs at the nexus of viral disinformation, digital misogyny, online conspiracy communities, and political populism and extremism, and it is also increasingly cross-platform. This demands sophisticated and collaborative responses to the problem. Ellen Tordesillas, President of Vera Files and columnist for ABS-CBN News in the Philippines, said: "When [the platforms] come up with preventive measures, it will not take long for [abusers] to come up with another way to circumvent. It's a continuous battle." For example, key among this study's findings is the shift by abusers to more subtle, less easily detectable, and less actionable forms of abuse, which nonetheless can be cumulatively devastating.

This chapter draws on over 714 women-identifying survey respondents, 15 country case studies (Kenya, Nigeria, South Africa, Pakistan, The Philippines, Sri Lanka, Lebanon, Tunisia, Poland, Serbia, Brazil, Mexico, the UK, the US, and Sweden) produced by the regional research teams attached to this study, and 184 long form interviews with journalists, editors, digital safety practitioners and experts conducted between July 2020 & October 2021. The chapter analyses the UNESCO-ICFJ survey responses and interview data collected as it pertains to women journalists' experiences of working on, and dealing with, the internet

PERCENTAGE OF WOMEN JOURNALISTS SURVEYED WHO REPORTED ONLINE VIOLENCE TO

39%

26%

16%

companies in the context of online violence. Research interviews were sought with senior representatives from Twitter, Facebook and Google to discuss the measures they were taking to address gender-based online violence. Twitter's Nick Pickles agreed, and he is therefore quoted in this chapter. Facebook declined the request, and Google said it could not identify an appropriate interviewee in time to meet the deadline set. In August 2021, UNESCO received a formal response from Facebook to a discussion paper previewing the study,[185] in which company representatives said they recognised the importance of the recommendations formulated in the paper.[186]

i. Not safe for work

The top five platforms or apps most frequently used for work by the 714 women journalists who responded to the survey were ranked as follows: Facebook (77%; n=550); Twitter (74%; n=528); WhatsApp[187] (57%; n=407); YouTube (49%; n=350); and Instagram[188] (46%; n=328). Although Twitter was used almost as heavily as Facebook by the respondents in the course of their work, Facebook was disproportionately identified as the service to which respondents most frequently reported online attacks (39%; n=279), with Twitter attracting complaints at the rate of 26% (n=186). Sixteen percent (16%; n=114) had reported instances of online violence to Instagram, while 50 women respondents had referred complaints to YouTube, and 43 to WhatsApp.

Interviewees in all 15 countries studied conveyed the same sense of futility as the survey respondents when it came to reporting online violence to the companies. "I feel nothing will be done, so I don't bother [reporting incidents to the platforms] anymore," Nigerian journalist Kiki Mordi said, echoing comments made by many other women interviewed.

Fatigue and frustration is further illuminated by the survey data identifying levels of dissatisfaction among the women respondents who had reported online violence to the services which they use in the course of their work. 17% (n=122) of survey respondents said they were "very dissatisfied" by Facebook's response. That was almost twice the rate of respondents who said they were "very dissatisfied" with Twitter's response to incidents they had reported to that company. Instagram was ranked third in the dissatisfaction stakes, followed by YouTube and WhatsApp. Facebook was also identified as the least safe of the high-use platforms globally among women journalists surveyed, with 12% (n=86) rating it "very unsafe" - almost double the number who rated Twitter "very unsafe".

---

185 The Chilling: Global trends in online violence against women journalists Research Discussion Paper elicited 26 preliminary findings on pages 12-18: https://en.unesco.org/sites/default/files/the-chilling.pdf
186 Internal Facebook communication to UNESCO received August 19, 2021.
187 WhatsApp is owned by Facebook/Meta.
188 Instagram is owned by Facebook/Meta.

ii. Inaction, inadequacy, and ineffectiveness

On International Women's Day in 2021, Director General of Sweden's Sveriges Radio Cilla Benkö wrote about the need to shift the onus for responding to gendered online violence to encompass an increased role for big techs: "[D]igital platforms need to take more responsibility for removing hatred and threats from their platforms. Threats targeting journalists as a group must also be taken more seriously. The level of measures in place and feedback have been far too little for far too long" (Benkö, 2021).

The women journalist interviewees largely regarded reporting incidents to the companies to be an act of futility. This assessment is the result of frequently frustrated attempts to get the companies to flag or remove misogynistic, racist, threatening and libellous posts, comments, memes, pictures or videos. The most frequently reported response they said they received from the companies when they did report abuse was that the material reported was not in breach of corporate policies and therefore unable to be addressed. Some also described waiting weeks, or even months, for threatening and abusive content to be removed – and some assessed that this might only happen after a certain (unknown) threshold of complaints is received about accounts or posts. Many said that they had reported incidents which were never even acknowledged.

While attitudes to the companies varied among the 113 women journalists interviewed for this study, they were almost universally critical of Facebook, and expressed little faith in the company's announcements and initiatives regarding online harassment and abuse. Many were also scathing about Twitter, but several journalists referenced what they perceived to be recent improvements in Twitter's reporting tools and abuse minimisation efforts, saying that they now feel "safer" or "less exposed" on the platform.

Overall, the dissatisfaction reported by the interviewees covers eight main areas of concern:

- **Inadequate and cumbersome abuse reporting processes.**

- **The absence of human-centred points of contact and response mechanisms.**

- **Poor, unidentifiable or inaccessible processes for escalation.**

- **Unresponsiveness - including non-responsiveness, and poor and inconsistent responses - to incidents reported.**

- **Concerns about an absence of gender-sensitivity and a lack of awareness about intersectional threats.**

- **Failure to develop moderation capabilities suitable for linguistically and culturally diverse communities of users.**

- **Failure to recognise the specific press freedom and journalism safety risks entailed, and respond appropriately on the basis of recognition that freedom from online violence and freedom of expression are not mutually exclusive.**

- **A lack of transparency and accountability demonstrated in policies and official responses to the problem of gendered online violence on their services.**

These identified concerns are elaborated below.

Shortcomings of abuse reporting tools and processes

The interviewees and survey respondents shared a common sense of frustration that they were left to block or mute abusive users themselves due to the failure of automated platform-based systems. Many also expressed concern that repeat perpetrators were often able to act with impunity. Further, when incident reports did elicit a response (usually after a significant time lag), the journalists said that their requests for flagging, muting or deleting offensive content or accounts were most often rejected.

Al Jazeera's Ghada Oueiss, who is the target of coordinated cross-platform disinformation campaigns allegedly involving State actors, said she had "lost count" of reports she had made to Facebook, Twitter, YouTube and Google. She described YouTube and Google Search as the sites of some of the worst abuse she has experienced. "You can never know who Ghada Oueiss is for her journalism. You only see attacks, attacks, attacks...you would think that either I'm a terrorist, or I'm a whore," she said of the smears prevalent on the sites. Oueiss also said Twitter was very slow to deal with tens of thousands of tweets sharing stolen and altered pictures of her that were part of a coordinated smear campaign.

Brazilian journalist Patricia Campos Mello said the companies' response to the online abuse triggered among their supporters was meagre. When she reported doctored images of her in 2018, she found Twitter to be "more agile" while Facebook "just ignored it". Campos Mello said she has virtually given up on using standard platform-based reporting systems.

In order to force action, Swedish magazine editor Susanna Skarrie enlisted help from an external consultant to liaise with Google and Facebook about removal of abusive content and the need to reduce traffic to websites targeting her, her family, and her colleagues. She said Google promised to stop search engine optimisation for websites publishing false information about her and her colleagues. However, this did not halt the websites coming back online: "Every time a new subpage emerges, we have to contact Google again," she said. Facebook told Skarrie that the abusive accounts she reported were "not illegal", but it did remove some. However, the accounts kept regenerating, "and are used to slander me and my family and other journalists who have investigated them," Skarrie said.

Doxxing and other digital security breaches which expose women journalists to increased offline threats are also rarely dealt with swiftly enough by the platforms.

For example, when Serbian journalist Jovana Gligorijević of *Vreme* said she was doxxed in YouTube comments in 2019 (Strika, 2019), she reported that her personal information was only taken down by the Google-owned company after the breach was reported more than 30 times.

Many interviewees also expressed concern about the companies only taking action after the prospect of immediate physical harm had become apparent. Facebook's standard for 'credible' violence, for example, requires language that incites or facilitates serious violence be judged "a genuine threat of physical harm" before moderators will act to remove content - although its requirements for assessing the seriousness of threats are not apparent (Facebook, 2020).

One example was provided by Guardian US investigative journalist Julia Carrie Wong regarding the 'Unite the Right' rally in Charlottesville, Virginia, which started as a Facebook event and resulted in the death of a woman. A few weeks before the event, Wong said she sent Facebook a spreadsheet with links to 175 neo-Nazi, white nationalist and neo-Confederate hate groups that were using the service to recruit and organise. She said Facebook had declined to take any action against the vast majority of them until after the woman, Heather Heyer, was killed during the rally. She said this chain of events raises questions about Facebook's role in facilitating and amplifying hate.

Various studies have made complementary points. An assessment of barriers to more effective responses from the internet companies by PEN America criticised Facebook's "byzantine" settings for allowing users to make profile or cover photos private, unlike the 'one-click' systems Instagram and Twitter have (Vilk, Vialle and Bailey, 2021). This functionality could help mitigate the problem of stolen images of women journalists being used by imposter accounts, or in the production of deepfakes. PEN America has also highlighted that purely reactive measures such as blocking and muting can mitigate online abuse once it is underway, but do not proactively shield targets. These points resonate with the argument that the companies need to "centre the voices of those who are directly impacted by the outcome of the design process" (Costanza-Chock, 2020).

Many interviewees and survey respondents said that the companies need to create rapid response units - staffed by multilingual employees with expertise in both press freedom and gender-based violence. They also wanted to be able to hand over their accounts to trusted colleagues for monitoring in serious cases. A series of practical product design solutions proposed by PEN America includes a privacy-preserving facility to allow trusted contacts to assume control of a journalist's accounts when she is under attack. They also recommended an 'SOS button' for those facing severe threats or large-scale pile-ons, which would allow them to access a support hotline and instantaneous in-platform protections (Vilk *et al.*, 2021). The World Wide Web Foundation's 'Tech Policy Design Lab: Online Gender-Based Violence and Abuse' report recommended a similar measure:
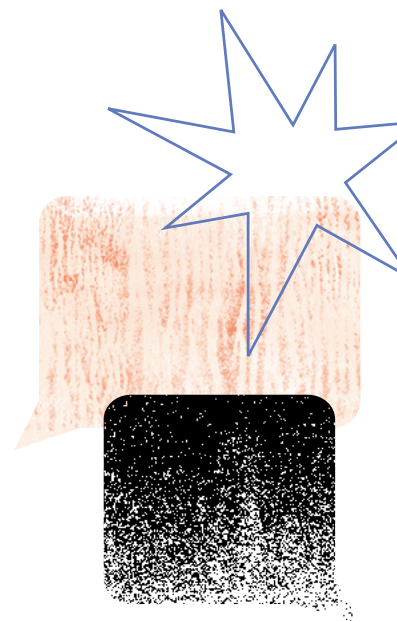
Users could assign different roles to trusted contacts, giving them authority to upload or delete content, or restrict and delete comments. Posts uploaded by trusted contacts could have the option to be marked with a verified 'trusted contact' badge….[However, it is vital] that companies do not shift too much responsibility onto people who volunteer to help others manage abuse (Dhrodia et al., 2021).

There was evidence from this study's research subjects that Twitter monitors some highly vulnerable journalists and intervenes technically to choke online abuse in real-time, sometimes making direct human contact to warn a target of impending threats. Julia Carrie Wong said: "I actually got alerted by a Twitter staffer who told me that [they] had seen some kind of threats or discussion…[So] I started to lock down quite seriously on my digital footprint to try to minimise things, and started using a Delete Me[189] service."

Former *The New York Times'* journalist Taylor Lorenz pointed to recent technical improvements on Instagram, such as comment filters and action against imposter accounts, which, while not being nearly enough to address the problem of violence against women journalists on the platform, are "a very good step in the right direction," she said. Lorenz has also tried contacting the companies' PR representatives via Direct Messages on Twitter, TikTok, Clubhouse, Substack and Instagram as a work-around to deal with the companies' non-responsiveness – although she said that they usually respond that "there's nothing they can do".

The need for human points of contact and proactive detection

The concerns expressed above underpinned the widespread call among the interviewees and survey respondents for the companies to employ many more human moderators and policy specialists with training in human rights, particularly in the areas of gender equality and press freedom, and in countries where the offline risks associated with online violence are most severe.

In the US, Al Jazeera's White House Correspondent Kimberly Halkett, who was doxxed and received death threats in 2020 across multiple platforms, articulated the research participants' shared demands for change. She said special response units which prioritise human contact are needed to deal with online violence complaints involving women journalists:

When you've been a target of sexual violence or any sort of gender-based violence of that magnitude, it needs a different phone number. It needs a different email address and it needs a quick reply. [But] these platforms make it very challenging to reach a human voice, and they do that deliberately. This insulates them from having to deal with the hassles of people like me.

---

189 At least a dozen interviewees based in the US or working for US organisations said their newsrooms had paid for them to use Delete Me, or they paid for it themselves. Founded by US company Abine in 2011, this tool removes personal data from websites, or 'scrubs the internet' of an individual's personal information.

Halkett suggested responses to address these deficits, such as a dedicated email address and phone number with human contact points, along with a dedicated ombudsperson for special categories of users (e.g., journalists and human rights defenders) to ensure appropriate action in cases of gender-based violence.

Overall, this research revealed the need for social media companies to respond proactively and pre-emptively to acts of online violence against women journalists. Instead of placing the onus for managing the abuse and harassment on those women targeted, many interviewees said that the companies needed to work harder to prevent such attacks at their point of origin, not wait for them to make complaints on a scale deemed necessary to trigger action.

When interviewed in June 2021, Twitter's Nick Pickles acknowledged that a shift was required by the platforms in responding to online violence incidents to avoid further burdening those targeted. He said the company was going in the right direction: "We definitely hear the feedback that the burden is way too much on victims. And that's something that we're working to change now in real time… Now we're at a point where more than half of all the content we removed is detected proactively by us". However, with Elon Musk's takeover of Twitter looming at the time of writing, there was deep concern among international human rights experts.

Company policy and human rights

While a State cannot legitimately mute a citizen permanently without disproportionate violation of the right to expression, a corporation can remove a user from its service. Such an individual can use, or set up, alternative channels to impart their opinions. The companies have no legal constraint to tolerate users who are routinely abusing the terms of use, injuring other citizens' rights in the process, and threatening democracy (Posetti and Bontcheva, 2021). By de-platforming abusers (based on due process of consumer rights, such as providing tiered warnings where appropriate, and appeal options), a company can end the prevalent impunity on its service, putting an end to a situation where online violence can continue to be committed without consequences for the perpetrators within this space. This depends in part on how the company interprets respect for human rights in general and freedom of expression in particular.

Within several US-based internet communications companies, there is a framed tension between allowing "free speech" (seen basically as a right to unfettered speech), and protecting other rights. But freedom of expression rights as enshrined in international human rights law do not uphold the right of a person to use online violence to limit the speech (or personal safety) of another person, especially not journalists - whose public interest service merits particular protection.

Even in the US, while the authorities are constitutionally restrained from restricting much speech, private sector entities are free to adopt restrictions that reflect the limits of what they permit in their realm. In addition, due to the limited liability offered them by the Communications Decency Act, they only need worry about legal consequences of carrying third-party speech that crosses the threshold of illegality, if and when such content is drawn to their attention and they take no corresponding action.

In this context, the companies have routinely deflected demands for action against much online violence on the basis that this is simply part of legitimate 'free speech'. This has led to numerous women journalists suffering the violation of their own freedom of expression - both as citizens and professionals. However, according to Brandy Zadrozny: "freedom to does not overtake freedom from [during] a harassment campaign. I say to the tech platforms 'why does this person's rights usurp this user's?'". This is a point echoed by PEN America's Viktorya Vilk:

Women journalists, journalists of colour, LGBTQ journalists are getting forced out of public discourse, which is increasingly taking place online, and sometimes getting forced out of their professions altogether. And so this whole idea that we can't do anything about online abuse, because if we do, we will damage people's free expression rights is wrong, because the online abuse itself is actually what's limiting the free expression rights of so many folks who are marginalised...already. Now they're getting marginalised online and in their professions to boot.

This perspective underlines the recommendation by the former Special Rapporteur for Freedom of Expression and Opinion, David Kaye, that companies align their conceptualisations of freedom of expression with international human rights laws and norms. This situation would then recognise that protecting against online violence is a legitimate restraint on speech. However, it would require much more investment by the companies if they were to accept and to follow such a voluntary commitment in practice.

There is a UN-level human rights framework for corporations that should guide their conduct. The UN-commissioned Ruggie principles (OHCHR, 2011c) are designed to prevent corporations from undermining human rights. Called the "UN Guiding Principles on Business and Human Rights" (OHCHR, 2011b), these require corporations to "avoid infringing on the human rights of others and... address adverse human rights impacts with which they are involved", while "taking appropriate steps to prevent, investigate, punish and redress such abuse through effective policies, legislation, regulations and adjudication". Meanwhile, the Rabat Plan of Action (OHCHR, 2012) is a UN operational framework that can serve companies seeking to balance freedom of expression rights against the need to curtail incitement to hatred, violence, hostility and discrimination. To date, there is no evidence of the companies agreeing to be held accountable in terms of commitments in this area.

Other initiatives include the B-Tech Project, facilitated by the UN's Office of the High Commissioner for Human Rights, which is designed to synthesise guidelines and tools for practical application to human rights questions in connection with business and technology (OHCHR, 2019b). The UNESCO-led UN Plan for the Safety of Journalists is a particularly relevant instrument (UNESCO, 2012) which social media companies could factor into their efforts to respond to gendered online violence against journalists. Further work by UNESCO seeks to promote transparency by the internet companies as a means towards accountability for

respect for human rights, and this will be followed in 2022 by specific focus on access to data concerning journalistic safety issues.

The 250 recommendations of the Information and Democracy Forum's Working Group on Infodemics[190] for structural reform to improve the platforms' governance, transparency and accountability could apply to the problem of online violence. Key among them: "Platforms should follow a set of Human Rights Principles for Content Moderation based on international human rights law: legality, necessity and proportionality, legitimacy, equality and non discrimination" (Forum on Information and Democracy, 2020). However, Brazilian journalist Patricia Campos Mello drew an effective comparison between the distinctions in the policy regarding disinformation and gendered online violence, whereas substantially more progress is being made with regard to dealing with disinformation in the context of the pandemic compared to gender-based online violence.

Additionally, there has been effective civil society work around content moderation standards such as the Santa Clara Principles of 2018, developed by a collective of human rights organisations, advocates, and academics (Electronic Frontier Foundation, 2021). The principles were updated in 2021 with an emphasis on respect for human rights frameworks and human oversight (ibid.). Facebook published its first Corporate Human Rights Policy in March 2021 - 17 years after the company was founded (Sissons, 2021; Facebook, 2021a). It makes an explicit commitment to the safety of journalists by promising to protect "professional and citizen journalists" (under the umbrella of 'human rights defenders') from online attacks. Twitter was working on a similar policy in late-2021.

In July 2021, Facebook, Google, Twitter and TikTok signed a World Wide Web Foundation (WWWF) pledge to tackle gender-based online violence (WWWF, 2021). The companies committed via the pledge to build better ways for women to curate their safety online by:

- **Offering more granular settings (e.g., who can see, share, comment or reply to posts);**

- **Using more simple and accessible language throughout the user experience;**

- **Providing easy navigation and access to safety tools;**

- **Reducing the burden on women by proactively reducing the amount of abuse they see.**

They also committed to implement improvements to reporting systems by:
- Offering users the ability to track and manage their reports;
- Enabling greater capacity to address context and/or language;
- Providing more policy and product guidance when reporting abuse;
- Establish additional ways for women to access help and support during the reporting process.

---

190 Lead author Julie Posetti was a member of the Working Group.

NBC News-MSNBC's Brandy Zadrozny has argued that a more radical transformation is required, with a need to de-platform propagators of online violence at first strike in serious cases in order to combat recidivism: "If [the platforms] were to adequately enforce their own policies against harassment, they'd lose half their users... The companies need to do a lot of soul searching and then come away with a vigorous plan. Their commitment to freedom from harassment needs to be larger than their commitment to say whatever the hell you want on the internet."

There is also the issue of whether companies will address recommendation algorithms for content, users and groups, which have been found to exacerbate the problem in some cases and to promote misogyny (see 'Gap 5' below for details).

Inconsistent application of standards

Even though there are the above-mentioned human rights standards and principles, and even though the companies have policies for dealing with online abuse (e.g., Twitter Help Center, 2021a), when women journalists report online attacks, these companies also often fail to adequately enforce their own rules (Amnesty International, 2018).

The unevenness of the social media companies' policy implementation and enforcement across different countries and languages worldwide is another significant challenge. One example is the investigation by the Balkan Investigative Reporting Network (BIRN) into how Facebook and Twitter deal with content violations in the region. This established that half of the posts they reported as hate speech, threatening violence or harassment remained online - even though they were in clear violation of platform policies (Jeremić & Stojanovic, 2021).

Another issue is response times. In Germany, where the Network Enforcement Act sets time limits for the removal of hate speech, the companies react much faster, and more decisively. Penalties for failing to act in a timeous manner include steep fines and the companies are required to report on the local personnel they employ, and the actions they take to remove hate-related content they assess to be evidently unlawful.

A further issue is double standards within community guidelines. For example, the deputy Editor-in-Chief of a women-oriented magazine in Poland, Monika Tutak, expressed frustration with the way her publication's content is censored by social media companies due to breaches of policies regarding nudity, while gender-based hate speech against her staff is not deemed to meet the requirements for removal. Tutak referred to the symptomatic 'ban of the nipple' that Facebook adheres to, resulting in their content being deemed inappropriate and removed or 'shadow banned' by Facebook on a number of occasions:

We don't trust this company...I very often report to them hate speech and they don't react. I am afraid of the opposite situation, when our journalists are being blocked by Facebook. For example, we had a topic about period poverty, and there was an illustration of a woman with a stained skirt and Facebook blocked our journalist for 24 hours for this
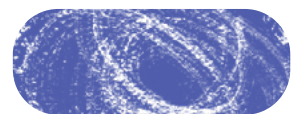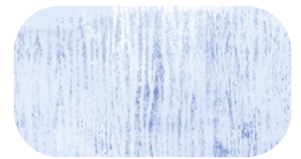
story... Facebook sometimes cuts the reach [of our posts] and it mostly applies to content about the female body.

Facebook has been accused of censoring the accounts of journalists in a number of other countries, and de-platforming them without providing justification. In Tunisia, 60 journalists and activists had their Facebook accounts deleted without warning or explanation in 2020. Anti-corruption watchdog Iwatch managed to get 14 deleted accounts reactivated after lodging complaints (Cordall, 2020). This pattern also highlights the double-edged sword associated with blunt content moderation policies, and the need to balance responses to gendered online violence against journalists with broader freedom of expression considerations.

Content removal is often requested in response to online violence cases, yet content removal can also work against women journalists unless there is oversight by teams with freedom of expression expertise and local contextual knowledge. Several interviewees criticised the social media companies for frequently failing in regard to balancing the need to protect women journalists against the need to respect freedom of expression. They pointed to arbitrary censorship, a lack of transparency and 'shadow banning'.[191]

The extent of transparency is also at stake. According to the transparency reports of Twitter, Facebook and Google, between 2017 and 2020 the Mexican authorities made more than 38,659 requests for the removal of content. In 95% of these cases, no information exists about the nature of the content, and there is no accountability associated with its removal (ARTICLE 19, 2021b). According to ARTICLE 19, however, only 6% of the requests for content removal to Google made by public officials in Mexico were granted.

Available tools

There have been many policy announcements from the platforms regarding online abuse and harassment. Most interviewees dismissed these efforts as 'PR exercises', while also welcoming a number of platform initiatives. For example, Twitter's 2021 rollout of features such as allowing users to limit the ability of non-followers to reply to tweets and the option to remove followers was welcomed by a number of journalists, who said they now feel "safer" on the platform.

> **Rianna Croxford** ✔
> @The_Crox
>
> Would definitely recommend "muting" notifications from Twitter users who haven't confirmed their phone numbers or email addresses. Means you don't see 99% of the messages from trolls.
>
> 7:30 PM · Jun 14, 2021 · Twitter Web App

---

191 Poland's Panopticon Foundation perceives shadow banning as a form of censorship where either users or the reach of their content can be blocked by social networks in a way that they are unaware. A user who is officially blocked or removed, at least theoretically can appeal the decision, while shadow banning is arbitrary and there is no question of transparency here (Obem and Głowacka, 2019a).

Figure 1: After being targeted, BBC Investigations reporter Rianna Croxford adjusted her Twitter settings so as not to be notified about all tags. "My [Twitter] notifications are now set in a way now that…if people comment and I don't know them, or I don't follow them, then I don't see it." [192]

Another Twitter feature that the interviewees found helpful in reporting and documenting online abuse is the ability to attach multiple tweets to a single report. This allows users to flag additional context and makes it easier and faster to provide proof that a particular account is being used in abusive ways, instead of requiring those targeted to submit a list of the attacks as they occur (Tang, 2016). Taylor Lorenz formerly of *The New York Times* described this approach as extremely important for tracking and tracing online abuse in real-time under extreme stress: "If we don't have the screenshotted receipts, it's like it never happened". NBC News-MSNBC reporter Brandy Zadrozny said she documented her own online abuse on a spreadsheet, specifically with regard to Google Voice messages, recurring email addresses, or phone numbers calling to abuse her.

There are several third-party documentation facilities specifically designed for recording abuse in various stages of development. These include JSafe[193] and DocuSAFE.[194] Both of these apps still require users to manually track and enter data, but they offer a single place to store and organise it. Google's Jigsaw is also experimenting with documentation and reporting tools (including a 'harassment manager' which was still in beta development in December 2021), leveraging their machine learning system Perspective API. This system, used by over 200 partners including *The New York Times*, detects toxic language to help targets of online violence take action in a more streamlined manner (RE•WORK 2021, Jigsaw 2021). Tune[195] is another Jigsaw tool developed to address online toxicity, while Jumbo,[196] Sentropy Protect,[197] Tall Poppy[198] and BodyGuard[199] are other offerings that help users change their settings on social media platforms. PEN America has recommended an abuse documentation facility that captures and aggregates screenshots, hyperlinks, and other publicly available data "automatically or with one click" (Vilk *et al.*, 2021).

Such tools are important to aid police investigations and legal action against perpetrators, but the need for them highlights the difficulty of navigating privacy and security within the platforms themselves. To be really effective, tools like this need to have in-built facilities to instantly submit abuse reports directly to the platforms for action and escalation. To date, the companies have resisted such recommendations on the basis that they do not have capacity to deal with alerts at scale.

---

192 Rianna Croxford tweet from June 2021: https://twitter.com/The_Crox/status/1404491531843670017; interviewed 03.03.21.
193 JSafe is a mobile app in beta developed by the Reynolds Journalism Institute at the University of Missouri with the Coalition For Women in Journalism: https://womeninjournalism.org/jsafe
194 DocuSAFE is a free app created by the National Network to End Domestic Violence in the US: https://www.techsafety.org/docusafe
195 Tune is a Chrome extension created in 2019 that employs machine learning to allow users to "control" the volume of the conversation they see, for example in"customising" toxicity in comments: https://chrome.google.com/webstore/detail/tune-experimental/gdfknffdmmjakmlikbpdngpcpbbfhbnp?hl=en
196 Jumbo is a third-party app using a subscription-based model allowing users to connect to various platforms and control their search history, messaging and other data: https://techcrunch.com/2020/06/24/privacy-assistant-jumbo-raises-8-million-and-releases-major-update/
197 Sentropy uses natural language processing and machine learning to "protect users and their brands from abuse, harassment, and malicious content". It was bought by Discord, the online chat platform, in July 2021: https://techcrunch.com/2021/07/13/discord-buys-sentropy/
198 Tall Poppy is described by Canadian founder Leigh Honeywell as a 'digital public health nurse' - a platform helping individuals to take charge of their cyber-security, also via incident response, and the service is sold to companies: https://www.youtube.com/watch?v=xpcb2n6uvzE
199 BodyGuard is an ad-free, free to use mobile application for individuals using contextual, algorithmic analysis to detect and avoid toxicity online. It also offers businesses API services, including digital media companies: https://www.lefigaro.fr/secteur/high-tech/haine-en-ligne-bodyguard-s-attaque-au-secteur-du-jeu-video-20211006

Twitter's Nick Pickles says the company takes enforcement actions based on their hateful conduct policy, which was extended to prohibit language which dehumanises others on the basis of religious affiliation, caste, age, disability, disease, race, ethnicity, or national origin, sexual orientation, gender, gender identity, or serious disease in 2009 (Twitter, 2020). Tweet or account removal is carried out by human moderators. This follows either reporting from the abuse target (Twitter Help Center, 2021c), other platform users, or through proactive detection through machine learning tools that flag posts for moderator review due to their similarity to known violating content. Pickles said the company is also experimenting with 'nudges': "...to say to people 'this reply may well be seen as abusive. Are you sure you want to post that?' [But] machine learning is definitely not a perfect science. And so one of the things that we just have to be really careful with is, for example, catching counterspeech, we don't want to be catching people who are quoting people in these filters".

From a user perspective, some of the women journalists interviewed reported that they still appreciated the 'block' option on Twitter,[200] which - despite placing the onus on the journalist to deal with the abuse - can serve as a means to stem online harassment. Most recently, Twitter introduced a silent block option.[201] This means that products and policies that empower users to mitigate the impacts of abuse, while also relieving those under attack from the onus of responding, could be optimally effective. The proactive tool Block Party[202] automatically mutes people at scale, and MegaBlock[203] blocks a person and every person that has 'liked' their tweet.

On Twitter, another effective feature which users have reported protects them from spam and abuse enables users to control who can reply to their tweets. There are also options to filter notifications from people not followed by the user. In March 2020, Twitter started a dedicated gender-based violence search prompt for hotlines and support in local languages in partnership with local NGOs, government agencies and UN Women. First launched in Mexico, the prompt later became available in 27 countries and 20 languages. In August 2021, a Safety Mode feature was released by Twitter which auto blocks accounts with potential harmful content. This has been followed by an experimental feature which flags heated conversations and prompts users to be 'respectful and truthful'.

The comment lock option on public Facebook profiles also gave some women journalists interviewed temporary respite, but the downside is that it also blocks audiences from contributing useful information or providing support. Similarly, Facebook Messenger's filter feature (that sends messages from non-'friends' to a separate folder) can stop journalists from receiving messages from genuine sources and contacts, with adverse implications for their journalism practice. Serbian journalist Jovana Gligorijević, who had to resort to filtering her Facebook account so that everything goes to spam unless from close friends, recognises

---

200 Details of account blocking on Twitter are here.
201 In October 2021, Twitter added this feature which offers people the option to remove followers without them noticing them or needing to block them.
202 Block Party is an anti-harassment paid subscription service founded in 2021 by software engineer Tracy Chou, who was stalked in real life by an online follower. The service currently works on Twitter by directing users to a "lockout folder": https://www.fastcompany.com/90686948/inside-the-life-of-a-tech-activist-abuse-gaslighting-but-ultimately-optimism
203 MegaBlock is a third-party service which allows you to block a tweet but also the accounts of anyone who liked that tweet. It is created on a Discord server called Gen Z Mafia, a community of young tech workers operating collaboratively: https://megablock.xyz/, https://www.nytimes.com/2020/09/15/style/gen-z-tech-mafia.html

that this excludes people asking her to investigate stories. "I miss the chance to help someone, but I have to do it to protect myself," she said.

A Facebook policy change from mid-2020 involves the ability for its users to register as journalists, to justify stronger security protections. The feature was initially only available to journalists in the United States, Mexico, Brazil and the Philippines (Facebook Journalism Project, 2020). In late February 2021 it was rolled out in 19 additional countries. Due to limited publicly available information about it, it is also not yet clear how effective this new feature is, or what the uptake rate has been. It was also not possible at the time of writing to determine if this system entails human contact points and rapid escalation of reports lodged by the registered journalists. Most of those interviewed in the first four countries where it was introduced were not aware of the feature. Several who were aware nevertheless expressed scepticism, saying that they did not trust Facebook with the process of registration for a range of reasons, including data privacy concerns. However, a Filipino interviewee highlighted Facebook's takedown of pages for what the company calls Coordinated Inauthentic Behaviour (CIB) as a "good start".[204] A 'Journalists' Safety Guide' developed as part of the Facebook Journalism Project (Facebook for Media, 2021) was also noted by one interviewee as useful in providing tips on how to stay secure and report abusive behaviour on the platform.

Dealing with the large-scale problem of online abuse via anonymous and fake accounts is also an ongoing challenge. At one point, Al Jazeera's Ghada Oueiss said there were 165 fake Facebook pages in her name which were used to shame and defame her. She said that Facebook advised her that the only way to counteract this would be to set up her own professional page - since she is a "public figure" - which they would then verify. Thus, they put the onus of verification on her, not the abusive imposters, stating that whenever a fake account posted something in her name, she could prove that "this is not me". In mid-2021 Oueiss' verified Facebook page had 2.2 million followers,[205] and the page remained a constant target for abusers.

165
FAKE FACEBOOK PROFILES IN GHADA OUEISS' NAME

User-based pressure applied to the platforms

In Tunisia, some interviewees said they had resorted to legal action in order to force social media companies to act against the online violence they reported. This is because abusive content is usually not removed by the companies until the court orders it to be removed, according to Ayoub El-Ghadamsy, a lawyer with the National Syndicate of Tunisian Journalists. While a number of other women journalists interviewed for this study were considering legal action against the platforms after their corporate incident reporting mechanisms failed, there are occasional examples of companies removing content associated with targeted online violence against journalists, even when the courts fall short.

Jessikka Aro, a reporter at Finland's public broadcaster YLE, has been a target of organised 'troll' campaigns linked to a foreign State actor for her reporting on

---

204 Under its Community Standards Facebook defines Coordinated Inauthentic Behaviour as "the use of multiple Facebook or Instagram assets, working in concert to engage in Inauthentic Behavior, where the use of fake accounts is central to the operation": https://www.facebook.com/communitystandards/inauthentic_behavior
205 Ghada Oueiss' official Facebook page: https://www.facebook.com/ghadaoueiss14/

disinformation networks since 2014 (Aro, 2016; Ireton and Posetti, 2018). In 2020, Aro lost an application for a restraining order against two individuals whom she described as "far-right YouTube livestream harassers", and whom she accused of stalking her (Pohjola, 2019). However, in this case, YouTube closed the offending video channel, which carried a message saying: "This account has been closed due to repeated or severe violations of YouTube's policy prohibiting intimidation, harassment, and harassing content" (ibid.).

In the experience of Al Jazeera's Ghada Oueiss, the companies' responses also depend on the amount of attention the journalist or incident is getting, the number of followers a journalist has, and the extent of involvement of international organisations. Oueiss said pressure does help to evoke a quicker and more definitive response from the platforms in her experience. Solidarity exercises can also include mass-reporting of abusive accounts. Two Tunisian interviewees - Najoua Hammami, investigative journalist and director of the Media Office at the Arab Institute for Human Rights, and Khaoula Boukrim, Editor-in-Chief of Kashf Media, said they deployed mass reporting tactics when they were under attack as a defensive measure. They recruited their friends and colleagues to report abuse against them to the companies en masse as a way of escalating their complaints. Similar tactics were described by Kenyan journalists.

However, automated content removal has been weaponised against women journalists, among others. Interviewees from around the world described being affected by manipulation of the platforms' automated reporting systems. This occurs through mass false reporting of legitimate content as abusive in coordinated campaigns designed to automatically 'de-platform' (i.e. have their accounts suspended) the targets and censor their journalism through blocking or platform takedowns.

Nigerian journalist Kiki Mordi said she had seen abusive users with a large follower base encourage mass-reporting of legitimate accounts and content they dislike, such as that shared by feminists like Mordi. Mordi noted: "Just one or two accounts have been brought down for harassing me, as opposed to the hundreds everyday that continually harass me". And even when an abusive account is blocked or suspended by a platform, the same user can make a new account under a new pseudo-identity or use a Virtual Private Network (VPN) connection to prevent tracing.

iii.Policy gaps

Six core company policy gaps pertaining to gender-based online violence were identified in the course of research for this chapter. They are elaborated below with a view to informing responses.

Gap 1: Lack of prioritisation of gender-sensitivity in policy and implementation

A key reform needed to more effectively address online violence against women journalists is the development of gender-sensitive policies that recognise the

increased risks that women are exposed to on social media platforms, along with the exponentially worse offline impacts. These include, but are not limited to, threats of sexual violence, but they also extend to sexist attacks that portray women as sexually immoral and/or highly sexualised beings, non-consensual sexual imagery, and sexually explicit content (e.g., graphic images of male genitalia sent via DM) that are all used to harass women journalists.

The 15 detailed country-level case studies produced and drawn on for this report consistently demonstrated the need for comprehensive gender-sensitive policies designed to ensure that women, and in particular women journalists, can work safely on the platforms. In Brazil and Poland, for example, the research highlighted content monitoring and interviewees' accounts that suggest that companies may censor feminist posts more than hate speech and gender-based attacks (Martins *et al*., 2020). Other research shows that Facebook in Sri Lanka in effect allows a culture of casual sexism and misogyny, expressed through sexual harassment and non-consensual dissemination of images, including intimate pictures and videos with derogatory, abusive and violent captions (Perera and Wickrematunge, 2019).

Existing policies do not appear to proactively cover instances of targeted online violence. Neither do the companies appear to deal with pernicious hashtags like #presstitute, which are used to discredit women journalists personally and professionally in tandem, while also exposing them to increased risk in some contexts. Former CNN editor Inga Thordar said there was a need for the platforms to make misogyny and intersectional abuse much higher priorities, and that platforms "should be taking much stronger action against people who are persistent perpetrators of online harassment".

In addition to gender-aware policy improvements, Pakistani journalist Benazir Shah said the companies also needed to be more responsive to women users. She echoed many calls from interviewees around the world for gender-sensitive and in-country contact points to be made available for women journalists because of their particular exposure to risk in the course of their work on the services. When Facebook opened an operation centre in Nigeria, the company worked with the civil society organisation Paradigm Initiative to promote their online safety features, although not specifically in relation to online violence against women journalists.

In countries like Mexico, where there are high femicide rates (SSPC, 2021; UN Women, 2017), and extreme risks faced by journalists in general (RSF, 2021k), substantive attention needs to be paid by the companies to combating gender-based online violence. For example, columnist and political scientist Denise Dresser, who has been a victim of orchestrated online violence including death threats, said that she perceives a lack of responsibility and inaction associated with the social media companies' responses which increases the offline risks she faces.

With regard to closed groups and/or encrypted communications, there is also a policy gap concerning gendered online violence. For example, in 2019, Facebook made a strategic shift, moving away from a focus on public sharing and instead promoting closed or semi-closed Facebook Groups (Ingram, 2019a). Such an approach presents a twofold problem. Firstly, although women journalists do indeed create closed online communities to support one another during attacks, this encourages them to retreat from visibility in order to protect themselves.

It amounts to putting the onus on targeted people (i.e. women) to withdraw in response to abuse and attacks, rather than addressing the business model and design failures at the core of the problem. Executive Director of the US-based Representation Project, Soraya Chemaly, told the researchers: "What does it say if on your own platform women are hiding away in special groups because they cannot speak openly in their own space?" Secondly, it is evident that closed Facebook groups and Facebook Messenger, as well as WhatsApp, are also subject to less scrutiny (external and internal) of online violence and disinformation content (Ingram, 2019b). This is one reason these channels are targeted by perpetrators. Since many of the research participants described receiving the worst gender-based threats and harassment via such closed channels, this points to a priority policy gap that needs to be addressed.

Finally, policy measures and processes also need to be more technologically sophisticated and transparent to respond to gender-based hate speech that takes the form of synthetic media attacks, such as doctored images and deepfake videos.

A short case study in systemic policy and process failures

Facebook's dealings with Northern Ireland's Patricia Devlin - a reporter with the *Sunday World* newspaper until early 2022, highlight serious gaps in the company's approach to dealing with online violence against women journalists. Devlin has received multiple death threats from figures associated with neo-Nazis and paramilitary extremism. She also received threats of sexual violence against her baby via Facebook Messenger after she had been doxxed. "I received a message via my personal Facebook account, and it said 'Don't go near your granny's house in Maghera, Tricia, or you'll watch your newborn get raped'. And it was signed off in the name of a neo-Nazi group called Combat 18, which in the past has had links to loyalist paramilitaries in Northern Ireland." This threat, sent in 2019, continued to be investigated by the Police Service Northern Ireland (PSNI) in late 2021. But Devlin received a renewed threat against her child (again via Facebook Messenger) in mid-2021.
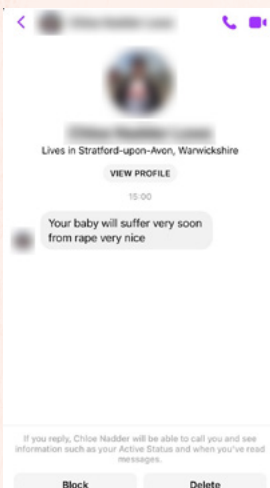

Figure 2: The threat sent to Patricia Devlin in May 2021, promising to rape her baby, which Facebook said it had difficulty investigating. The message was still visible to Devlin in July 2021. The name and face of this user have been obscured in accordance with research ethics protocols.

Devlin, whose experiences were initially documented in a UNESCO-published discussion paper associated with this study in April 2021 (Posetti *et al*., 2021b), contacted the researchers for assistance regarding this threat, after she reported the profile to Facebook. However, a senior Facebook representative, who had earlier asked for the company's self-reported efforts to combat online violence against women journalists be included in this study, told the researchers and UNESCO representatives that they had no record of Devlin having reported the incidents to Facebook.

In an email sent to the lead author on 13 May 2021, Devlin described the process of trying to get Facebook to take action to protect her by removing the threats and investigating the account as "absolutely staggering". She said that while she and others had reported the profile within the Messenger app multiple times, Facebook had not acknowledged the reports. As intermediaries, the researchers supplied Facebook with screenshots of the threat and confirmed details of Devlin's profile. But Facebook responded that screenshots did not help - they needed the researchers to supply the URL address for the Facebook profile of the person attacking Devlin because, they said, it was too difficult to trace the Facebook user through screenshots of their profile. This put the burden to document and investigate the threats on the journalist under attack, but the researchers supplied the URL required by Facebook

Facebook also suggested that Devlin apply for a "blue badge verification" to verify her profile, and said she should use Facebook Messenger's existing tools to protect against unwanted messages, such as disabling unsolicited messages, or block the user within Messenger (Facebook, n.d.b). However, Devlin noted that these measures were not enough - blocking would not make the threat disappear. She wanted the incident investigated and the perpetrator de-platformed. In addition, disabling unsolicited messages was professionally unadvisable for Devlin because sources routinely contacted her via Facebook Messenger. Facebook also told the researchers to advise Devlin that she could use a "powerful FB anti-harassment feature" - controlling who can comment on your public feed (Facebook, n.d.c) - but this particular threat did not come via her public feed.

Patricia Devlin

Then, the Facebook representative told the researchers that they were in direct contact with Devlin. But the journalist denied this was the case in a message she sent to the lead author on 24 May 2021: "They did respond to an email from my company where they AGAIN asked what I had done to report and block the profile; when I did it; how I did it; and claiming what I had already said, to you and my company, did not match their internal records."

This episode followed serious instances of Facebook-based online violence experienced by Devlin in 2020, which remained under police investigation in mid-2021. Police visited the homes of Devlin and a colleague to warn them of credible death threats following publication of false content about the journalists (including blaming Devlin for a bombing) on a Facebook page used by loyalist paramilitary groups. Devlin had reported the attack on 27 November 2020 to both Facebook and the Police Service of Northern Ireland (PSNI). After a second death threat, police again attended Devlin's home to inform her that they had received intelligence that she would be shot in the following 48 hours. Devlin made a statement to the police and provided the names of two individuals associated

with the attack, but the PSNI could not go and speak to these people because Facebook said it was unable to trace the page. It was still active and running in late 2021, heightening the offline risks she faces. Devlin reported the posts separately to Facebook but she was told that they did not breach its 'community standards'.

Additionally, Devlin said she also reported a page to Facebook and the police which had accused her of being behind her own death threats - a classic disinformation tactic deployed in orchestrated online violence against women journalists. This so-called 'community page' is being run by operatives linked to a paramilitary drugs gang, according to Devlin: "The individual who is involved with this page has been at the centre of the abuse and I have named him previously. He uses multiple fake profiles and 'anon' pages on Facebook to target me and I made numerous statements to police about this which are still being looked into".

Offline threats - including death threats taking the form of graffiti painted onto brick walls in Belfast - were also connected to these online attacks on Devlin.

Conclusion: In the specific case of Patricia Devlin, Facebook failed to respond effectively to some of the most serious examples of online violence surfaced by this study. Their response mechanisms were disjointed at best. The obfuscation, victim blaming and deflection that many research participants described as re-traumatising in the context of online violence attacks were also in evidence. The proprietary tools that they provided the researchers to share with Devlin were ineffective or inapplicable, and required her to do the labour of verifying and documenting the abuse. Even when the Police Service of Northern Ireland called on Facebook for assistance to trace the source of the threats, they were deflected. Despite its technical expertise, Facebook was apparently unable to trace the account of a user threatening to rape a journalist's baby despite being handed the account holder's name and screen grabs. When URLs connected to the profile were supplied by the researchers at Facebook's request, the company said it had no record of Devlin's multiple incident reports made via Facebook Messenger.

Facebook declined to provide a representative for a research interview for this study.

Gap 2: Inadequate responses to 'below-the-radar' abuse and linguistic and cultural content moderation challenges

This study points to inconsistencies in how the platforms detect and act on abuse across different languages and countries. While abuse moderation in languages spoken in the companies' major markets is somewhat better addressed, this is far from being the case in less prominent local languages, including in countries such as Kenya, Nigeria, Sri Lanka, the Philippines and Pakistan. At an intersectional abuse level, interviewees highlighted that being abused in languages such as Persian,

Sinhala, Tagalog, Malay, Urdu or Tamil adds to the frustration of trying to report abuse to the companies which have very limited abuse moderating capability in these languages, and therefore they often do not respond appropriately.

The cultural and linguistic context of online violence is currently not captured by the platforms' algorithms and policy implementation workflows. This failure was strongly criticised by the women journalists and policy experts interviewed for this study. This reflects that "87% of Facebook's global budget for time spent on classifying misinformation goes towards the United States, while 13% is set aside for other countries — despite the fact that North American users make up just 10% of its daily users" (Popli, 2021). Languages such as Arabic are particularly poorly-served. For instance, according to one study, only six percent (6%) of Arabic-language hate content was detected on Instagram before it made its way onto the photo-sharing platform, and just 40 percent (40%) was proactively detected and taken down from Facebook (Scott, 2021b). Women and the LGBTQ community were primary targets (ibid.).

A key problem identified both by the interviewees and other recent recent research (Posetti *et al*., 2020; Popli, 2021; Scott, 2021b) is that the companies have insufficient numbers of human content moderators and local policy staff who know and understand the nuances of international socio-cultural and political contexts, with linguistic capabilities that extend to minority languages and dialects. Women journalists interviewed said they found it hard to explain the specific cultural implications of abusive language and the type and severity of the attacks against them in relation to universal platform community standards when they reported abuse to the platforms, resulting in delays or a lack of action.

According to interviewees, in Pakistan, these companies are particularly weak in responding to hate speech and harassment against women journalists in Urdu - one of the country's two official languages. This was a point underscored by Sabahat Zakariya, who has reported for various news outlets in Pakistan (including BBC Urdu). She said the companies are not responsive to specific cultural problems and their repercussions:

> *They don't sometimes understand how crucial it may be for somebody's life. There is no response at all, especially if it comes with [some kind of] a hashtag or abuse in the Urdu or Punjabi language, or another local language. You try to explain to them the nature of the abuse or threat, but they don't get it. They simply don't have [enough] human resources from these countries to understand the nuances of what is going on. So it's a very, very, one-size-fits-all approach, and it's really wrong.*

The lack of an Afrocentric approach by social media also needs to be urgently addressed according to Kenyan editor Catherine Gicheru:
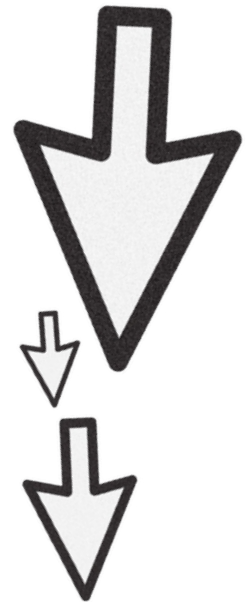
Not all black women online face online harassment the same way...If you're African, you're black, you're a woman, you will experience online harassment differently than if you are an African American woman online...And just the fact that you're Zimbabwean might attract some kind of harassment.. And then it goes even further, it can go lower because you're from Bulawayo or because you are a Shona and not a Ndebele.

Broadcast journalist Cecilia Maundu pointed out the difficulty of reporting on the platforms using different languages in Kenya, where groups online are formed in vernacular languages. According to Catherine Gicheru of the Africa Women Journalism Project, much of the online harassment she witnesses, such as body shaming remarks or subtle threats, is not in English and so is not monitored or "understood" by the platforms' algorithms. The Media Council of Kenya has lobbied the social media companies - including Facebook and Twitter - to create tools that flag offensive content, aid the reporting of it, and enhance moderation in multiple languages, and with appropriate cultural context.

A particular example is the experience of the Al Jazeera journalist Ghada Oueiss, who reported to Twitter a fraudulently altered video purporting to show her naked in a jacuzzi, screengrabs from which she said were retweeted 40,000 times. The video and a series of photos showing her eating a meal with colleagues were stolen from her phone, she alleges, as part of an orchestrated attack designed to discredit her.[206] They were distributed with messages alleging she was an alcoholic, drug-addicted prostitute. Twitter's initial perspective was that the content did not violate their policies. However, being a journalist working in the Arab States, the wide circulation of this stolen and manipulated material put Oueiss at risk of retribution and significant reputational damage.
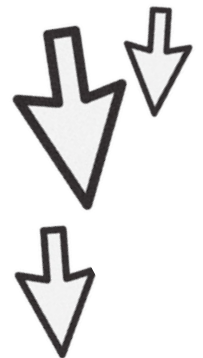
Lawyer Caoilfhionn Gallagher KC said that getting the social media companies to take action against those perpetrating online violence against her BBC Persian service clients in Farsi was extremely difficult due to language issues. Similar points were made by women journalists in the Philippines, Sri Lanka, South Africa, Lebanon, Mexico and Brazil.

Twitter's Nick Pickles stated that building capability to moderate content and respond to threats in diverse cultural and linguistic context on a global scale is extremely challenging and accepted that human intervention would also likely be required to deal with more nuanced cases.

The failure to effectively address multilingual content moderation also has negative freedom of expression implications. There is the need to counterbalance such interventions against the potential for overreach and censorship justified on the grounds of curtailing abuse. In particular, linguistically deficient algorithms not only miss abuse, but also frequently delete abusive posts. For instance, in one study

206 Ghada Oueiss said she knew she had been hacked when private photos on her phone were shared online, and her claim was confirmed by Forbidden Stories and Amnesty International. Leaked documents showed that Pegasus spyware, created by Israeli surveillance technology company NSO Group, had been installed on Oueiss' phone, turning it into a surveillance device (Solon, 2021).

(Scott, 2021), 77 percent of deleted content on Facebook in the Arab States was found to be non-violent and legal, meaning the algorithms harmed people's ability to express themselves online, and limited the reporting of potential war crimes.

In Nigeria, journalist Kiki Mordi said algorithms need to be localised, "…because some of the offensive words that Nigerians use aren't picked up". Adeboye Adegoke, Senior Program Manager at Nigeria's Paradigm Initiative, said that understanding cultural context is important when designing algorithms so that cultural nuances, and the different ways abuse manifests itself around the world, may be better detected and prevented. For him, that means major tech companies must employ locals to ensure there is diversity of personnel and diversity of algorithmic engineering.

In Brazil, Gabi Coelho of daily newspaper *O Estado de S. Paulo* noted that black women are particularly at risk of hate speech on Twitter: "Twitter, the network I use the most, is very violent for black women, especially, so there is the gender issue and there is also the race issue … There was a week when I came to report more than 40 profiles."

In addition, the current algorithmic processes for content moderation and removal appear to be prone to bias and reinforcement of inequality while also being opaque. ARTICLE 19's Thiago Firbida noted that while people can see the impacts of online violence and its disproportionate effects on women, Black and Indigenous people, there is an "invisible structure that is built by white men, straight Europeans and Americans" which means that it is not possible to "see how the platform's algorithms facilitate the spread of attacks".

The failure to remove misogynistic posts from public fora entails a risk of entrenching cultural, racist, sexist and abusive bias into the next generation content moderation machine learning models, which are being trained on present-day abusive content (Hao, 2021). Twitter's Nick Pickles said: "The reason for investing in countries like India, like Ghana, is to build out that capacity where, being totally frank, we don't have the right diverse capacity that we need. And so those investments are explicitly intended to start building that capacity".

Failure to detect abuse 'below the radar'

As the companies' abuse prevention methods have begun to improve with regard to addressing highly explicit abuse and hate speech, abusive behaviour has started to shift towards harder to detect implicit cases. Such implicit abuse is often contextual in nature and grounded in cultural and political specifics, which makes it hard not only for the platforms' machine learning algorithms to flag, but also for employees who are not from the same culture and country to moderate. Additionally, rape and death threats go undetected when expressed via implicit wording, imagery and memes. One example is when the so-called Islamic State evaded moderation on Facebook via the use of local Arabic slang to spread hate speech (Scott, 2021).

One particular source of complaints from the interviewees was the companies' failure to address online violence that occurs via their private messaging services, such as Facebook Messenger, Direct Messages on Twitter (DMs) and Instagram
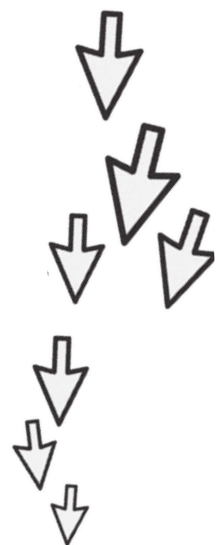
Direct Messenger. These avenues are frequently used by perpetrators to deliver death threats, or sexual harassment through unsolicited sexually explicit messages and images (one of the top attack modes identified by the survey respondents). Where there is privacy preserving end-to-end encryption, this makes the task of monitoring and moderation challenging – although meta-data can reveal behavioural patterns and flows of content, which the company can control if it so decides.

The BBC's Marianna Spring described the frustrating process of trying to report threats and abuse on Facebook Messenger. This has involved her compiling and emailing evidence of the abuse (including links and screengrabs), then being asked to repeat the process of logging the abuse using Facebook's on-platform reporting tools (which rarely generate a response), and finally being told that nothing could be done because the content was "private". At the time of writing she had escalated cases to the police for investigation, but meanwhile the reported accounts remain active on Facebook.

Vice's UK Editor-in-Chief Zing Tsjeng reported a user on Twitter Direct Messages who called her a "C-word who eats dogs" in May 2021, but Twitter did not take any immediate action against them. Twitter's Nick Pickles acknowledged that his company places the onus for monitoring online abuse sent via Direct Message on the receivers: "We do rely on users flagging to us things in their DMs that we should look at... This is also something that, potentially longer term, we can use technology for. But I think it's fair to say that the complexities of moderating private spaces, and the legal frameworks that go around them, is greater than the public space on Twitter." He added: "We expanded our policies to cover unwanted sexual advances, which I think for female journalists do often come through DMs. And if they're reported as a DM, we can investigate them, and look at them."

In Sri Lanka, women journalists also described voice-based harassment received via these closed channels that are even harder to police. In addition, according to our interviewees, some abusers use intimate verbal threats, calling reporters via Google Chat or Facebook Messenger and shouting "murderer" or "racist". Twitter Spaces was also identified by one user in Pakistan as a place where she had experienced organised audio-based abuse.

A further area of online violence that is delivered 'below the radar' concerns subtle and coded online violence. In Sri Lanka, for example, discriminatory speech is often conveyed with 'humour', via memes and cartoons, in order to avoid moderation triggers for removal. Similarly, captions are often placed over images, to avoid text-based detection.

Freelance journalist Thulasi Muttulingam, who created a Facebook page called Humans of Northern Sri Lanka to document people's personal stories of war, said that social media perpetrators are "cunning" and use passive rather than active threats to avoid moderation. "I hope you get hit by a bus and die. I hope a white van picks you up and you're tortured and raped and murdered, all that sort of nonsense," she said. Another example is a death threat sent to journalist Maria Ressa in a tweet on February 21st, 2021 (Posetti, Maynard, & Bontcheva, 2021) where the text of the tweet itself was not abusive, and the text in the accompanying

image is not easy to process by automated tools, but taken together the menacing underlying meaning is clear to human readers.

Other forms of more subtle abuse include body-shaming and framing women journalists as 'controversial' or 'divisive' figures as well as drawing attention to their employers in the hope of getting them fired. Nigerian documentarian Ruona Meyer, whose trolls targeted the BBC following the broadcast of a programme she made, said: "Talking about somebody's body parts isn't necessarily hate speech, but it's devastating." She feels Twitter's reporting processes lack nuance and therefore fail to recognise and respond to less overt forms of abuse. Twitter's Nick Pickles affirmed that online abuse tactics have morphed significantly with some forms of attack being harder to combat, such as when the attacks are designed to undermine credibility or trust in critical reporting. He added that "this kind of far more pernicious, subversive attack" on the credibility of journalism had been "particularly challenging when we've had public figures making these sorts of statements".

In Northern Ireland, when Patricia Devlin received an indirect death threat referencing an investigative journalist from her now former newspaper (*Sunday World*) who was assassinated in 2001 (RSF, 2020a),[207] the contextual nuance contained in the threat, along with the limited press freedom and journalism safety expertise at the platform, meant that it was not detected nor understood as a death threat until Devlin was able to get this account taken down. Lawyer Caoilfhionn Gallagher KC noted a similar incident involving the BBC Persian service journalists she represents (OHCHR, 2020f, BBC, 2021c), saying that the companies "...simply don't spot it when it's a pattern of threatening behaviour against women - a pattern of threatening behaviour which involves making subtle references to other cases. They don't get picked up at all... The key thing here is prevention is better than cure".

Another hard-to-address case is that of discriminatory abuse served through emojis with racist connotations, such as the monkey emoji which is weaponised in racist posts (MacInnes, 2021). Detecting contextualised racist cases automatically is currently beyond the platforms' technological capabilities. However, Twitter's Nick Pickles highlighted that reporting such contextual cases is helping them to improve their content moderation guidelines: "The challenge you've got is basically the scale of emojis on the platform. If you had some technology that surfaced every use of the monkey emoji, for example, you're going to get a huge amount of false positives. And so that's where I think the technology still has some refinement to do".

This shift from explicit threats and abuse towards more subtle kinds of abuse like networked gaslighting was also apparent in the findings of the big data analyses focused on Maria Ressa and Carole Cadwalladr, in Chapter 3.0 snapshots from which were also published in a 2021 UNESCO discussion paper. Long-range attacks are also associated with these less overt forms of online violence. However, sharpening anti-abuse policies to capture more subtle forms of online violence can risk catching strong opinions and legitimate critique in the net. Twitter's Nick Pickles: "One of the biggest challenges for any company in this space is trying to

---

207 It read: 'You're going to end up like Marty O'Hagan'. Since the assasination of Martin O'Hagan near Belfast in 2001, the first killing of a journalist in the line of duty in the UK was Lyra McKee in Derry in 2019. No one has ever been convicted of O'Hagan's murder (RSF 2020a).

separate out harassment from very strong opinions. And that is something which I think, given the sort of polarisation we see around the world and the political controversy we see, has also made that job harder".

Gap 3: Need to address the cross-platform nature of online abuse

The companies' policies and processes also fail to account for the cross-platform nature of online abuse experienced by women journalists. The BBC's Marianna Spring described coordinated campaigns of abuse and harassment that begin on YouTube in comments that trigger escalating abuse, which then spread to other social media platforms: "It's always cross-platform...that YouTube link is sent to me on Instagram and on Facebook and on Twitter and [it floods] my mentions. It's never limited to a single platform and this ecosystem is very open".

In the US, former *New York Times* reporter Taylor Lorenz found that a one year-long episode of online violence she experienced may have been fuelled by tweeting a story from The Verge about a tech company CEO. A high-profile investor then incited harassment against Lorenz on the audio app Clubhouse, as well as other platforms:

> *Once they...found me, they've never let go. And it's like all of these far-right actors have aligned themselves with the worst people in Silicon Valley. And then it's just been a year of harassment and abuse, on every platform...Instagram, TikTok. Twitter, Facebook Messenger... They were calling me and sending me crazy messages and stuff... Thousands of emails...thousands of direct messages.*

Brandy Zadrozny of NBC News said she was "a big topic of discussion" in "white nationalist spaces" and it was mentally exhausting to monitor all the sites where she was being discussed and threatened. Julia Carrie Wong, Senior Reporter at Guardian US, described being targeted by 'Q', the reputed figurehead of the QAnon conspiracy (Wong and Collins, 2018), who "specifically sent people to my Twitter account in a 'Q-drop'" which was "overwhelming and unpleasant".[208] When Serbian journalist Jovana Gligorijević from *Vreme* tries to take 'mental breaks' from Twitter, abusive tweets are uploaded as screenshots on Facebook which she is alerted to via that platform's facial recognition technology. Brazilian researcher Claudia Lago described the manifestation of the cross-platform trend in Brazil: "Online attacks originate as fake news in the underworld of Whatsapp groups; when they reach Twitter they have already done real damage".

Another shortcoming is the predominant focus on responses by the "big four" online platforms (Facebook, Twitter, Instagram, and YouTube) which overlooks abuse targeting women journalists through other platforms, networks and apps. One example is a first-person blog post written about the BBC's Marianna Spring in

---

208 QAnon is an internet conspiracy theory which began in October 2017 on 4chan claiming politician Hillary Clinton would be arrested. Q is the pseudonym of the anonymous posters and "a motivating factor for many of the insurrectionists who attacked the US Capitol on 6 January 2021". Twitter banned 70,000 related accounts after the event (Wong, 2021).

her 'voice' on a fringe blog, which attracted around 3,000 unmoderated comments, including rape threats and other highly sexualised threats. And at the time of her interview for this study in March 2021, Taylor Lorenz, then still with *The New York Times*, received an alert of a "disgusting hate comment" in the comments section of her Substack newsletter.

Individually, the platforms have much more detailed information, compared with independent researchers and the targets of gendered online violence themselves, as to the origins of abuse (e.g., IP addresses from which a given user has posted abusive messages), possible coordination between users during pile-on abuse, and the deployment of 'sock puppets' and bots in orchestrated attacks. Collectively, they could work very powerfully to combat online violence incidents in real time, through information sharing and collaborative responses to coordinated attacks. They could also work jointly on developing more effective and gender-sensitive policies and tools to improve reporting of cross-platform instances of online violence against women journalists and human rights defenders.

Gap 4: Platform policies need to cover abuse from prominent political figures

A question raised by several interviewees, also evidenced in the companion big data case studies, is whether it is justifiable to exempt politicians and other political actors from content moderation policies and actions, given that some are prominent perpetrators of online threats, abuse and harassment.[209] Both women journalists and the platforms have acknowledged that in some cases prominent political figures (some of whom have had their social media accounts suspended for periods of time), and influencers play a major role as instigators of online violence. However, remedial actions from the social media companies in such cases tend to be more lenient due to policy exemptions on the grounds of "newsworthiness" and prominence (Reuters, 2021).

Based on her own experience, the BBC's Marianna Spring suggested that companies should better honour their duty of care to users by: "...breaking down these ecosystems, which generally have a few central figures at their heart, and as we've seen multiple times it tends to be pretty effective when those people disappear". Former UN Special Rapporteur David Kaye said: "It really does go back to the platforms and their treatment of political figures who engage in this generic kind of incitement against journalists. The platforms should consider that as they take action either to protect or promote political pages like Trump's."

In response to the recommendations of its own Oversight Board, in June 2021 Facebook announced changes to its policies and committed to acting quickly on posts by influential users which may lead to harm (Facebook, 2021b; Posetti and Bontcheva, 2021). As this policy change was still new at the time of writing, it was not possible to independently evaluate the extent to which it could help. However, it is important that all platforms adopt such policies and enforce them in a consistent manner.

---

209 As of June 6 2021, the debate as to whether Facebook will end its policy of exempting political figures from content moderation rules was still ongoing, despite reports of an imminent change: https://www.theverge.com/2021/6/3/22474738/facebook-ending-political-figure-exemption-moderation-policy

Gap 5: Failures of algorithms for content recommendation and moderation

Algorithms for prioritising and recommending content, users and groups have been found to promote misogynistic hate.

Researchers have identified, for example, deficiencies in the metrics used by some companies to prioritise content: "...the top performing domains were those that surfaced in users' feeds over and over—including some highly partisan, polarising sites that effectively bombarded some Facebook users with content" (Faife, 2021). In 2020, *the Wall Street Journal* revealed a 2016 internal company presentation by Facebook researchers that stated "64% of all extremist group joins are due to our recommendation tools", showing that most of the activity came from the platform's "Groups You Should Join" and "Discover" algorithms. "Our recommendation systems grow the problem," according to the presentation.[210]

Due to the lack of transparency, it is hard for external researchers to establish the full role of these algorithms in online abuse. Therefore, a BBC documentary in 2021 reported by one of this study's research subjects, Marianna Spring, created a fake troll persona on YouTube, Facebook, Instagram, Twitter and TikTok to test algorithmic referrals (Spring, 2021). "Barry" - the invented persona - was interested in anti-vaccination and conspiracy content and "he" initially engaged with only a small amount of misogynistic content. After two weeks of following recommendations on each of the platforms, the top recommended pages to follow on both Facebook and Instagram were almost all misogynistic, whereas TikTok suggested no anti-women content and Twitter and YouTube only a small amount (ibid.). The fact that a small scale and brief study can elicit such striking results points to ongoing problems on some platforms, potentially also exacerbated by insufficient moderation of misogynistic content. The lack of independent access to companies' data means that large scale, multi-platform, longitudinal, independent studies that can monitor, evaluate, and compare effectiveness of platform algorithms and approaches is also a gap that needs addressing.

Deficiencies in content moderation algorithms include a limited ability to detect nuanced online violence against women journalists. But simultaneously, they can also lead to censorship of non-offensive content, due to the algorithms' inability to interpret the wider conversational and cultural context. Journalists in Eastern Europe, Asia, Latin America and the Arab States interviewed for this study reported experiencing algorithmic censorship applied to feminist content. Beyond images, the algorithms for detecting misogynistic and other abusive posts have also been found to censor legitimate conversations, due to lack of local context. One such reported example involved Facebook users being muted, banned, or warned when posting about a well-known UK landmark (Plymouth Hoe), as its name was erroneously flagged as misogynistic.

These cases highlight the crucial importance of a human-in-the-loop approach to content moderation, which is not only sensitive to local cultural and linguistic context, but also complemented by an effective appeals process - including the

64% OF ALL EXTREMIST GROUP 'JOINS' ARE DUE TO OUR **RECOMMENDATIONS TOOLS**

FACEBOOK RESEARCHERS, 2016

---

210 "Facebook Executives Shut Down Efforts to Make the Site Less Divisive. The social-media giant internally studied how it polarizes users, then largely shelved the research": https://www.wsj.com/articles/facebook-knows-it-encourages-division-top-executives-nixed-solutions-11590507499

ability to appeal refusals to remove abusive content - and robust transparency policies. Moreover, as Twitter itself noted in a policy paper: "Content moderation is more than just leave up or take down. Regulation should allow for a range of interventions, while setting clear definitions for categories of content."

Gap 6: Deficiencies in collaboration and multistakeholderism

There is little systemic engagement by the companies with civil society, media, governments and experts in developing policy responses to gendered online violence. Some hopeful signs of collaboration are slowly emerging, but there are still often gaps in stakeholder representation. Interviewees said, for example, that in Mexico, where company representatives meet periodically with government officials and civil society representatives to discuss trends in digital attacks and share good practices, there are only limited exchanges with independent news media and academic experts.

More needs to be done to bridge the industry divides between big tech and news outlets to proactively address the issue of gender-based online violence against journalists, as indicated by Sveriges Radio CEO Cilla Benkö. Reflecting on International News Safety Institute (INSI)-facilitated conversations between (mostly Western) news organisations and the platforms about addressing the online abuse of women journalists, she said: "It is good that as industries we communicate with each other in an open and constructive manner. But much more needs to be done" (Benkö, 2021). Exploring the "more" could lead to country-level facilitated workshops designed to improve understanding and trigger more effective responses.

At the level of process, Facebook has established a 'Trusted Partners' programme to provide more context to the flow of requests to flag, block or remove content and people on its platform or on WhatsApp. This programme has been operational in select countries in Asia, the Arab States, Latin America and Europe according to research for this study, and parallel research (Sinpeng *et al.*, 2021). The programme offered journalists and human rights defenders the option to escalate complaints to the platforms via civil society organisations and provided more support to help counter the automatic decisions being made by moderation software. But there is scant public information about this programme and awareness is low, particularly in the Asia Pacific (ibid). Twitter also provides a one-on-one service via their Latin America Policy team, in partnership with ABRAJI (Associação Brasileira de Jornalismo Investigativo, also known as Brazilian Association of Investigative Journalism), ARTICLE 19 and Reporters without Borders (RSF), with a fast-track channel to prioritise complaints from these organisations. However, according to interviewees, many cases fall through the cracks, and freelance women journalists along with others who lack the contacts, are at a disadvantage.

Facebook told researchers in an email that the Digital Security Helpline of the civil society group Access Now (see Access Now, n.d.) has high priority access to their systems. However, there was limited awareness among the research participants about this service. For example, although the service has a Tunis-based contact point, according to Faouzia Ghiloufi - executive member of the National Syndicate
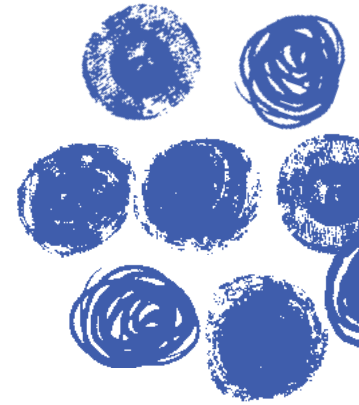
of Tunisian Journalists and Vice Chair of IFJ Gender Council - Tunisian civil society organisations do not have direct contact with social media platforms, and that the onus falls on the individual journalists to reach out and report to the social media platforms. While civil society contact points for rapid escalation in cases of gendered online violence targeting journalists could be valuable, this does not exempt the companies from having staffing capacity to cover local languages and contextual understanding.

iv. The need to move fast and fix things

Most interviewees assessed that the companies' enabling role in gender-based online violence against journalists could not be properly addressed until the business models and technical design were overhauled to prioritise safety and human rights over profit. The dominant players set up their highly profitable business models to maximise traffic and engagement, rather than to protect journalists, human rights defenders, democracy, or marginalised communities. Therefore, these models favour inflammation rather than accurate information and this principle is embedded in the design decisions of the platforms.[211] This underpins the view of Dr. Michelle Ferrier, executive director, Media Innovation Collaboratory, that the companies are complicit in maintaining an environment of tension that serves business purposes and also reinforces patriarchal norms in digital spaces.

A core problem is reliance on attention, not least that centred around harassing or abusive engagements, which grows market share and boosts already astronomical profits. This can be damaging to women journalists in particular, even if legal. South African investigative journalist Qaanitah Hunter linked the platforms' business models, which "incentivise and reward trolling", to their reluctance to effectively combat gendered online violence. "They have no interest in getting hate off [their platforms]," her compatriot, editor Ferial Haffajee said. ARTICLE 19's Legal Officer Paulina Gutiérrez said "the business model is a problem in itself" and that it enables attention-driven pile-ons against women journalists and many other groups at risk and those facing discrimination.

The scale of the gender-based online violence problem and its networked nature mean that it can go viral very easily and quickly, with wide-ranging impacts including on democratic deliberation and digital citizenship more broadly. These developments, along with UN-level acknowledgement of Facebook's "determining role" in human rights abuses against the Rohingya in Myanmar (Miles, 2018; McPherson, 2020), and its facilitative function in the 2021 Capitol Hill insurrection (Silverman et al., 2021; Zuckerberg, 2021; Reuters, 2021) highlight the urgency for action.

A report by civil society monitoring organisation, GLAAD, titled the "Social Media Safety Index" has assessed how companies detect abuse against LGBTQ users, and argues that "bad actors" have learned how to game AI systems (GLAAD, 2021).

---

211 One example is a recent study of the role of highly active, abusive users in shaping content recommendation on U.S. Facebook pages https://www.theatlantic.com/technology/archive/2022/02/facebook-hate-speech-misinformation-superusers/621617/

This is one reason for the NGO urging the "need for human moderation — as well as a corresponding need for ethical and responsible employment practices in relation to these workers". According to GLAAD, Reddit conducted a study on hate and abuse on its services and made these findings public – and recommended that Facebook do likewise. In addition, GLAAD called on Facebook to allow an "independent audit, specifically focused on its lack of transparency" (ibid.).

Guardian US' Julia Carrie Wong pointed to the need for an overhaul of Facebook's ethical and normative frameworks - led from the top:

> *I don't necessarily need Facebook to provide me with personal on-platform protection, but my society is being damaged by what's happening on Facebook...the QAnon movement was allowed to grow so big on Facebook with Facebook support, rather than with Facebook trying to limit it... it's a lack of being willing to make moral judgments that would protect fundamental features of a liberal democracy.*

However, former HuffPost UK Editor-in-Chief[212] Jess Brammar said that expecting the platforms to not just 'clean house', but rebuild the infrastructure was a futile wish: "I feel like it's too late to come from the platforms... We are years into this now". She called for broader social and political reform to address the problems at the root. "[I]t lets off the people who have got us into this situation...if we just look at the platforms."

Conclusion

While structural sexism and misogyny, populist political instigators and partisan news media contribute to gendered online violence, the platforms bear a major responsibility for enabling and facilitating the problem – and for addressing it. For women journalists to be able to work safely online, the policy gaps identified must be addressed. Business models and algorithms must be restructured and redesigned. And more effective and comprehensive tools and protocols for detection, reporting, moderation and countering of online attacks on journalists are required. Additionally, there is a strong need for transparency; for independently defined and evaluated measures of the effectiveness of abuse countermeasures. **It is time to move away from the current approach of largely ineffective self-regulation.**

---

212 At the time of the interview Brammar was Editor-in-Chief of HuffPost UK. In September 2021 she was appointed the BBC's executive news editor of news channels.- a move which led to further pile-ons: https://www.pressgazette.co.uk/bbc-boss-fears-jess-brammar-effect-will-affect-hiring-of-journalists-with-diversity-of-views/
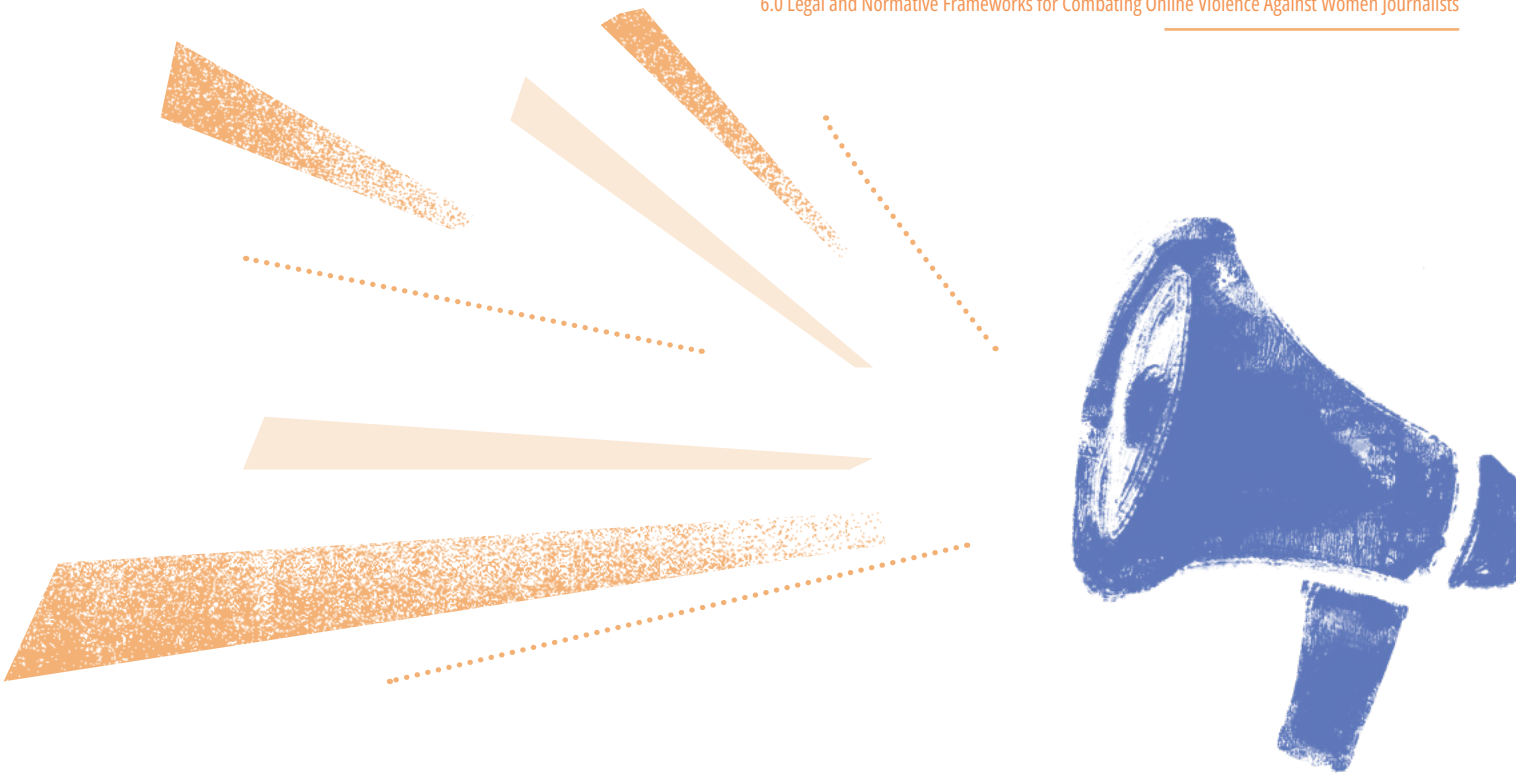
6

Legal and Normative Frameworks for Combating Online Violence Against Women Journalists

Angelique Lu, Julie Posetti and Nabeelah Shabbir

The right to freedom of expression, with its corollaries of press freedom and freedom of access to information, is protected by Article 19(2) of the International Covenant on Civil and Political Rights (ICCPR) (OHCHR, 1976), and Article 19 of the Universal Declaration of Human Rights (UNDR) (UN, 1948). These articles require States to guarantee the right to seek, receive and disseminate information for citizens generally and, by extension, journalists and news publishers who benefit from press freedom protections. In 2011, the UN Human Rights Committee recognised that these rights also included "electronic and internet-based modes of expression" and called on Governments to protect any attack on an individual's right to their freedom of expression (OHCHR, 2011a). Since 2016, the UN has made it clear that "...the same rights that people have offline must also be protected online" (A/RES/71/199) (UN GA, 2016; UN GA, 2018b).

The UN Human Rights Council and the UN General Assembly have also recognised that women, and especially women journalists, are disproportionately affected by online violence, acknowledging that they are particularly exposed through intersectional factors such as race, ethnicity, sexual orientation and age (OHCHR, 2019; UN GA, 2017d; UN GA, 2017e; UN GA, 2018b; UN GA, 2020b; UN GA, 2021b).

In this chapter, international, regional and State-level legal and normative frameworks for responding to online violence against women journalists are examined, while exemplar judgements are catalogued, and gaps in law enforcement are highlighted. Here, insights gleaned from 184 in-depth interviews, and responses to the relevant survey questions in the main data corpus are supplemented by relevant examples from other countries, surfaced through extensive desk research.

Additionally, the 15 country case studies underpinning the broader study are drawn on to contextualise the discussion.

i. An assessment of relevant UN-level responses

The UN's specific measures to address the issue of violence against women journalists

The United Nations' legal and normative frameworks addressing violence against women journalists have grown and evolved over the decades. UNESCO, the UN organisation that commissioned this research, has a mandate to protect freedom of expression. It also plays a central role in the UN's work to improve the safety of women journalists, spearheading the implementation of the UN Plan of Action on the Safety of Journalists and the Issue of Impunity. The UN Plan of Action was developed in 2012 to coordinate the efforts of actors - both within and outside the UN system - in promoting a safer environment for journalists worldwide. In its 10th anniversary year, the Plan is under review to render it fit for purpose in the Digital Age.[213]

A 2016 Human Rights Council Resolution (33/2) on the Safety of Journalists specifies that States must prosecute attacks of all kinds, including gender-specific attacks, create protective measures for journalists, facilitate independent investigations, and ensure victims have access to appropriate remedies.[214] Additionally, the resolution explicitly refers to the specific threats faced by women journalists, and the need to take a "gender-sensitive" approach when considering mechanisms to improve safety.

In 2017, the UN Secretary-General addressed the problem of online violence in his report to the UN General Assembly on 'The Safety of Journalists and the Issue of Impunity' (A 72/290). Shortly afterwards, the UN General Assembly adopted resolution (A/C.3/72/L.35/Rev.1) (UN GA, 2017e) on the safety of journalists with a particular gender focus, "condemning unequivocally" all "specific attacks on women journalists in the exercise of their work, including sexual and gender-based discrimination and violence, intimidation and harassment, online and offline."

The UN General Assembly passed another Resolution on the Safety of Journalists (A/RES/74/157) in 2019 that again condemned attacks on women journalists online and offline, including sexual harassment, intimidation and incitement to hatred. It also called upon Stated to "...tackle these issues as part of broader efforts to promote and protect the human rights of women, eliminate gender inequality and tackle gender-based stereotypes in society".

In 2020, the UN Special Rapporteur on violence against women, Dubravka Šimonović, submitted a thematic report on the issue of violence against female journalists to the UN Human Rights Council (UN GA, 2020a). The report found that

---

213 See this 2022 report from the UN Special Rapporteur for the right to freedom of expression for an assessment of digital era gaps in the UN Plan of Action on the Safety of Journalists and the issue of impunity: https://documents-dds-ny.un.org/doc/UNDOC/GEN/G22/323/44/PDF/G2232344.pdf?OpenElement=
214 "Resolution adopted by the Human Rights Council on 29 September 2016: the safety of journalists": https://documents-dds-ny.un.org/doc/UNDOC/GEN/G16/226/24/PDF/G1622624.pdf?OpenElement

the rise of the internet also coincided with an increase in online attacks against women journalists, disproportionate to those faced by their male colleagues, as well as a rise in online violence such as doxxing, 'sextortion' and 'trolling'.

It also made a series of recommendations to States, including prohibiting and criminalising gender-based online violence against women journalists, creating investigative units or independent commissions to investigate these issues, as well as developing training protocols for law enforcement to better prosecute cases.

Also in 2020, the UN Human Rights Council passed a Resolution on the Safety of Journalists (A/HRC/45/L.42/Rev.1) which calls on States to:

> *... take measures to prevent sexual harassment and other forms of sexual and gender-based violence, including threats of rape, intimidation and harassment against women journalists, to encourage the reporting of harassment or violence by providing gender-sensitive investigative procedures...and to prohibit incitement to hatred against women journalists, online and offline, and other forms of abuse and harassment through relevant policy and legal measures that comply with international human rights law* (UN GA, 2020b).

Then, in mid-2021, the Human Rights Council passed a Resolution (UN GA, 2021b) that condemned unequivocally:

> *...online attacks against women and girls, including sexual and gender-based violence and abuse of women, in particular where women journalists, media workers, public officials or others engaging in public debate are targeted for their expression, and calls for gender-sensitive responses that take into account the particular forms of online discrimination.*

States, therefore, can be understood to have an obligation to prevent and stop online violence against women journalists, human rights defenders, and other public figures under their broader duty to protect freedom of expression (including press freedom), and end discrimination and violence against women and girls - online as well as offline.

However, States may limit or restrict freedom of expression if a three-part test is satisfied, in line with the provisions of Article 19 of the ICCPR, as explained in General Comment 34 (ARTICLE 19, 2020a):

- **The restrictions must be "provided by law": any restrictions on the right to freedom of expression must be explicitly drafted, to enable people to adjust their conduct appropriately;**

- **The restrictions must pursue a specific "legitimate aim": any restrictions must have the purpose of protecting the rights and reputations of others; and**

- **The restrictions must be "necessary and proportionate" to its intended goal.**

In this context, freedom of expression is not an absolute right. On the one hand, this means that freedom of speech defences cannot be used to justify abuses of the rights of others, and neither can they be used to fend off justifiable restrictions by a State acting within the international standards for legitimate limitations on expression.

Nor can 'freedom of speech' be used to excuse failing to act against online violence by those private actors whose facilities and platforms are exploited by attackers. The notion that a person's right to 'free speech' therefore entitles them to undercut another person's right to freedom of expression (including press freedom) is contrary to international standards on freedom of expression.[215]

On the other hand, as former UN Special Rapporteur for freedom of expression David Kaye warned in a 2017 report, any attempts by States to stop and prevent gendered online violence must also avoid censorship: "Censorship and undue restrictions on content could end up undermining the rights of the very women for whom governments and corporate actors may seek to provide redress" (OHCHR, 2017a).

In other words: at the UN level, countering online violence against women journalists while respecting freedom of expression is a 'balancing act' (Bontcheva and Posetti, 2020). But it would be a false binary argument to suggest that it is not possible to both defend freedom of expression while also working to prevent and stop online violence against women journalists - a form of attack which is ultimately designed to chill their reporting and undercut press freedom.[216]

UN-based frameworks to protect women and girls online

Women journalists experiencing online violence can also look to UN-level protections more broadly enshrining the rights of women and girls online and offline. Women's rights are protected generally under a number of other UN articles. For example, Article 2 of the International Covenant on Civil and Political Rights (ICCPR) requires states to guarantee human rights to all people "without distinction of any kind", including gender and sex. Further, the Convention on the Elimination of all Forms of Discrimination Against Women (CEDAW) reates specific obligations to end discrimination based on gender and sex characteristics that would restrict a woman's human rights (UN Treaty Collection, 1979).

In 2013, the UN Commission on the Status of Women (CSW) called on States to develop mechanisms to combat violence against women online (UNCSW, 2013). The General Assembly went further later the same year, recognising that female human rights defenders were at risk of violence both online and offline by State and non-state actors, calling on States to bring perpetrators to justice (UN GA, 2013). The Human Rights Council confirmed that domestic violence could include acts such as cyberbullying and cyberstalking in 2015 (OHCHR, 2015).

---

215 For a detailed discussion regarding the limits to freedom of expression, please see: https://unesdoc.unesco.org/ark:/48223/pf0000378755?posln-Set=4&queryId=30ca5706-029b-4911-b525-e067c8e66a87
216 See the 25-point plan for States to respond to online violence against women journalists while respecting freedom of expression rights in Chapter 7.0 (Conclusion and Recommendations).
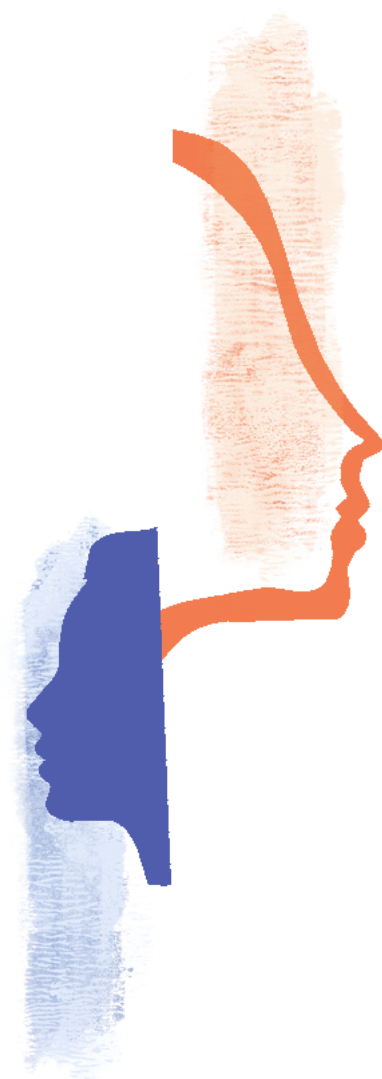
In 2017, the Special Rapporteur on Violence Against Women, Dubravka Šimonović, presented a report to the General Assembly, in which she suggested the formulation of a new legal framework for addressing violence against women (UN GA, 2017e). In the same year, the UN Committee for the Elimination of Discrimination Against Women (CEDAW) also recognised that gender-based violence occurs in technology-enabled settings (ARTICLE 19, 2020a).

The following year, the UN Human Rights Council (HRC) commissioned the aforementioned thematic report into gendered online violence (UNGA, 2018a). That report began by looking at legal issues linked to the terms "Information and Communications Technology (ICT)-facilitated violence against women" and "online violence against women" (ibid.). It defines the latter term as extending "to any act of gender-based violence against women that is committed, assisted or aggravated in part or fully by the use of ICT…because she is a woman, or affects women disproportionately" (ibid.). The report concluded that the internet is being used in a broader environment of widespread and systemic structural discrimination and gender-based violence against women and girls, and it drew a number of other relevant conclusions:

- **International human rights law and other UN-related instruments pertaining to women in public and private life are fully applicable in digital spaces;**

- **Any legal and policy measures used to eradicate online gender-based violence should be framed within the broader framework of human rights (i.e. encompassing the right to freedom of expression);**

- **States should enact new laws and measures to prohibit new and emerging forms of online gender-based violence.**

Possible solutions proposed included:
- Increased education and training on the issue of online abuse and violence;
- Lobbying technology companies to develop mechanisms that allow individuals to control and define their online experience. These include tools that permit them to block specific individuals, control their privacy, or tailor their interaction to protect themselves against abusive behaviour;
- Increased funding for research into the scale of the issue (UN GA, 2018a).

Further, in 2018, the UN Human Rights Council adopted Resolution 38/5 on accelerating efforts to eliminate violence against women and girls in digital contexts (OHCHR, 2018c). The resolution acknowledged the effects of gender-based violence on the participation of women in the digital realm, and the obligations of States to prevent and protect such abuses from occurring, as well as highlighting the role that businesses need to play in addressing the issue.

In 2019, launching the UN Strategy and Plan of Action on Hate Speech (UN, 2019), Mr Gutteres said: "Around the world, we see a groundswell of xenophobia, racism and intolerance, violent misogyny, antisemitism and anti-Muslim hatred" (UN News,

2019). The Plan acknowledges the exploitation of social media as platforms for bigotry, which see public discourse weaponised for political gain by stigmatising and dehumanising women and other targets such as minorities and refugees.

At a normative level, therefore, the UN has increasingly recognised the problem of online violence against women journalists, putting it forward as a matter of serious concern and calling for action in the international arena.

ii. Third party intermediary legal obligations

There are also a number of relevant human rights instruments relating to the human rights obligations of third-party platforms, such as Facebook, Google, TikTok and Twitter. As discussed in the previous chapter, social media companies have an obligation to protect human rights as set out by the UN 'Ruggie' Guiding Principles on Business and Human Rights (OHCHR, 2011b).[217] These principles require companies to respect the rights established in core human rights treaties, including women's rights and the right to freedom of expression, press freedom and the safety of journalists. In particular, businesses must avoid violating human rights or facilitating human rights violations, and remedy them if they occur. They must also seek to prevent and mitigate any issues linked to their operations, products or services (ARTICLE 19, 2020a).

There is a question of the extent to which States have made internet companies legally liable for third party content that violates human rights. Given the debate and the diversity of jurisdictions on this issue, a number of different approaches have also been developed. In the US, where there is no legal liability for third party content, the Federal Trade Commission fined Facebook $5billion for consumer privacy breaches in 2019 (FTC, 2019). In France, Reporters Without Borders (RSF) instigated a consumer law case against Facebook in March 2021. RSF alleges that Facebook is guilty of "deceptive commercial practices" on the grounds that the company's promises to provide a "safe" and "error-free" online environment are "largely mendacious", and "that it allows disinformation and hate speech to flourish on its network (hatred in general, and hatred against journalists), contrary to the claims made in its terms of service and through its ads" (RSF, 2021p).

iii. Regional legal frameworks and instruments

Various regional bodies have recognised the generic need for protection of human rights on the internet companies' services. For example, in 2013, the Special Rapporteur for Freedom of Expression of the Inter-American Commission on Human Rights (IACHR) recommended that private actors establish services that are consistent with human rights laws (Botero Marino, 2013).

Other recommendations included the publication of transparency reports about government requests for user data or content removal. It was also recommended

---

217 See also Chapter 5.0, "Platforms and vectors: Assessing big tech responses to online violence".

that efforts be made to notify individuals who may have their rights violated, granting them access to non-judicial remedies and creating protective measures and business practices consistent with human rights protections (ARTICLE 19, 2020b).

Similar recommendations were made by the subsequent IACHR Special Rapporteur in 2017 in his report 'Standards for a Free, Open and Inclusive Internet,' which noted "the relevant policies and practices must be based on respecting and guaranteeing human rights" (Lanza, 2017).

Towards the end of 2021, the Inter-American Court of Human Rights found the State of Colombia responsible for the kidnapping, torture and rape of El Espectador journalist and former UNESCO World Press Freedom Prize winner Jineth Bedoya Lima by far-right militia (AFP, 2021), in a judgement which referenced research for this study to highlight the online-offline nexus (IACHR, 2021b).

The specific need to recognise and address gendered online violence against women journalists has also been examined by regional bodies such as the Organization for Security and Cooperation in Europe (OSCE), the Council of Europe (CoE), the European Court of Human Rights, the European Commission and the IACHR. Each of these has explored similar themes about the safety of women journalists (ARTICLE 19, 2020a).

Article 17 of the Council of Europe Convention on preventing and combatting violence against women and domestic violence (Istanbul Convention) requires Member States to actively encourage the private sector and the news media to help prevent violence against women (Council of Europe, 2011). In 2016, the Council of Europe adopted a recommendation on the safety of journalists, which recognised that women journalists faced gender-specific threats which were increasingly taking place online. The guidelines called on States to "take appropriate preventive operational measures'', such as police protection, taking into account "gender-specific dangers" (Council of Europe, 2016).

The Organization for Security and Cooperation in Europe (OSCE) has also been significantly involved in developing norms to protect women journalists from online violence. In 2016, the OSCE Representative on Freedom of the Media recommended that States recognise threats and harassment of female journalists as an attack on freedom of expression as part of its Safety of Female Journalists Online project (OSCE, 2016b). The report suggested a number of mechanisms to address the issue, including the improved training and strengthened capacity of law enforcement, as well as the collection of data to determine the extent of the issue. OSCE's work on this issue continues, with a 2020 study making a series of concrete and tailored recommendations for action, including within legal and judicial arenas, directed at both State and non-State actors (Chocarro et al., 2020).

The nature of the internet means that any person across the world can engage in online violence. However, in the landmark 'right to be forgotten' case against Google, the Court of Justice of the European Union, found that Google was bound by EU laws. The court ruled that even if physical servers were outside of the EU, laws applied to search engine operators if they have branches in a Member State (EUR-LEX, 2014).

In late 2021, the European Parliament voted in favour of stronger regulatory responses to the problem of harms such as online violence and the spread of disinformation as part of the proposed Digital Services Act (DSA) which focuses on content moderation and aims to "create a safer digital space in which the fundamental rights of all users of digital services are protected" (European Commission, 2021a). It is "building a new framework, so that what is illegal offline is also illegal online" (European Parliament, 2021). The potential extra-EU dimensions of this initiative will become evident over time.

Under Articles Two and Three of the Council of Europe's European Convention on Human Rights, there is an obligation to conduct an investigation when there is a death of a journalist which may implicate a government, or where the State may have had a protective obligation with which it failed to comply. Article Three - the right to freedom from inhuman and degrading treatment - could also extend to many acts of gender-based online violence. However, this concept is not yet well developed at the international level.

iv. Practical legal challenges experienced by women journalists under fire

Traditional legal routes are often costly, emotionally taxing, convoluted and ill-equipped to provide adequate redress for the targets of gender-based online violence. Tort law-based suits[218] could potentially provide some form of remedy for victims, however there are limitations. Legal routes such as privacy torts, copyright, or even laws that were written for analogue communications methods like telephone and mail are being tested, but they are often deemed to be unfit for purpose in regard to online violence against women journalists, both domestically and across jurisdictions.

Tort actions require being able to identify perpetrators, while laws criminalising harassment and stalking usually require proximity and repetition - making legal remedy in a case involving a 'pile on' instigated by a perpetrator who posts a one-off threatening comment from the other side of the world very challenging. Such laws are also often interpreted as applying only to the physical realm.

The jurisdictional issues that exist in criminal law are also found in civil actions. Secondly, civil law proceedings may bring more attention to a matter, which can deter complainants. Thirdly, a victim may not be seeking an award of damages, but rather an injunction to remove content or de-platform a perpetrator, meaning that conventional remedies in civil law may not be appropriate. Finally, civil law remedies may only be used after the event has occurred: they cannot prevent the publication of abusive, threatening or harassing content at the outset (ibid), limiting their role as preventive measures, although successful prosecutions and verdicts may be effective deterrents in some arenas.

Research conducted for this study identified the following **12 legal, judicial and law enforcement factors presenting challenges to, and potential opportunities for, legal remedy:**

---

218 Tort law is the branch of law that imposes civil liability for breach of obligations imposed by law e.g. negligence cases. See: https://uk.practicallaw.thomson-reuters.com/6-107-7397?transitionType=Default&contextData=(sc.Default)&firstPage=true

1. Platform accountability

The internet services that primarily enable online violence are legally protected from liability for content in most instances, providing very little incentive for them to police the behaviour of users or recalibrate their algorithms in ways that prioritise human rights over profit. Nevertheless, companies like Facebook, Google and Twitter have come under increasing pressure over their role in facilitating and enabling the online harassment and abuse of women journalists. However, due to being based in the US, these companies are currently shielded by Section 230 of the Communications Decency Act (CDA) which generally protects them from liability for the speech of others published on their platforms, even if said speech is found to be illegal (Lipton, 2011).

Section 230 protects companies from tortious actions involving their users. It was passed in 1996, and has been lauded as "the most important law protecting internet speech" (Harvard Law Review, 2018). Some commentators credit this law for allowing the internet to thrive. However, many researchers have noted that the law was introduced before the advent of the social web, and it actively prevents victims of gendered online violence from pursuing legal action against the internet companies for the actions of their users.

In 2015, journalist Alex Chu wrote a public appeal to the then US President Barack Obama to repeal Section 230. It continues to resonate:

> *Right now you can't sue digital platforms for enabling harassment on their services, even if they enable harassment through flagrant, wilful neglect. If your harasser is able to take fairly basic steps to keep himself anonymous — and if the platform he chooses enables and enforces that anonymity — then there is literally nothing you or the government can do, even if his actions rise to the level of major crimes like attempted murder. Closing this loophole wouldn't require giving the internet "special treatment" compared to other forms of communication. Nor would it require a sudden, major deviation from the standards of free speech most of the developed world respects. It would require the exact opposite — it would require the United States to remove a law that specifically mandates special treatment for internet service providers and platforms that no other communications medium has...* (Chu, 2015).

Some scholars have suggested including a take-down provision in the CDA, similar to the safe harbour provision found in Section 12 of the Digital Millennium Copyright Act (DMCA), which protects services that inadvertently host material that infringes copyright (MacAllister, 2017). A number of interviewees for this

study also pointed to Germany's Network Enforcement Act as a potential model for dealing with the issue.

In a 2021 report, the Aspen Institute's Commission on Information Disorder also recommended that Section 230 be amended to a) Withdraw platform immunity for content that is promoted through paid advertising and post promotion, and b) Remove immunity as it relates to the implementation of company product features, recommendation engines and design (Aspen Digital, 2021).

However, ARTICLE 19's Legal and Policy Officer Paulina Gutiérrez expressed concern about subjecting the platforms to regulatory systems governing news media as a means of achieving improved responses to online violence. She argued that as internet intermediaries, they should not be regarded as news publishers held responsible for third party content on their sites and apps: "This is a very important freedom of expression principle that we need to uphold even in these cases. We need to start looking at other regulatory models, like what can we ask [of] them that is free speech compliant; privacy compliant". But this view runs counter to that expressed by many international editors and journalists interviewed, who generally argued in favour of the platforms being legally required to fulfil the same obligations met by publishers.

Blowing the whistle

Whistleblowers exposing Facebook's failure to deal with predictable harms on its platforms have also increasingly called for reforms to the CDA. In October 2021, former Facebook product manager Frances Haugen disclosed a cache of internal company documents and urged the US Congress to amend Section 230 with the aim of regulating companies to redesign algorithms that emphasise engagement and inflame hate, as distinct from their liability over user-generated content.

In late 2021, Facebook shareholders resolved to commission an independent human rights impact assessment on the targeted advertising systems that drive the company's profits. The resolution cited data from Ranking Digital Rights' Index on the company's lack of transparency around how it enforces its advertising policies (Investor Alliance for Human Rights, 2021). Facebook derives 99% of its revenue from advertising, but the company publishes nothing about how many ads it takes down, how effectively it detects (and deters) bad actors, and how often its various ad rules are broken. "Facebook continues to evade transparency on this topic while making every effort to block independent research that aims to unearth more information on how it enforces its rules", tweeted Ranking Digital Rights (RDR, 2021).

Time for news organisations to take legal action against the platforms?

Several of our interviewees expressed scepticism about the prospect of regulatory reform making the platforms accountable for online violence against women journalists, including Inga Thordar, a senior editor at CNN at the time of her

interview. She suggested that a class action lawsuit brought by major international news organisations might be a more effective approach.

Kenyan editor Catherine Gicheru of the Africa Women Journalism Project told the researchers:

> *We need to start taking people to court for their online behaviour because it is affecting our capacity to report freely as female journalists, impeding our right to free expression... The media should test the laws that exist and take these perpetrators to court. I know that many of these cases may be dismissed, but it is an excellent starting point for engaging the police and government to enact or strengthen current anti-harassment laws to recognise harassment in online spaces. As it is, if you report online harassment to the police today, they are frequently at a loss as to how to document your complaint unless you can provide them with physical evidence of the harassment.*

Such legal action is not easy, as demonstrated by Sweden's Cyber Hate Crime Monitor, which pursued domestic legal action to get Facebook to take responsibility for toxic content in closed and 'secret' Facebook groups in 2019. The organisation documented 80 clear examples of criminal statements published within one such group. They submitted a 'legal removal request', with 80 statements attached, asking Facebook to close down the group, or alternatively that the company moderate the group - but Facebook rejected the application. Instead, the company suggested that the Cyber Hate Crime Monitor should go through the various groups identified and report each criminal activity spotted. The organisation then reported Facebook to the police for violating the act on responsibility for electronic bulletin boards, but they did not proceed with an investigation. The Cyber Hate Crime Monitor continues to pursue legal action against Facebook in the case (Cyber Hate Crime Monitor, 2020a; Cyber Hate Crime Monitor, 2020b; Swedish Television, 2020).

2. Identifying the perpetrators

At the individual State level, conventional legal remedies (including law enforcement tools like protective orders) are often dependent on being able to identify the perpetrator, which is difficult when anonymous usernames and VPN devices are designed to prevent identification. When perpetrators are unfamiliar to the victim, or go to great lengths to conceal their identity, significant resources are required to investigate, locate and identify harassers (Fenwick, 2021). Police and law enforcement officials may also have to rely on the cooperation of the platforms to carry out their investigations, which can often be time-consuming, and may not lead to successful identification of perpetrators. These factors may reduce the likelihood of a successful prosecution, which could deter prosecutors from taking on cases involving the harassment of women journalists.

Some proposed legislative responses involve 'real-naming conventions' for social media users to enable easier identification of perpetrators. This raises privacy issues, however, and also could weaken the important function of source confidentiality and anonymity in investigative journalism (Posetti, 2017a). One example of this approach is the Australian legislative push of late 2021 designed to require social media companies to supply details of users who defame or harass others. It will incorporate a complaints mechanism with takedown enforcement powers. "They have created the space and they need to make it safe, and if they won't, we will make them (through) laws such as this," the Australian Prime Minister said (Reuters, 2021c). The proposed legislation comes in response to a ruling by the country's High Court which held news organisations, not Facebook, liable for defamatory third party comments made on a news publisher's post. It also follows defamation action by a senior cabinet minister against a human rights defender over a critical tweet focused on the minister's public statements regarding a high profile rape case (Karp and Remeikis, 2021).

3. Cross-jurisdictional challenges

Jurisdictional issues remain one of the biggest challenges in combating gendered online violence. For example, UK based human rights lwyer Caoilfhionn Gallagher said in an interview that the Council of Europe mechanisms are generally only effective when dealing with Council of Europe Member States. So, when non-Member States are associated with attacks on journalists in Europe - online or offline - there is very little that can be done to hold them to account.

One significant case carried by Gallagher involves online violence being experienced by women journalists at the BBC Persian language service.[219] They are UK residents who are being targeted in an orchestrated online violence campaign that they believe is emanating from Iran, a non-EU Member State, and removing such content from view online only within Europe would have limited effect.

Successful prosecution of gender-based online violence requires cooperation between the platforms, law enforcement agencies, and often multiple jurisdictions (across countries or federal states), which may pose significant challenges. Even if perpetrators can be identified and located, assuming devices like VPNs are not used to hide their IP addresses, they could be inaccessible due to jurisdictional issues which dramatically reduce the likelihood of a successful prosecution against individuals, or US-based social media companies.

Law enforcement across international borders may depend on mutual legal assistance treaties (MLATs) to gain access to evidence and to identify perpetrators through international cooperation (ARTICLE 19, 2020a). Although research shows they are a 'resilient' form of obtaining data, they are rarely used by law enforcement agencies. MLATs can take months, are often costly, require complex administrative legal processes, and the cooperation of other countries and third-party platforms.

---

219 See Chapter 2, "Global Overview: comparative analysis of incidence, impacts and trends" and Chapter 5.3, "Policy gaps" for more about BBC Persian journalists.

Civil society organisations have also raised concerns about possible privacy violations and a lack of transparency about their application (ibid).

The case of Qatar-based journalist Ghada Oueiss [220] is relevant to jurisdictional limitations. She is taking legal action in a US court against residents of two other States, the UAE and Saudi Arabia, in connection with what she alleges is an example of orchestrated online violence by foreign political actors and their agents. The action is being pursued in a Florida court and it names various State actors, State media, and US-based social media users (Shilad, 2021a).[221]

One Twitter account targeting Oueiss grew to 50,000 followers in three weeks, prompting her to take legal action. "I had no choice but to file the lawsuit. That was the only way to respond, because it was becoming more and more vicious," she said. However, the case and others like it, face major technical hurdles, including jurisdictional issues (Clary, 2021).

These cases raise the issue of international legal processes dealing with cross-border harm. For example, in response to the murder of US-resident Jamal Khashoggi in Saudi Arabia's Istanbul Consulate, Reporters Without Borders (RSF) applied to the German courts for remedy for 'crimes against humanity' under the German Code of Crimes against International law (VStGB). German laws offer jurisdiction over core international crimes committed overseas, and "German courts have already shown readiness and willingness to prosecute international criminals" (RSF, 2021p).

Inconsistent laws within a country can also make the prosecution of online violence against women journalists more difficult. For example, all US states criminalise the stalking or harassment of a person, but inconsistent laws relating to the internet have been highlighted as a key impediment to developing a national approach to the problem. Some legal researchers argue that the existing framework of disjointed laws in the US is ill-equipped to deter conduct that crosses state and national borders (Lipton, 2011).

One possible approach is to draft legislation so that any legal actions are based on the jurisdiction where the victim lives, rather than the perpetrator.[222]

4. Criminal Remedies for Harassment, Stalking and Threats

Laws covering online violence are also often drawn from offline stalking and harassment legislation. This may be problematic, given many harassment laws were drafted before the advent of the internet and may not be fit for purpose. Additionally, the internet arguably makes harassment easier to execute, enables it to attain virality and can elevate physical risk.

The absence of clear legal definitions as to exactly what behaviour constitutes actionable digital stalking, harassment or threats is one of the key legal issues

---

220 See case study in Chapter 2.2.1: "Racial vilification and structural racism".
221 Ghada Oueiss' original court filing: https://www.courthousenews.com/wp-content/uploads/2020/12/1-20cv25022-002.pdf
222 See for example: Dow Jones & Company Inc. v Gutnick, Joseph (2002) HCA 56 which held that defamation occurs where the loss to reputation occurs (High Court of Australia, 2002).

affecting the prosecution of online violence against women journalists (ARTICLE 19, 2020a). Even when there is consensus that criminal sanctions should be applied, there are also issues about how the crime should be defined and when the threshold for criminal liability should be reached. Courts have struggled to determine the nature of online threats, or provide guidance on which comments or content are worth prosecuting. One problem is the view that if online violence has not manifested offline, it is not sufficiently serious. There is also the problem of misogynistic speech being protected by free speech laws in the many countries where misogyny is not considered a hate crime, for example.

A complicating factor involved with laws covering stalking offences is the need for 'repeated' communications. This element of the law, examined in reference to online 'pile-ons', means that someone who posts a threat on a single occasion may not be covered by the provision.

In this section, we highlight challenges and opportunities regarding legal redress under criminal law encountered by women journalists in select countries.

In another UK case, in March 2021 a court issued a temporary stalking prevention order against a far-right figure to protect the *Independent*'s Home Affairs correspondent Lizzy Dearden and her partner. It was alleged that the man threatened and harassed both Dearden and her partner online and offline, including at their home, in a bid to chill her reporting. The order prevented the man from contacting the couple, or publishing anything about them on social media unless referring to Dearden as the author when responding to any story written by her. Issuing the order, the Deputy Chief Magistrate said: "What the police say in this case is he has embarked on all of this to persuade her not to publish the story" (PA Media, 2021).

In Finland in 2018, three people were convicted of "stalking", in relation to coordinated attacks on investigative journalist Jessikka Aro. She was the subject of a four-year harassment campaign, after she published stories about Kremlin-aligned troll factories (McCully, 2019). Aro filed a complaint with the police in Helsinki in 2016, and three people were ultimately prosecuted under three sections of the country's criminal law, with nine other people listed in the prosecution case as victims of the harassment campaign[223]

The Helsinki District Court held that the defendants had violated the country's criminal code by repeatedly contacting Aro by "tagging" her in social media posts, in

a way that fit the legal definition of "stalking", among other factors (see discussion about the function of defamation law in this case below). One of the defendants was sentenced to 22 months in jail, while the other two received suspended jail sentences.

The criminal code - which covers the offences of stalking prosecuted in this case - requires "contact", which could imply that only direct communications with a victim (i.e. DMs or posts in which they are tagged), and not messages to third parties about the target, or posts that make only oblique references to the victim, would be covered.

---

223 This case has some similarities with cases explored in this research, including Maria Ressa (the Philippines), Carole Cadwalladr (UK), Ghada Oueiss (Lebanon), Karima Brown (South Africa), and the BBC Persian language service group of journalists.
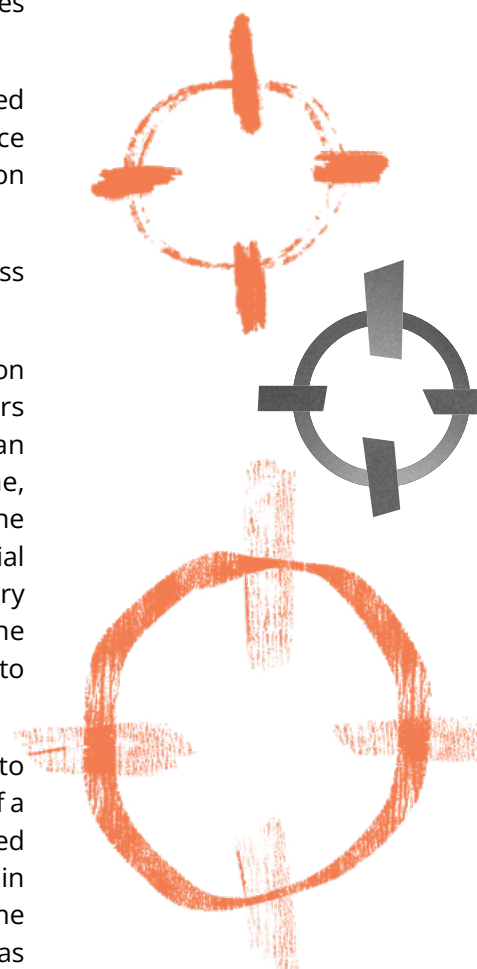
Jessikka Aro's case also highlighted the issue that prosecutors are unable to launch proceedings without a complaint from a victim. Complaints forwarded by third parties, such as colleagues, employers, family members or bystanders are unable to compel a prosecutor to investigate unless the victim complains

Prosecuting deepfakes in Ireland

In Ireland, a man was sentenced to four and a half years in prison in 2018 for harassing RTÉ journalist and presenter Sharon Ní Bheoláin. The case involved an early example of a harasser producing what are now known as deepfakes. The man uploaded 32 pictures to a website which had been doctored to show Ní Bheoláin's head on pornographic images which could be searched for online. The site was shut down during the police investigation. The officers also uncovered private messages in which he named Ní Bheoláin while discussing torture, murder and extreme sexual violence (McCully, 2019).

When sentencing the man, the judge described his actions as an "insidious form of harassment" and "debasing behaviour", noting that the "information on [Ms. Ní Bheoláin] will be out there forever" and "no doubt it caused considerable distress… It was reprehensible and he should be thoroughly ashamed" (INSI, UNESCO and TRF, 2021).

The following year (2019), another man, a self-described "internet troll" was sentenced to five years in jail after targeting six Irish women journalists with hundreds of abusive emails. While noting the internet's "wonderful advantages", the judge said it also had a "dark side which allows a man sitting in his house to inflict huge amounts of trauma on six women" (ibid).

New Irish legislation - Harassment, Harmful Communications and Related Offences Act 2020, known as the 2020 Act - specifically includes online harassment and cyberstalking; whilst Section 10 of the Non-Fatal Offences Against the Person Act 1997 (the 1997 Act) is the primary legislative provision in Ireland for the prosecution of incidents of harassment (INSI et al., 2021). Under Irish law, it would need to be established that a perpetrator intended for the threat to be believed, and there would have to be persistent abuse to prosecute cases of online harassment which did not include threats to kill or cause serious harm (ibid.).

Successful prosecution of rape and death threats in France

In France, in 2018, radio journalist Nadia Daam became the target of online violence (considered a form of moral harassment under French law) after she criticised the actions of members of an online forum during a broadcast. Her employer, Europe 1, filed a complaint on her behalf to the police, who then identified seven possible perpetrators, leading to two people being brought before the court. The charges concerned a rape and death threat which superimposed Daam's face onto an image of a victim of the so-called Islamic State . Both men were given six-month suspended prison sentences and were fined 2,000 Euros (McCully, 2019).

The Paris Court of Appeal ruled that the incriminating messages were "intended to 'punish' Daam as a journalist" covering women's rights, and "deserved a sentence sufficiently dissuasive to prevent a new offence, particularly via the internet, a communication tool perfectly mastered by the accused who, acting under a pseudonym, which proves his cowardice, could only be identified thanks to the cyber investigations of the police" (INSI et al., 2021). This case was successfully prosecuted under Article 222-17 of the French Penal Code, which criminalises threats to commit a crime. However the Article has a high evidentiary burden, with the prosecution needing to demonstrate a person making a threat knew and intended to create fear in the victim.

Cyberstalkers jailed in the UK

In the UK, two men were jailed in 2020 and 2021 for attacking journalist Amy Fenton[224] on Facebook in two separate episodes (Sharman, 2020; Tobitt, 2021a). In the first instance, the man posted on her newspaper's Facebook page that she "needed raping". He was sentenced to five months in jail (Tobitt, 2020; 2020c). In the second case, the man sent Fenton Facebook messages threatening to shoot her. He was jailed for nearly six months for "sending by public communication [Facebook] an offensive, indecent, obscene or menacing message" (Tobitt, 2021). The court also issued restraining orders against the perpetrators in both cases.

In another UK case, in March 2021 a court issued a temporary stalking prevention order against a far-right figure to protect the Independent's Home Affairs correspondent Lizzy Dearden and her partner. It was alleged that the man threatened and harassed both Dearden and her partner online and offline, including at their home, in a bid to chill her reporting. The order prevented the man from contacting the couple, or publishing anything about them on social media unless referring to Dearden as the author when responding to any story written by her. Issuing the order, the Deputy Chief Magistrate said: "What the police say in this case is he has embarked on all of this to persuade her not to publish the story" (PA Media, 2021).

Difficulties with legal definitions of targeted online violence in Sweden

In 2014, a Swedish journalist reported a series of threats she received to the police. However, the court found that the following threatening statement was protected by freedom of speech provisions, due to its general nature: "To me gender equality is when you take a sexist feminist whore in the vagina with a large knife" (Edström, 2016). The court also ruled that it would not be acceptable to explicitly name the woman. Feminist legal scholars in the country argued that the court was enabling perpetrators to threaten women in a more generic way: "Those who hate adapt. Now they know they should not put the name of the person they are threatening in the postings" (ibid.)

---

224 As noted in Chapter 2.5.5 "Increasing offline security in response to online attacks", Amy Fenton, former chief reporter for The Mail in the north of England, has been subjected to extreme online violence with offline impacts in connection with her reporting on grooming gangs. She has reported more than 100 such threats to law enforcement (Pidd, 2020).
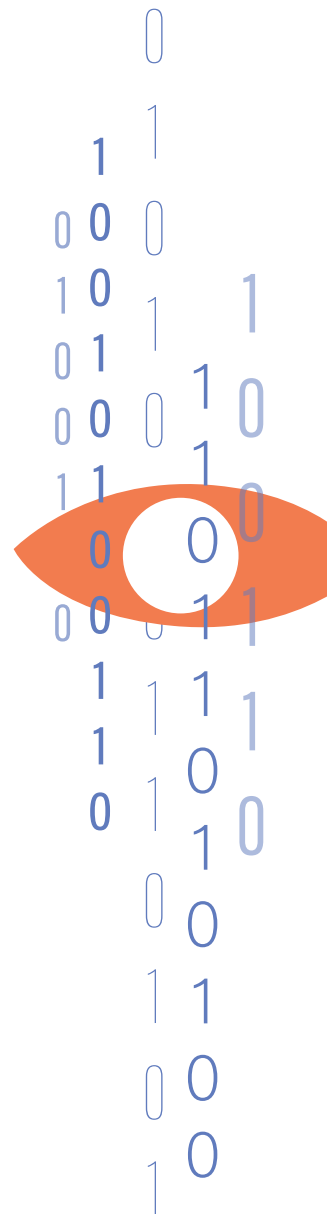
In December 2020, the Lebanese parliament passed a law that criminalises sexual harassment. The penalties include up to two years in prison and a fine of up to 20 times the value of the minimum wage, 675,000 Lebanese pounds (Al-Arabiya News, 2020). The penalties can be increased to a four-year prison sentence and a fine of 50 times the minimum wage if the crime is more 'serious' (such as related to a work relationship). The law defines sexual harassment as "any bad and repetitive behaviour that is extraordinary, unwelcome by the victim, and with sexual connotation that constitutes a violation of the body, privacy, or emotions". Sexual harassment can take the form of offline and online speech and actions. It can also be a single or repeated occurrence that enforces psychological, moral, financial or racist pressures to obtain sexual benefits (Human Rights Watch, 2021a).

Issues with witness testimony in Lebanon

The Lebanese legislation states that sexual harassment cases need to be heard in a criminal court instead of a civil court. Advocates have criticised this requirement, suggesting that victims would be less likely to report incidents as a result (Human Rights Watch, 2021a; Al-Arabiya News, 2020). However, these concerns may be resolved if criminal proceedings were held in camera, or in private, in an attempt to protect would-be complainants, a process common in other international jurisdictions (Mhaidly, 2018).

The double edged sword of Nigerian cyberharassment laws

Section 24 of Nigeria's 2015 Cybercrime Act penalises 'cyberstalking' or messages that are "false, for the purpose of causing annoyance, inconvenience, danger, obstruction, insult, injury, criminal intimidation, enmity, hatred, ill will, or needless anxiety to another". In theory, this could be viewed as a legal mechanism that could be used by women journalists experiencing online harassment. However, civil society organisations say this law has been used to prosecute journalists, bloggers and activists (Freedom House, 2019; CPJ, 2018; CPJ, 2019b; CPJ, 2022).

In 2019, journalist Obinna Don Norman was arrested and charged for allegedly harassing and defaming a politician. Nigerian authorities allege the offences were committed in 2018, but made no specific reference to articles or comments related to the charges. He was charged with four violations of the 2015 Cybercrime Act, from Sections 24 and 27: cyberstalking, sending defamatory messages using a computer, using a computer to send messages "...for the purpose of causing public hatred", and using a computer to "bully, threaten and harass". Norman told advocacy groups he believed the charges were in response to his critical reporting ahead of an upcoming election (ibid). This case highlights the potential for anti-online violence laws to be used to chill freedom of expression.

Law enforcement weak in Kenya where risks lurk

A 2020 report commissioned by the Media Council of Kenya found low levels of abuse reporting to the police and poor follow-through by them. One reason may

be the lack of knowledge among police officers about social media platforms and online violence, as suggested by interviewees for this study. Law enforcement may diminish or normalise the experiences of online violence and not see it as an urgent threat when no physical violence has yet taken place, according to interviewees. Similar patterns were described by South African interviewees.

But freedom of expression concerns have also been raised in Kenya, about the Computer Misuse and Cybercrimes Act proposed to regulate against the use and abuse of digital technologies which could cause harm or any form of criminal behaviour (Nitsche, 2019). The Act also sets out provisions to protect Kenyans or anyone living in Kenya against 'false publications', circulation of 'false information' and cyber harassment, with guidelines on protecting all citizens against crime, child pornography and false information (Articles 22, 23 and 27).

Taking legal action over online violence in Serbia is a complex and fraught process

In Serbia, sexual harassment is reported to take place online at rates above average for the region (OSCE, 2019). The Criminal Code of Serbia was amended to add the crime of sex-based harassment (Article 182a) and stalking (Article 138a) in June 2017, with penalties including a fine and imprisonment of up to six months. The Code defines sex-based harassment as any "verbal, non-verbal or physical conduct that aims to violate the dignity of the person in the sexual sphere, which causes fear or creates hostile, degrading or abusive environment" (AWC, 2019).

Legal experts interviewed for this study emphasised that the Serbian courts might see online threats as lacking "the character to induce a sense of endangerment" (especially in cases where they come from powerful figures). Other criminal offences have been identified by the Standing Working Group for Journalists' Safety as applicable to online harassment of women journalists in Serbia, and are part of the Group's initiative for amending the Criminal Code. These include 'computer sabotage' (Article 299) and unauthorised access to computers, computer networks or electronic data processing (Article 302); and racial and other discrimination (Article 387 para. 2, 4 and 6), (Babović and Reljanović, 2018).

In Serbia, online attacks are investigated and processed through special departments such as the Special Prosecution for High Tech Crime in the Higher Public Prosecutor's Office, and under the scope of the Ministry of Interior's Special Department for Combating Cybercrime, established under the Law on Organisation and Competences of Government Authorities in the Fight against High-Tech Crime. However, interviewees described difficulties connected to digital capability among investigators in online violence cases.

In 2018, Serbian journalist Verica Marinčić filed criminal charges against a man who was allegedly threatening her online and physically stalking her in connection with a news story (Djurić, 2019; Mapping Media Freedom, 2018; Apro, 2020). The accused man reportedly swerved his motorbike into her after she commenced legal action (Stojanovski, 2019). When she lost the case, she said: "I cannot report cases of harassment and intimidation anymore while I see how they are turning a victim into a fool" (Apro, 2020).

According to the Slavko Ćuruvija Foundation[225] and the Centre for Judicial Research (CEPRIS), 70% of cases of reported attacks against journalists and media workers in Serbia end with the prosecution rejecting criminal charges. Nevertheless, Ana Lalić, journalist at Nova.rs, a Serbian news website, had six online violence-related lawsuits in motion at the time of writing. "I am fighting here for my name and for my personal integrity, which they are trying to tarnish," she said.

Barriers to justice in Sri Lanka

Some sections of the Sri Lankan Penal Code could be used to prosecute acts of cyber violence including sexual harassment, criminal intimidation, criminal breach of trust, blackmail, extortion, and impersonation. However, a fundamental problem is that women have described great difficulty reporting violence to the authorities in general: "With this legislation in place, it should theoretically be easy for victims to access justice for crimes committed against them online. However, most victims do not pursue these solutions due to flaws that exist in the system of reporting online violence to authorities" (Perera and Wickrematunge, 2019).

An anonymous Sri Lankan interviewee catalogued the problems she experienced in pursuit of justice as: a lack of specific laws to deal with gendered online abuse and harassment; the absence of a designated law enforcement authority to handle complaints; a lack of knowledge of these crimes within the police force; and a lack of legal support to prosecute cases. She also argued that there needed to be more effective ways of reporting, documenting and investigating such crimes, including a hotline for complaints or a database for indexing and searching cases.

5. Defamation action as a defence

Defamation is a remedy found in common law jurisdictions to protect individuals against the publication of false statements about them that harm their reputation (Harvard Law Review, 2018).[226] Defamation and libel are key features of online violence against women journalists - especially in cases where disinformation tactics are deployed. However, there is a tension involved for many journalists here. This tension is based on the professional norm that journalists should not sue for defamation because defamation law is frequently used as a tool to suppress critical reporting, chill press freedom and limit the public's right to access information. However, freedom of expression protections do not extend to hate speech or disinformation which could damage someone's reputation, or intimidate them into discontinuing their reporting. After a due process of law, where the three-part test is applied,[227] could prove to be one of the more effective responses to online violence against women journalists.

In Finland, three people were convicted of "aggravated defamation" and "incitement to commit aggravated defamation" in 2018 in relation to the aforementioned coordinated attacks on investigative journalist Jessikka Aro. One defendant was the editor-in-chief of a website that published sexist abuse and racial slurs. In

---

225 Civil society organisation founded in memory of the murdered Serbian journalist, which works to advance media freedom, and member of the Standing Working Group of Journalists.
226 As long as they are applied in line with General Comment #34: https://www2.ohchr.org/english/bodies/hrc/docs/gc34.pdf
227 For more information on "The Legitimate Limits to Freedom of Expression: the Three-Part Test", see: https://www.youtube.com/watch?v=Wg8fVtHPDag

addition to criminal convictions for stalking (see point 4. above), the Helsinki District Court found that the defendants had:

- Published a series of false and defamatory articles about Aro;

- Encouraged others to publish defamatory statements about Aro;

- Committed copyright violations.

The three defendants in the case were convicted and ordered to pay a total of EUR €238,625 in damages, costs, and legal fees, and Aro herself received EUR €94,000 in damages.

Brazilian journalist Patricia Campos Mello won two relevant defamation actions in early 2021. She successfully sued powerful political figures for defaming her in connection with an onslaught of online harassment and abuse that followed her reporting on disinformation associated with Brazil's 2018 election. The judge ordered her attacker to pay Campos Mello US $5,500, saying that he "should have more caution with his statements — something that is expected from all those with some sense of responsibility to the nation" (Neder, 2021). In March 2021, Campos Mello also won her case against another attacker, who was ordered to pay her damages. According to the judge, the attacker's remarks had damaged her honour (BBC News, 2021b). Despite national and international attention and her court victories, online attacks against Campos Mello nevertheless continued on social media in mid-2021, highlighting the need for further preventive legal measures, and a stronger response from the social media companies serving as vectors.

Pakistan introduced the Prevention of Electronic Crimes Act (PECA) in 2016 with the stated objective of contending with online hate speech, extremist content and harassment against women. However, the legislation was criticised by human rights groups which argued that it could lead to the violation of freedom of expression rights and enable censorship. The law, administered by the Federal Investigation Agency (FIA), is perceived by some as a tool that can enable harassment of women, according to Human Rights Watch (Gossman, 2020). Nine complainants were charged with criminal defamation in 2020 under Section 20 of PECA, while many complaints about online violence against women journalists have remained unaddressed (ibid).

6. Privacy-based legal action

Gender-based online violence against journalists often violates the privacy of victims, and there may be legal avenues found in privacy laws. In Scotland (United Kingdom), for example, a person can be sued for 'breach of confidence' which seeks to protect violations of a victim's 'autonomy, dignity and self-esteem' (Hill, 2015). Under the tort, the court would consider whether the subject had a reasonable expectation that content (e.g. intimate images) would remain private. It is arguable that this could allow women to sue for so-called 'revenge porn' forms of gendered harassment. However, suing for privacy breaches would be after the fact, and may be of limited value in preventing the threats associated with privacy breaches.

In Canada, there have been attempts to address a number of relevant issues, including prohibiting the non-consensual distribution of intimate images, allowing courts to order the takedown of such images, and awarding compensation (Government of Canada, 2014). In Ireland in 2020, it became a crime to share intimate images without consent under the Harassment, Harmful Communications and Related Offences Act (2020). Known as 'Coco's Law', the Act criminalises such behaviour - online and offline. Importantly it also enables the protection of victims' anonymity and expressly applies to "electronic communications". There are two new offences created under the act relevant to the use of 'revenge porn' as a tactic to threaten or shame women journalists into silence (Kelleher and Daly, 2021):

- Distributing, publishing or threatening to distribute or publish intimate images without consent with intent to cause harm or being reckless as to whether harm is caused. Crucially, it is irrelevant that a person may have consented to the taking of an image if it is subsequently published or distributed without their consent. This offence will carry a maximum penalty of an unlimited fine and/or seven years' imprisonment.

- Recording, distributing or publishing intimate images without consent. This is a strict liability offence as the person who records, distributes or publishes the image without consent, does not need to have intended to cause harm. The maximum penalty for this offence is EUR €5,000 and/or 12 months' imprisonment.

The inclusion of a strict liability offence in Section Two of the Act is also notable, as the prosecution does not need to prove intention, knowledge, recklessness or even negligence (Hashmall, 2017). Under Section 4 of the same act, it is also now a criminal offence to distribute, publish, or send a threatening or grossly offensive communication (Harassment, Harmful Communications and Related Offences Act, 2020). Convictions under this Section of the Act attract fines and/or a prison sentence of up to two years.

In Brazil, Law No. 12,737/2012, popularly known as the 'Carolina Dieckmann law', was introduced in 2012 after nude photos of the Brazilian actress were hacked and shared after a failed attempt to blackmail her. The law classifies digital crime, including hacking a computer device to obtain, tamper with, or destroy data or information without the authorisation of the device's owner (Glickhouse, 2013). This broadened the scope of the computer intrusion crime under this law, by punishing any form of unauthorised access into a third-party device (Freedom House, 2016a). However, the law's penalties have been described as too "weak to be a deterrent - with just three months to one year in prison and a fine" (Thompson and Muggah, 2015). Law no. 12.965/2014, also known as the Marco Civil Law or 'Constitution for the Internet' (Presidência da República, 2014) offers detailed privacy protections pertaining to personal data, guarantees net neutrality, and promises to uphold the participatory nature of the internet (Freedom House, 2016a). Some online violence victims have filed claims under these laws relating to privacy breaches.

In Mexico, the Olimpia Law on online safety was passed in 2018. Named after Olimpia Melo, a survivor of online violence and advocate for a free and safe internet, it bans crimes against privacy (the dissemination of intimate content without consent) and cyberbullying (including online sexual violence). New crimes were recognised in the state Penal Code, including 'violation of sexual intimacy'

(Article 182) and 'crimes against sexual intimacy' (Article 225). However, the laws have been criticised by feminists as failing to respond to women's needs, both theoretically and practically. And there are also concerns about the gap between legislation and implementation (Aguirre et al., 2020). Between 2017-2020, 2,143 investigations were launched into the crime of disseminating intimate images without consent. 83% of the investigations were still in process at the end of 2020. Only 175 of the case files were settled through alternatives to justice, such as conditional suspension of the process, reparatory agreement, or abbreviated procedure (ibid).

7. Prosecuting acts of doxxing

Doxxing as part of online violence against women journalists is also correlated with additional offline risks. Some academics argue that conventional legal instruments do not provide a reliable remedy for victims of doxxing. While doxxing often implies threats by releasing personal information that could invite physical harm, the act of releasing personal data often does not explicitly include threats, meaning that conventional harassment and stalking laws do not apply (MacAllister, 2017). Also, some information revealed by doxxing might already be on the public record, and current legal frameworks in many jurisdictions effectively give immunity to perpetrators (ibid). However, at least one successful prosecution of a perpetrator who doxxed a woman journalist - South African editor Karima Brown discussed in point 8. below - has been identified.

In Sri Lanka, some digital crimes may, in part or whole, fall under general laws, such as the Computer Crimes Act (No. 24 of 2007) which also prohibits hacking, or the Obscene Publications (Amendment) Act (No. 22 of 1983). But these provisions are focussed on e-commerce, and so are not clearly useful to journalists trying to secure prosecution of doxxing for example (Samaratunga and Hattotuwa, 2014). Two agencies currently investigate cybersecurity related crimes. One of them, the Sri Lanka Computer Emergency Readiness Team (CERT), is the first point of contact for individuals reporting threats and vulnerabilities in computer systems and online networks, such as fake accounts, hacking, image-based abuse, and cyberbullying amongst others. CERT provides advice on information security, privacy violations and identity theft, and refers victims of online violence to the police or the Criminal Investigation Department (CID); however, there are no specific provisions for addressing journalists' complaints within these systems, according to a Sri Lankan journalist interviewed for this study who wished to remain anonymous. CERT told her to take her case of identity theft to the CID, who then referred her to their Cyber Crimes division, who told her that they could not take action because they had never encountered such a complaint before, and it had not caused her financial loss. She said she then took the issue to the Telecommunication Regulatory Commission of Sri Lanka, who referred her back to the CID.

8. Electoral and equality law

The South African Electoral Code was the basis for successful legal remedy in a novel legal challenge on behalf of an editor who was doxxed by a political leader and experienced a torrent of online abuse in response. In 2019, the Gauteng High Court in Johannesburg ruled in favour of the late political editor Karima Brown, whose phone number was published on Twitter by the leader of the Economic Feedom Fighters party, Julius Malema, along with claims that she was a State intelligence operative, triggering threats of rape and other acts of violence from self-described party's supporters. The court found that the party's failure to condemn the harassment of Brown breached the Electoral Code's requirement to respect the rights of women and the news media, and instructed the party and its leaders needed to take reasonable steps to condemn and stop the harassment experienced by the journalist (Columbia Global Freedom of Expression, 2019).

Brown was awarded the equivalent of USD $7,000 in damages and the court ordered the party to formally apologise and delete offending messages from all platforms. Following the verdict, Brown said: "This is a victory for media freedom, a victory against sexism, and it is a victory for women in journalism, and protection, and freedom of the media" (Chabalala, 2019a). This judgement highlights the ways in which electoral law could be used to prosecute cases of online violence and, importantly, potentially work as a deterrent against attacks instigated by political actors.

In another test case from South Africa involving the same political party, the Equality Court ruled against a claim brought on behalf of several journalists, including Daily Maverick investigative journalist Pauli van Wyk who participated in this study. The South African National Editors' Forum (SANEF) took the party to the Equality Court in 2019 to argue that the party had enabled an environment of intimidation and harassment of journalists. SANEF lost the case, with the Equality Court deciding the case did not fall within its terms of reference (Chabalala, 2019b).

9. Copyright Violations

Some legal practitioners have used copyright laws in an attempt to address gendered online violence. In the US, for example, lawyers have used notice and takedown provisions under the Digital Millennium Copyright Act (DMCA) (MacAllister, 2017). However, this is limited to content owned by a victim of harassment, such as selfies or other content they produced themselves. Copyright remedies would not extend to content filmed, consensually or not, by someone else (Chen, 2016). This also makes the role of copyright problematic as a remedy for harassment against women journalists involving so-called 'revenge porn' (i.e. sexual imagery shared without consent to damage a woman's reputation, or to cause shame). There are instances of women journalists being threatened with the release of such content in several countries (Columbia Global Free Expression, 2020; Walsh, 2020).

Facebook was widely criticised for asking users to send their intimate pictures to the platform in its attempt to automate the removal of such images during a 2017 experiment. The company asked users to pre-emptively submit their pictures so

that it could detect images posted elsewhere on the site and automatically take them down (Solon, 2019). Some critics raised privacy concerns, others noted that a third person would have to see the images, but some supported the move, saying it gave users agency (ibid).

The copyright remedy is also limited by the nature of the internet. Once images or private information are published online, the author can quickly lose control of the content as it is screen-grabbed, copied and disseminated, limiting the capacity of copyright takedowns to remove it. Further, some legal academics note that using copyright laws in this way distorts its rationalisation, which is to stimulate the creation of new works by ensuring fair compensation, rather than the suppression of content (ibid).

10. Limitations of conventional communications legal instruments

In the US, for example, laws governing traditional communication methods could arguably address cases of gendered online violence. The Interstate Communications Act provides that anyone who "transmits in interstate or foreign commerce any communication or any threat to kidnap any person or any threat to injure the person" will be fined or imprisoned up to five years, or both (Lipton, 2011). However, this arguably would not extend to other kinds of threats made about a person, nor necessarily apply to social media content. Such limitations are familiar in many jurisdictions internationally.

11. Heavy costs

Other obstacles to legal remedy identified by legal academics researching gendered online violence include the prohibitive costs of bringing a case or justifying prosecution. Quantifiable costs include legal fees, lost income and protection services. Some economic impacts are harder to quantify, including the time, effort and mental well-being imposed by reporting online harassment and abuse to the police. For example, in circumstances where prosecution for acts of online violence requires proof that the conduct was repetitive or ongoing, evidence gathering can be particularly onerous: "One victim described spending 'countless hours' over four years logging the online activity of one particularly committed cyberstalker, 'just in case' he carried out his threats" (Marshak, 2017).

Several journalists interviewed also raised concerns about the capability of law enforcement and judicial investigators to investigate online violence. In Serbia, journalist Jovana Gligorijević was required to submit supporting evidence (such as screenshots, time-logs, etc) on a CD-rom to verify online threats from far-right actors because the Prosecutor for High Tech Crimes was not able to access social media (including YouTube) from their premises, she said. After months of delay and a lost evidence dossier, a police officer asked for Gligorijević's help to review the materials documenting the threats because he did not know how to use social

media. In Lebanon, former investigative journalist Myra Abdallah filed a complaint with the Cybercrimes Bureau about death threats she had received, but the case was dropped. "I was told we don't know who [the fake accounts] belong to and the real accounts we cannot do anything about! I was told 'you seem to be messing with the wrong people'," she said. BBC disinformation reporter Marianna Spring had similar experiences in the UK, when she reported rape and death threats to the police in 2020.

This underscores the need for law enforcement to significantly improve the digital investigative capabilities of officers and units assigned to deal with cases of online violence against women journalists. But it also highlights the value of news organisations and civil society actors investing in systems that enable secure third-party documentation of abuse to relieve the victim of the burden. Related to this, is the burden imposed on victims to 'relive' their online violence experiences for the public judicial record in the course of conventional legal redress. Victims may be reluctant to make complaints or give evidence in court as a result (Lipton, 2011).

The international legal NGOs Media Defence and ARTICLE 19 have also been exploring options for strategic litigation against perpetrators and facilitators of online violence against women journalists as of the time of writing in mid-2021. However Joanna Connolly, former Legal Officer for Media Legal Defence Initiative, noted that they were having difficulty surfacing cases involving women journalists who were prepared to put themselves through onerous litigation processes.

"In the cases that we've taken, the women have faced greater threats and greater reprisals, specifically because they tried to take legal action, specifically because they approached the police...in the environments we work in, that worsens their situation," she said. Meanwhile, irrespective of legal developments, there is mounting public and industry pressure on the social media companies to do much more, and more rapidly, to protect the rights of their users, and in particular those of women journalists, as part of their obligations under the Ruggie Principles.

12. Legislating against online violence

The UK's Online Harms White Paper (DCMS, 2020) which led to the Draft Online Safety Bill (DCMS, 2021) is an example of an attempt to protect people from online violence, which could be relevant to the case of women journalists. The UK government would be obligated to make social media companies uphold their 'duty of care'. This would in turn be overseen by the UK's main independent communications regulator, Ofcom. As the bill stood at the end of 2021, internet communications companies could be fined up to GBP £18 million or 10% of their global turnover for 'harmful' but 'lawful' content that violates binding corporate commitments to deal with abuse. 'Journalism and democratic political debate' are covered under the protections afforded (Tobitt, 2021b). In mid-2021, the UK Culture Secretary said she would follow a recommendation by the Law Commission - an independent body that reviews laws in England and Wales - to

include a 'psychological harm' crime in the bill,[228] with reference to online pile-ons (Law Commission, 2021; Milmo, 2021).

The original UK Online Harms White Paper was also followed by the National Safety Action Plan for Journalists (NUJ, 2021), which includes training and support for police representatives, newsroom leaders and student journalists. But Michelle Stanistreet of the UK's National Union of Journalists, noted: "Even with the laws as they are today, there are mechanisms to tackle [online harassment] and for it to be taken seriously and robustly. They're not deployed or utilised now as much as they should be." There are concerns that the draft law could also be abused by bad faith actors to jeopardise freedom of expression, according to UK-based human rights lawyer Caoilfhionn Gallagher KC. ARTICLE 19 has raised similar concerns that the draft lacks an effective notice and appeals process for content moderation decisions, transparency and accountability. "The draft Bill not only addresses various types of illegal content but also introduces the extremely problematic concept of 'legal but harmful,' which threatens protected expression" (Caster, 2021).

Nigeria's 1999 Constitution guarantees the right to freedom from discrimination on the grounds of sex and recognises women's equal rights. While the country has a National Gender Policy aimed at protecting women from all forms of oppression (Tijani-Adenle, 2019), there is a lack of a legal framework for safeguarding women online. This is seen as contributing to the under-reporting of gender-based harassment (World Wide Web Foundation, 2015). Journalist Kiki Mordi received online death threats for a documentary which exposed her to escalating online harassment (Asamoah, 2019). She chose not to report the attacks, even though she knew she had legal grounds: "The legal process isn't straightforward...I wasn't ready for that. Plus, I didn't have the money to start legal battles or the time to stay off work. I'm freelance, so every second counts".

A Group of Experts on Action against Violence against Women and Domestic Violence, who evaluated legislative measures introduced in Serbia to meet the requirements of the Istanbul Convention, found both the implementation of sexual harassment and stalking laws have often been hampered by significant media backlash. There had been a trivialisation of these offences within public discourse "as the criminalisation of flirting", and a general lack of understanding of the essence of stalking and sexual harassment (GREVIO, 2020).

Germany passed a law in January 2021 reinforcing police powers to investigate online hate speech. Likewise, in France, a new Act called Reinforcing Respect of the Principles of the Republic, which includes provisions on online hate was passed. In September 2021, the European Commission also adopted a Recommendation on the protection, safety and empowerment of journalists which is intended to encourage Member States to take further legal steps to ensure safer working conditions for all media professionals, free from fear and intimidation, whether online or offline.

---

228 At the time of writing the UK's Online Safety Bill sat with a Joint Committee to assess the legislation (Dawood, 2021).

13. Legislating against misogynistic hate speech

A number of jurisdictions have attempted to legislate against cyberbullying and cyberharassment, which may extend to the harassment of female journalists. However, one of the key legislative gaps identified in this research was the exclusion of sex and gender from anti-hate speech legislation, which routinely covers race, ethnicity, religion, sexual orientation and sometimes transgender identity. Some of the states studied in this research, where gender and sex were excluded from existing hate-speech legislation, include Poland, Sri Lanka, Sweden, and the UK.

In 2021, the US Congress enacted a hate crime bill giving specific protection to Asian-Americans and Pacific Islander people, but this does not recognise misogyny as a hate crime and so does not specifically help women journalists to take action in reference to sex- or gender-based hate speech (Sprunt, 2021; White House, 2021).

From January 2020, the Sri Lankan Defence Ministry began work on new cybersecurity legislation to tackle online defamation as well as ethnically and religiously sensitive posts that incite hatred and pose a threat to national security. This Cyber Security Act was proposed following the Easter 2019 terrorist attacks on churches and hotels which resulted in anti-Muslim violence in the country (Bemma, 2019). It would also address crimes including revenge porn and hacking (Ministry of Defence, 2020). However, there is no specific legislation to protect women journalists from digital violence in Sri Lanka, according to Professor Prathiba Mahanamahewa, Dean of Law and former Human Rights Commissioner of Sri Lanka. While Sri Lanka's Penal Code Article 153 criminalises hate speech which promotes enmity between groups on the basis of religious, racial, language, region, caste or community difference, it does not address gender-based hate (Sri Lanka Government, 1885).

In Brazil, hate speech can be framed as a crime against 'honour', including false accusations and defamation, under Articles 138, 139 and 140 of the Brazilian Penal Code (IRIS, 2019), and it can also be understood as a crime against the public peace (Articles 286 and 287). The Anti-Racism Act (Law No. 7716/1989) criminalises the practice or incitement of discrimination on grounds of race, religion, or national origin (Presidência da República, 1989), and there have been prosecutions for online hate speech under these laws. In 2018, a high profile misogynist was sentenced to 41 years in prison for inciting racism, and making death and terror threats, particularly against women (Uchoa, 2019; Declerq, 2018). His most prominent target was the author of feminist blog 'Write Lola, Write', Lola Aronovich, who has endured years of harassment and threats since 2011. At the time, Aronovich's local Women's Police Station told her it was "unable to carry out investigations", as they involved complex actions such as accessing a website hosted abroad. The Federal Police told the professor of English Literature at the Federal University of Ceará it was "not their job to investigate this type of crime" (Declercq, 2018). But following the man's conviction, 'Lei Lola' [229] ('Lola's Law') (Presidência da República, 2018) was introduced allowing "the federal police to take over any investigation

---

[229] 13.642/2018

into online crime of a misogynistic nature" and makes hate speech against women illegal (Evans and Coelho, 2019).

France has also adopted a gender-sensitive law addressing 'cyberharassment', which "criminalises the repeated targeting of an individual with both sexual and sexist statements that harm the victim's dignity through their degrading, humiliating, intimidating, hostile or offensive nature" (McCully, 2019). This was in part a response to a Facebook group, created by male journalists in 2009, called the 'League of LOL' (Ligue du LOL) which was found to be harassing women journalists, among others, and encouraging pile-ons (RFI, 2019). Illegal content can also be reported and complaints filed on the platform PHAROS,[230] hosted by the French national police force where officers are trained to track the IP addresses of online attackers that hide behind a pseudonym to obtain their IP address (INSI et al., 2021).

It appears evident that to be an effective deterrent to online violence, hate speech legislation needs to cover misogyny if women are to protect themselves from gender or sex-based hate speech, if they are not able to claim protection under other hate speech categories.

Conclusion:

The UN and other international and regional bodies set clear guidelines and frameworks about the obligations of states to introduce laws to protect women journalists against online harassment. However, in practice, women journalists face a series of barriers and obstacles when seeking to prevent and/or remedy online harassment.

Existing laws are often out-dated and ill-equipped to deal with the modern realities of reporting in the internet era. The need for physical proximity, the need for repetition, defining what constitutes "real" threats, and the inability to identify perpetrators make achieving justice for victims of online harassment difficult to attain.

The additional emotional, financial and time costs faced by women is another barrier. Online harassment of women journalists is a global issue, which makes the prosecution of harassment across jurisdictional and global border lines extremely difficult. Women journalists have often had to resort to creative legal remedies and procedures, which were not designed for online harassment, such as copyright legislation and privacy laws.

Third party platforms, where the harassment often takes place, are often uncooperative or protected by freedom of speech laws or corporate interests. And while more and more countries are beginning to legislate against online violence, misogyny is often omitted or excluded in anti-hate speech laws. Against this background, recommendations at the end of this report, and a 25-point assessment tool, are provided as ways forward for States to improve legal protection of women journalists subjected to online violence.

---

230 'Platform for Harmonisation, Analysis, Cross-referencing and Orientation of Reports': https://www.interieur.gouv.fr/Archives/Archives-publications/Archives-infographies/Securite-des-biens-et-des-personnes/Securite-des-biens-et-des-personnes/Cybersecurite/PHAROS

7

Conclusions and Recommendations

Julie Posetti and Kalina Bontcheva
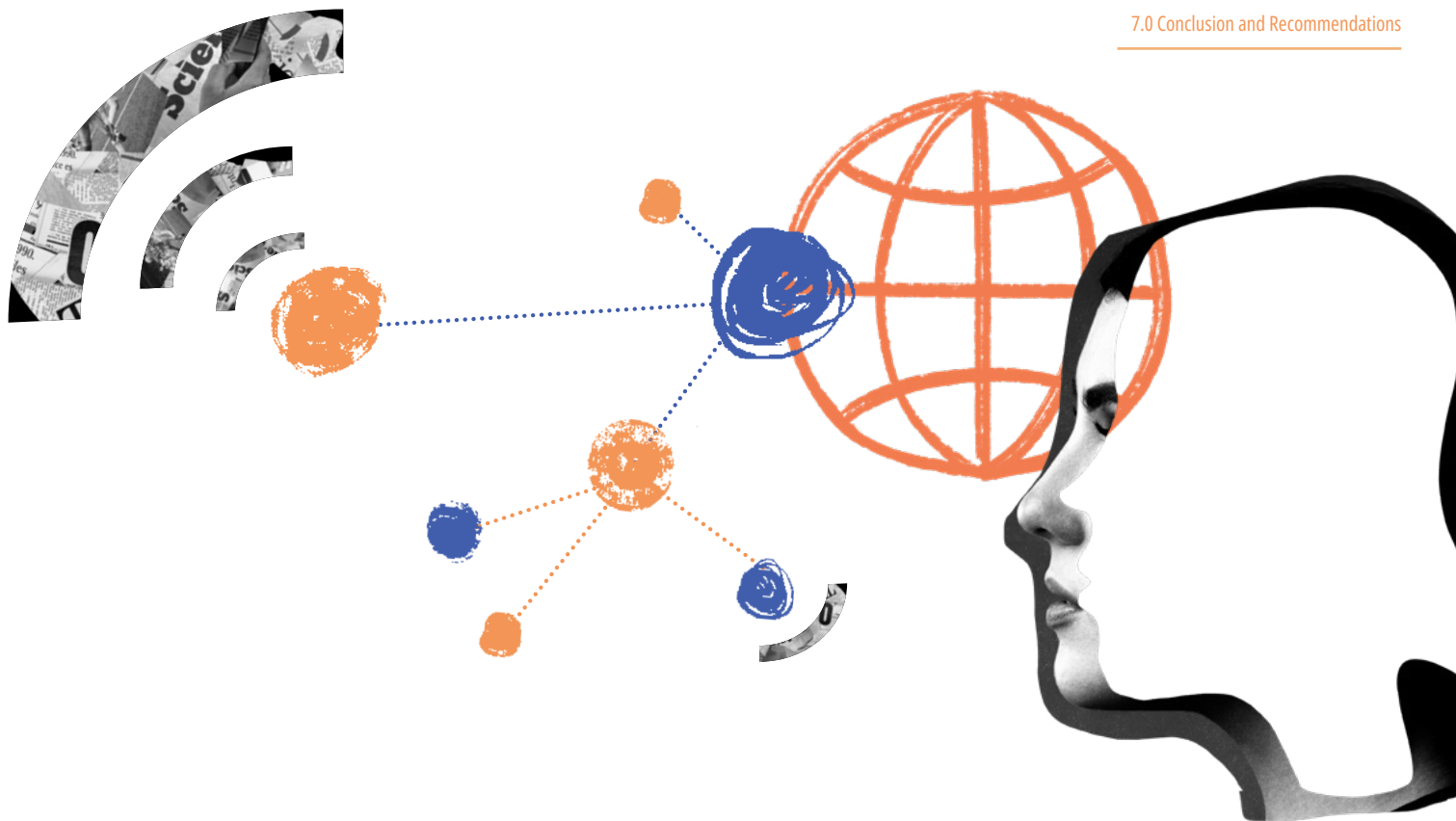
> "We call on States to prohibit, investigate and prosecute online and offline gender-based violence against women journalists. Investigators must presume such crimes are connected to the journalist's work unless proven otherwise. States as well as companies must include a gender perspective in all initiatives aimed to create a safe and favourable environment for independent, diverse and pluralistic media."

Joint statement from the UN Special Rapporteurs, November 2021 (OHCHR, 2021)

This study has demonstrated that online violence against women journalists is a debilitating global phenomenon, albeit one with diverse manifestations regionally, and uneven impacts that are heightened at various intersectional points. While misogyny, sexism, and patriarchal norms are the main identifiable driving forces behind gender-based online violence against women journalists, the incidence and effects are more severe at the nexus of racism, religious bigotry, antisemitism, sectarianism and homophobia.[231]

As evidenced, the types and methods of attack deployed against women journalists online are growing more sophisticated, and they are evolving with technology such as Artificial Intelligence. For example, synthetic media formats are increasingly used to defame and misrepresent women journalists in an effort to discredit them personally and professionally, thereby eroding trust in their journalism and increasing their exposure to threats, harassment and abuse.

---

231 The recommendations and 'online violence response assessment framework' included in this chapter will also be published separately by UNESCO here: https://en.unesco.org/publications/thechilling

As elaborated in this study, targeted online attacks on women journalists are also increasingly networked, and at times State-linked. This points to the need for responses to online violence to grow equally in technological sophistication and collaborative coordination.

This study has described the correlation of online and offline violence and pointed to causation in a number of cases involving women journalists in the US, UK, the Philippines, Mexico, Pakistan, and Sri Lanka. It has also mapped the vicious circular trajectory of online violence: digital attacks can fuel offline violence, while offline abuse by prominent figures can trigger online pile-ons.

This highlights the possibility of treating online attacks on women journalists as an early warning system (ICFJ, 2021; Posetti *et al.*, 2021) to help predict offline violence and target responses to impacts such as psychological injury. Investment in collaborative solutions that support intergovernmental organisations and States to monitor, report, condemn and prevent attacks against journalists in a way that integrates online violence data is therefore recommended.

The research has also underscored the climate of impunity surrounding online violence against women journalists, which must be more urgently and effectively addressed. Impunity emboldens the perpetrators, demoralises the targets, and erodes the foundations of journalism (Harrison and Pukallus, 2018). When perpetrators of online violence are able to act with impunity, the threat of impunity for offline harassment and crimes against journalists becomes further entrenched.

This study has illuminated the evolving challenges faced by women journalists dealing with prolific and/or sustained online attacks around the world. It has called out the victim-blaming and slut-shaming that perpetuates sexist and misogynistic responses to offline violence against women in the online environment, where patriarchal norms are being aggressively reinforced. It has also demonstrated the inadequacy of responses from social media companies that enable gendered online violence spread at scale on their platforms. And, it has identified and analysed obstacles encountered by news industry managers and civil society organisations seeking to respond more effectively to the crisis.

Additionally, the researchers have assessed international examples of legal and legislative responses, identifying key gaps in process, and emphasising the need to ensure that rights afforded to women journalists offline are also upheld online.

Finally, this study has registered the role of political actors and State officials, in perpetrating and fuelling online violence against women journalists in order to chill their critical reporting.

Not only does online gender-based violence endanger the careers, health and well-being of women journalists, it also has chilling effects on press freedom and the public's right to access information, while further inhibiting gender equality in, and through, the news media. Therefore, when it becomes entrenched, it can contribute to the erosion of democracy.

There is, therefore, a very urgent need to address this onslaught against freedom of expression and gender equality. It is time to reverse the onus on individual women journalists to manage and respond to gendered online violence and make it the responsibility of social media companies, news media employers, policy makers and legislators, law enforcement and the judiciary.

This global crisis demands proactive solutions, rather than reactive responses. To that end, we offer 107 recommendations for key responders alongside a 25-step assessment protocol designed to assist UNESCO Member States and intergovernmental organisations to formulate effective legislative, regulatory and policy responses to counter gender-based violence against women journalists while respecting freedom of expression, access to information and privacy rights. It is also of relevance to big tech and news media companies.

Online Violence Response Assessment Framework

This 25-step tool[232] **can guide responses to online violence against women journalists at the legislative, legal, and policy levels with regard to international human rights laws and norms.**

1. Do the responses recognise online **violence** as *violence*, and the psychological impacts as **real?**

2. Have the responses been developed under a **gender-sensitive lens,** which takes account of the increased risks facing women journalists, and especially those at the **intersection of misogyny, racism, religious bigotry, sectarianism** and other forms of discrimination? And are they framed in a way to **alleviate discrimination?**

3. Do the responses recognise **misogyny and sex-based discrimination as forms of hate speech,** acknowledging that misogyny should be treated as seriously as hate speech that is focused on race/ethnicity, religion, disability, sexual orientation, and gender identity?

4. Are journalists able to **receive effective judicial protection from hateful content received in the course of their work** which incites hostility, violence and discrimination, even if it has not **(yet)** triggered **offline attacks?**

5. Do the responses clearly and transparently identify the specific problems such as **threats of physical and sexual violence;** the **enabling role of social media companies;** and **practices by officials or foreign State actors that impact negatively** on the safety of journalists? Do the responses recognise **impact on critical independent journalism** (e.g., reporting focused on gender-based violence, election integrity, or disinformation networks)?

6. Do the responses impinge on, or limit, **freedom of expression, press freedom, access to information and privacy rights?** If so, and if the circumstances triggering the response are considered appropriate for such intervention (e.g., a threat to human life or incitement to violence), is the interference narrowly-defined, necessary and proportionate?

---

232 This tool builds on a similar protocol developed as part of a global study into disinformation responses published by UNESCO and the UN Broadband Commission in 2020 (Bontcheva and Posetti, 2020).

7. Does a given response (unintentionally or otherwise) risk **restricting journalistic functions** such as reporting, publishing and confidentiality of source communications; does it limit the **right of access** to public interest information? (refer to no. 6 above) Responses in this category could include: hate speech laws; communications interception and surveillance; data retention and handover; anonymity and encryption overrides. If the measures do impinge on these journalistic functions, or on accountability of duty-bearers to rights-holders in general, do they provide **exemptions for acts of journalism**?

8. Do responses include an **impact assessment** as regards consequences for international human rights frameworks that support freedom of expression, press freedom, access to information or privacy? Do such assessments take account of the fact that women journalists have **a right to work online free from hate speech, including misogynistic abuse and other online violence?**

9. Are the responses (e.g., legislative, normative, legal, etc.) considered **together and holistically** in terms of their different roles, complementarities and possible contradictions?

10. Do the responses avoid the **false binary position that the right to freedom of expression cannot co-exist with, or be balanced against, the right to be protected from online violence** which inhibits free expression?

11. Have the responses been the subject of **multi-stakeholder engagement and input** (especially with civil society organisations, industry representatives, specialist researchers, and press freedom experts) in their **formulation, implementation and review?**

12. In the case of **legislative** responses, has there been appropriate opportunity for **deliberation prior to adoption,** and do the laws and regulations provide for **independent oversight of implementation and recourse to appeal?**

13. Are responses primarily **restrictive** (e.g., criminalisation of misogynistic online attacks; regulatory interventions targeting social media companies), or is there an appropriate **balance with enabling and empowering measures** (e.g., increased capability among judicial and law enforcement actors through training and development; investment in support for digital safety and security training for journalists; requirements for social media companies to support women journalists under attack)?

14. Do **legal responses** come with **gender-aware guidance and training** for implementation by law enforcement, prosecutors and judges, concerning the need to protect the core right of freedom of expression and the implications of restricting this right as regards online violence against women journalists?

15. While the impacts of online gender-based violence can vary in seriousness, does the response recognise the **range of manifestations** - from one-off threats of an extremely violent nature, through to the **cumulative effects of sustained lower-level harassment and abuse?** And does it recognise the **distinctions** between an

individual perpetrator with limited means to execute a threatened act of physical violence, and an orchestrated mob or State-linked attack, for example?

16. Is the response designed to be **transparently assessed,** and is there a process to systematically **monitor and evaluate** the freedom of expression impacts (such as through reports to the public, parliamentarians, specific stakeholders)?

17. Is a given response able to be **rolled back** if it is found that any benefits are outweighed by **negative impacts on freedom of expression, access to information and privacy rights** (which are themselves potential antidotes to gender-based violence online)?

18. Are measures relating to the internet companies developed with **due regard to multi-stakeholder engagement,** and in the interests of **promoting transparency and accountability,** while **avoiding privatisation of censorship?**

19. If the response is targeting **internet companies**, does the measure consider the **implications globally?** For example, are journalists who use the platform in other countries going to benefit from mechanisms that cater to local languages?

20. Do the responses maximise the **openness and availability of relevant data,** with due regard to personal privacy protections, held by the social media companies (e.g., evidence of incidents reported; rates of response; time taken to remove content deemed to be in breach of policies; and justifications for content to be removed or retained)? Do they enable **independent research and reportage about the scale** of the problem and the companies' responses to it?

21. Is there **assessment** (informed by expert advice) of both the **potential and the limits of automated technological responses** to gendered online violence (while keeping freedom of expression and privacy rights intact)?

22. Are **civil society actors** (including NGOs and researchers), **women's advocacy groups, and the news media** engaged as **autonomous partners** in regard to combating online violence through knowledge sharing and facilitation?

23. Are the response measures accompanied by initiatives, programmes or campaigns designed to effect and embed change in the medium to long term, rather than being short term measures?

24. Do the responses proposed to address online gender-based violence against journalists work in tandem with disinformation responses, recognising the frequent links, intersections and overlaps between the two?

25. Do the responses entail parallel measures to combat misogyny, structural sexism and patriarchal norms present in the social context offline?

Recommendations for Action

The following research-based recommendations are proposed for key responders to online violence against women journalists globally. They build on the chapter analysis of the role of news organisations, the internet companies and State actors, as well as recommendations in the companion country studies published in Appendix 1.

Cross-cutting recommendations for multiple stakeholders:

1. Recognise that online violence is **"real",** as is the **psychological injury** it causes.

2. **Avoid blaming women journalists** for the online violence they experience, and do not expect them to bear the responsibility for managing or combating the problem.

3. Recognise that the problem of online violence manifests itself in the context of **powerful and wealthy internet companies** that should be held to account for enabling, and responding to, threats, harassment and hateful abuse directed at women journalists.

4. Facilitate and encourage **coordinated, global multi-stakeholder cooperation** and exchange of good practice **between States, internet companies and news organisation**s in the interests of effective implementation of holistic measures.

5. Foster and fund **collaborative** responses involving civil society organisations, journalists' networks, news organisations and researchers to gain more granular **knowledge** about the problem.

6. Ensure that **research** conducted on the issue is genuinely **independent,** and that it covers the **fast-moving nature and scale** of online violence and social media company responses, as well as addressing abuse on **closed networks** (e.g., private messaging apps and direct messaging platform functions) and coordinated cross-platform trolling.

7. Recognise the **intersectional** threats associated with gendered online violence, such as **racism, religious bigotry, sectarianism, antisemitism,**

homophobia and transphobia, and respond accordingly - including through policy development and training.

8. Encourage and aid women journalists to **report** online violence to the social media companies, their employers, and law enforcement where appropriate, while also recognising the priority need to **emphasise preventative and protective responses** that do not rely on the targets of abuse to manage the problem.

9. Encourage development of more **effective responses,** and ensure that these are **aligned with international human rights standards,** per the 25-step protocol outlined above.

News organisations should:

10. Recognise gendered online violence as a **workplace safety issue** experienced by their journalists (whether staff or freelance). Understand that **this applies regardless of whether or not the abuse is directed** at the journalist on their own news website, or a digital service owned by a third party.

11. Ensure that online violence is understood as **"real"** and that **psychological injury** suffered by women journalists under attack is recognised as **serious.**

12. Acknowledge the **increased intersectional risks** and impacts facing women journalists at the **nexus of misogyny, racism, religious bigotry, homophobia, transphobia** and other forms of discrimination which require recognition in editorial guidelines and online violence response protocols.

13. Recognise the correlation (and potential causal link) **between online violence and offline attacks,** and respond accordingly by ensuring that defensive strategies integrate physical safety, digital security, psychosocial support (including access to specialised trauma-aware counselling), editorial responses and legal assistance.

14. **Avoid making women journalists responsible** for their own protection and defence.

15. Develop or improve **newsroom protocols** that can address online violence against women journalists, recognising their **additional exposure** to risk.

16. Such protocols should be designed to **identify, monitor, prevent and respond** to online violence. They should be sensitive to intersectional threats, and it should also be regularly reviewed so that it is **responsive to the changing nature** of gendered online violence.

17. Ensure that these protocols take account of **contexts** of weaponised social media platforms, viral **disinformation,** far right extremism and conspiracy

networks.

18. Establish procedures and systems that **cover both staff and freelancers.**

19. Call on **law enforcement officers** to protect those targeted and prosecute perpetrators in those countries and circumstances where it is safe to do so (bearing in mind intersectionalities and global realities).

20. Collect and analyse **data** related to online violence and its effects, as experienced by staff and freelancers, and create a **gender-disaggregated database** of specific occurrences, and any follow-up. Use this internally to keep protocols under review, and make it **available** to trusted researchers internationally, so that the changing nature of attacks can be monitored, and the efficacy of remedial action can be evaluated.

21. Provide **targeted education and training initiatives** to staff and freelancers.

22. Appoint a **Digital Safety Editor** with capabilities and responsibilities that bridge editorial functions, digital security, and journalism safety. This position should include selection criteria that reflect the need for gender-awareness and understanding of intersectional threats and impacts.

23. Assign a **point person/team** to deal with the monitoring and reporting of attacks across platforms, private messaging, email, and across different devices when a woman journalist is under attack.

24. Lead from the top: Create a **company culture of gender equality and zero tolerance** for threats and harassment (online or offline) against staff, or women journalists at other outlets.

25. Put in place clear and transparent procedures related to **content and comment moderation on corporate websites,** along with clear **community guidelines,** and **train** relevant staff accordingly. Apply these principles - where possible - to the social media communities created and curated by the news organisation.

26. Hold social media companies to account through **investigative reporting,** and through **advocac**y for media freedom and journalism safety, regardless of commercial ties to the platforms.

27. Use **investigative and data journalism as countermeasures** to both raise awareness of gendered online violence, and to investigate and expose perpetrators (including orchestrated and/or State-sponsored attacks).

28. Ensure that coverage **avoids inflaming online mobs** targeting women journalists by amplifying and legitimising their attacks.

29. Avoid "victim-blaming" and speech-restrictions when responding to gendered online violence cases, recognising that the **target is not to blame** for the

abuse, harassment, or threats to which she is subjected. **Empower** her to speak, recognising that **"don't feed the trolls" is an inadequate response.**

30. Ensure that **policies on social media use represent a 'two-way street'** - where the obligations of the journalist to behave professionally on social media are **matched by a commitment to support and defend** her when she comes under attack.

31. Work **collaboratively** with other media organisations, professional associations and civil society organisations to monitor online violence, create robust integrated **models of risk assessment,** evaluate **recovery models,** and create **industry-standard guidelines, support systems and training.**

32. Lobby governments to formally recognise that online violence directed at journalists is an **attack on freedom of expression (including press freedom),** and that it has a **disproportionate impact** on women and marginalised journalists.

33. Lobby social media companies to recognise the **special needs and status** of women journalists – with sensitivity to intersectional risks – and introduce **rapid response units** focused on the safety of journalists, with human points of contact.

34. Support regulation to **make social media companies accountable** for the safety of women journalists on their services.

35. Act on the **November 2021 recommendation from the Council of Europe Expert Group on Action against Violence against Women and Domestic Violence,** which encourages media organisations and journalists unions to: "take concrete steps to **eradicate** gender-based discrimination, **victim-blaming attitudes** and **violations of the privacy** of victims of gender-based violence against women and their children in all their journalistic activities. Further efforts should be undertaken to **uproot male-dominated power dynamics** in media landscapes".

Internet companies should:

36. Continuously review their **policies, algorithms and moderation processes,** to address the evolving nature of gender-based online violence, while working closely with women journalists and civil society groups to **co-design new solutions.**

37. Develop more **sophisticated abuse reporting systems** with capacity for **escalation** for women journalists under attack (and their employers), recognising their particular vulnerabilities along with the implications for press freedom.

38. Implement a **coordinated multi-stakeholder approach** to protecting women journalists from online harms, which brings together all platforms, female journalists, civil society, news organisations, governments, and independent experts - at national and international levels.

39. Initiate **platform-platform cooperation,** since online violence often jumps across platforms and exploits the weaknesses of each.

40. Implement proactive countermeasures which **reverse the onus** on women targets having to report online violence to start with. This might involve using **human moderators and artificial intelligence technology** to more effectively filter out threats, abuse and harassment at the point of origin.

41. Retain **data documenting attacks** to aid targets wishing to access and use it for research or legal actions. Such proactive steps could link to monitoring processes to develop an 'early warning system' so as to better protect women journalists at the outset, or in the midst of an attack.

42. Build shields that enable users to **proactively filter abuse which could be quarantined for review and response.** Such systems should also provide prioritised pathways for women journalists under attack and news organisations seeking to report online violence.

43. Provide authorised independent researchers with secure and privacy-preserving **access to archives of moderated content and user appeals** in a standardised format, to enable **transparency and independent audits of moderation decisions** about threats made to women journalists.

44. Use the findings of such independent audits to **adjust both human and algorithmic moderation practices,** to strike a better balance between protecting freedom of expression and prohibiting abuse.

45. Implement an effective **human-in-the-loop approach** to content moderation coupled with a timely and effective appeals process - including effective **systems to appeal against company refusals** to act against online violent content and perpetrators.

46. Report transparently on **how human moderators and artificial intelligence algorithms are trained** to detect online abuse.

47. Define effective **policies for detecting and penalising repeat offenders,** to stop the same abusers assuming new online identities after action taken such as suspension or de-platforming.

48. Develop **markers for abuse perpetrator accounts,** similar to systems used to identify disinformation purveyors.

49. Establish **clear and transparent community rules** on what constitutes online violence and cease making exceptions **for influencers, public figures and**

242

other high-profile actors, whose high number of followers makes it easy for them to instigate abuse pile-ons.

50. Create **more effective content moderation tools** that provide sufficient support for all languages in which their services are offered (including vernacular or slang), and which are sensitive to contextual and cultural norms.

51. Technical solutions should be supported by **human contact points** who are familiar with a country's cultural, linguistic, and religious context and are well versed in local languages. These people should also possess press freedom, gender and journalism safety **expertise,** and be able to assist women journalists under attack.

52. Establish **task forces** and carry out **proactive programmes** to protect women journalists from certain abuse types, such as the dissemination of intimate images and doxxing.

53. Take effective **steps against the use of bots, false accounts and sock puppet networks** to prevent coordinated attacks and pile-ons that are frequently used in targeted online violence against women journalists.

54. Conduct **regular human rights impact assessments** as well as retrospective studies into the problem, including review of company policies and responses to gender-based online violence, and make the findings **public.**

55. Provide **detailed transparency reports on actions taken** against online violence against women journalists, broken down on a national level and including meaningful quantifiable metrics, beyond the total number of accounts removed and posts moderated. Reports need to also include **appeals and their outcomes,** along with **data** about notifications and responses to online violence reported by women journalists. They should also include **statistical representation and analysis of content that stays up after being reported** by journalists as abusive, offensive or threatening - not just on what is taken down.

56. Monitor the **intersectional nature of attacks** on women journalists who are targeted more than others because they belong to religious or ethnic minorities, indigenous groups or identify as members of the LGBTQ community.

57. Strike a **better balance between supporting freedom of expression and prohibiting online violence,** and recognise that international human rights law requires that women journalists be able to work online free from threats and harassment.

58. Support **independent research** (i.e. with no strings attached) on campaigns of violence against women journalists, and responses to these.

Individual States should:

59. Ensure that laws and regulations that could protect women journalists offline are **applied equally online.**

60. As urged by UN GA A/RES/74/157 (2019), collect and analyse "...concrete quantitative and qualitative data on online and offline attacks or violence against journalists, that are disaggregated by, among other factors, sex...". Create a **national evidence database** tracking perpetrators of online violence against women journalists.

61. Consider introducing **protocols and guidelines to restrain elected representatives, their staff, and other officials** who engage in gendered online violence against women journalists, with punitive measures attached, and **ensure prosecution** of those who perpetrate attacks. (See also recommendations for political parties below).

62. Make social media companies more clearly accountable for combating online violence against women journalists. Arrive at a **clear legal definition** of what social networks and messaging services are, and how they are regulated under national laws, with a view to regulating for the protection of women journalists and other human rights defenders working on these platforms (in alignment with the 25-step protocol presented above).

63. Consider **taxing social media companies** to provide revenues that could help **fund the work of monitoring, protection and training** relevant to online attacks on women journalists.

64. Make the companies more clearly **accountable** even in countries where these entities are not directly incorporated. This could include **a requirement to provide adequate reporting and response mechanisms in the languages on their services,** as well as adequate provision of a **timely appeals mechanism** and recourse to an independent national **ombudsperson** to help arbitrate cases where platforms and journalists cannot reach a settlement.

65. Regulate for the availability and comprehensive functionalities of **tools that enable users to easily report** online violence to the platforms and escalate appropriately, but ensure such regulatory and legislative interventions **respect freedom of expression** (refer to the 25-step protocol above).

66. Require social media companies to **notify users** who have reported online violence, on what actions have been taken, when and why/why not. These responses could include **referrals to informed civil society organisations and effective resources** (e.g., the Online Violence Response Hub)[233].

67. Introduce clear and effective **transparency regulations** for the companies with respect to: **gender disaggregation** in their reporting content moderation statistics; changes in **detection and moderation algorithms;** the number and

---

233 One such hub is co-convened by IWMF and ICFJ: https://onlineviolenceresponsehub.org/

types of **notices received and acted upon** in a given period; the **volume and topics of local content** that have attracted labels, distribution **restrictions, warnings, demonetisation** measures, or content that has been **removed or restricted** in circulation, and the **numbers and types of users who have been suspended or de-**platformed. Additional useful data points could include the number of users and engagement on a national level, as well as revenues in the national market.

68. Regulate to require **transparent and gender-disaggregated reporting regarding 'takedown' notices** connected to targeted online violence against women journalists, and protection of victims of doxxing and the distribution of sexual imagery shared non-consensually.

69. Establish or reinforce **independent national bodies/regulators** to oversee compliance with the relevant national and international laws and regulations designed to defend the safety of women journalists.

70. Introduce regulation that provides victims of online violence with **access to appeals against company (in)action** through an independent, national ombuds facility.

71. Regulate against the **social media 'black market',** which enables coordinated attacks **through sale of** accounts, views, likes, and comments.

72. Strengthen **labour laws and universal health care** to help support women journalists, especially those in precarious employment, when they are targeted in online violence campaigns which involve attempts to get them fired from their jobs.

73. Remedy possible jurisdictional issues by allowing **legal action based on the victims' location,** rather than the alleged perpetrators', to allow for action against harassment that originates in different locations.

74. Consider **introducing legislation** such as Ireland's Harassment, Harmful Communications and Related Offences Act 2020, which criminalises the publication and distribution of threats or "grossly offensive" messages with the intention to cause harm. (Any such legislation should **reflect the 25 principles for preserving freedom of expression in the context of legislative countermeasures** that are laid out above, emphasising **transparency, necessity and proportionality).**

75. Regulate, where needed, to **preserve the anonymity of complainants** and offer **closed court proceedings** for trials, to encourage more targets of gendered online violence (including acts of 'revenge porn') to come forward without fear of drawing further attention to the abuse;

76. Review the utility of **'shield laws'** that protect third-party internet platforms hosting harassing content from civil liability.

77. Ensure **hate speech legislation covers both gender and sex** (in addition to race, ethnicity, religion, and sexual orientation) to combat misogynistic expressions of online violence, and provide access to additional **opportunities** for legal redress for women journalists subjected to misogynistic hate speech.

78. Review laws in order to deal with 'pile on' forms of harassment through a **'proportionality' requirement** in online harassment, indicating whether a one-off comment could cause lesser or greater harm to the victim.

79. **Criminalise doxxing** and threats to dox women journalists.

80. Allow legal action on the basis of **complaints from third parties** (e.g., bystanders or employers) to **avoid the onus** being on the victim of gendered online harassment to file a complaint.

81. Help fund **pro bono legal services** specially equipped to deal with gendered online violence, so as to alleviate the costs of litigation, and increase the likelihood of successful court action brought by women journalists against online violence perpetrators.

Political parties and other political actors should:

82. **Desist from mounting attacks** (on- and offline) on women journalists, recognising that such conduct can trigger or dangerously inflame threats to their safety.

83. Develop **policies, procedures and guidelines requiring party members and officials to avoid instigating, facilitating or fuelling attacks** against women journalists.

84. **Sanction members and officials** who take part in acts of online violence in general and particularly against women journalists.

85. Introduce **training modules for party members,** including highlighting responsibilities as stakeholders.

Law enforcement agencies and judicial actors should:

86. Acknowledge the connection between online violence and offline harm for targeted journalists, including the risk of escalation to sexual assault and murder, but also serious psychological injury.

87. Understand that online violence is not 'virtual'. It does not have to inflict physical harm to be serious: it causes significant pyschological injury, economic impacts, and reputational damage.

88. Participate in expert-led education programmes for judicial actors and law enforcement agents to improve their media and information literacy as regards digital freedom of expression and the implications of online violence for press freedom and the safety of women journalists.

89. Participate in expert-led education programmes for law enforcement officials, including police, on the best gender-aware responses to initial reports of targeted online violence against women journalists.

90. Improve social media literacy to support basic knowledge of the operation of contemporary digital media systems, and develop basic digital investigative skills.

91. Recognise the targeted harassment of women journalists online as a workplace safety issue.

Civil society organisations should:

92. Reinforce the **call for effective responses** to online abuse of female journalists, and monitor how these conform to **international human rights standards.**

93. **Partner** with journalists, news organisations and researchers on investigative and monitoring projects about online violence and responses to it.

94. Help **educate** women journalists in particular in online safety, effective use of technological tools, engaging in employer advocacy and securing legal support.

95. Collaborate on the development of a **global online violence response hub** (see footnote 152).

96. Commission **interdisciplinary big data case studies,** modelled on those produced in parallel with this research, to inform and tailor responses for **diverse national, regional, cultural and linguistic settings, and intersectional experiences.**

97. Ensure **all media development and journalism safety programmes** and projects feature r**esponses relevant** to the threat of gendered online violence.

98. Work with the companies to establish a **cross-platform response system** to high-level threats against women journalists, recognising the cross-platform nature of abuse and its implications for offline violence.

Intergovernmental organisations, including UNESCO, should:

99. Ensure that **mechanisms and protocols** to defend the safety of journalists and end impunity address online violence against women journalists, including use of **cross-border and cross-platform** dimensions.

100. Monitor, research, record and publish evidence, and speak out against, online violence associated with crimes against journalists, with findings presented in a **gender-responsive and gender-disaggregated** manner.

101. Recognise and respond to the **problem of State actors using force extraterritorially** through online attacks on women journalists.

102. Consider initiating a **multi-stakeholder, research-informed 'early warning system'**[234] (drawing on datasets such as those underpinning the two big data companion case studies in this study) to trigger **interventions** (including from UN Special Rapporteurs), in cases where there is a significant and/or repeated risk to the target under attack online.

103. Ensure that **key officials** have an **appropriate understanding of causes and consequences** of online violence targeting women journalists, and encourage them to make appropriate representations to social media companies and political actors.

104. Consider a **UN-level conduit** to channel complaints against social media companies and State actors engaged in targeted online violence campaigns transnationally.

105. Develop and provide **gender-sensitive training and education for lawmakers, law enforcement agencies and the judiciary** to enable them to deal more effectively and appropriately with online violence against women journalists.

106. Monitor the implementation and effectiveness of member States' responses to gender-based online violence against women journalists, recognising that **legislative, legal and policy-based responses are one thing and implementation is another.**

107. Ensure that programmes focused on **media development** incorporate holistic education and training to deal with gendered online violence against journalists.

REFERENCES: An extensive list of references consulted for this study has been published by ICFJ.[235]

234 A comprehensive bibliography accompanying this study is published separately by ICFJ here
235 https://www.icfj.org/media/31615

Appendix 1: Country Case Study Summaries

This appendix presents summaries of 15 detailed individual country case studies produced by the international research team. Country-specific recommendations for action are provided at the end of each summary, drawing on the qualitative research underpinning the case studies.

1.1 Brazil

Luisa Ortiz Pérez, Carolina Oms, Eunice Remondini, and Kate Kingsford

[Online violence] is the new form of censorship. In the past, in the military dictatorship of our country, there was a classic thing that was the censor in newsrooms. Nowadays, you can censor in two ways: on the one hand, you flood social networks with misinformation... on the other hand, you run campaigns of intimidation, character assassination.

Patrícia Campos Mello, journalist at Folha de S.Paulo

This summary draws on a detailed case study about online violence against women journalists in Brazil, involving extensive desk research and in-depth interviews with 14 journalists and experts. This research provides qualitative insight (rather than generalisable data) which informs the country recommendations as proposed by the authors.

Challenges faced by women journalists

There is a trend of press freedom erosion in Brazil that has been further exacerbated by the COVID-19 pandemic. And in this context, online violence against journalists – particularly women journalists – appears to have surged. Black, Indigenous, and LGBTQ women are at particular risk of hate speech.

Brazilian politicians have accused reporters of spreading "fake news", modelled aggressive behaviours towards them, and fed an environment that encourages hostility towards the press. Violent attacks on journalists more than doubled in a single year: 428 attacks were recorded by the National Federation of Journalists (FENAJ) in 2020, compared with 208 in 2019. According to FENAJ, the main perpetrators are men in positions of power, often elected officials – including those at the most senior positions.

One significant example is the case of *Folha de São Paulo* journalist Patricia Campos Mello. After publishing a story on how digital marketing companies had stolen the identities of thousands of elderly Brazilians to spread propaganda during the election campaign, Campos Mello suffered a wave of extreme online attacks and personal threats. Her face appeared in pornographic montages, shared widely through social media, and she received messages telling her she should be raped. Campos Mello fought back with a social media campaign, supported by activist and feminist networks, and she took legal action against the current President

and one of his sons. The Brazilian judicial system decided in her favour for both cases in 2021.

Deep fake videos are a serious new weapon to target women journalists. Talita Fernandes, formerly of *Folha de São Paulo*, suffered severe online abuse after a video of her interview with the current President was manipulated to make her questions seem ridiculous, and to make it appear that she was attacking him. The video was widely circulated on social media. As a result, Fernandes experienced such extreme online violence that her health was affected and she could no longer continue working on daily news coverage of the Presidential Palace beat.

In Brazil, online abuse is disseminated through social media, most frequently on Twitter, often involving fake accounts, as well as WhatsApp, Facebook, and Instagram.

Institutional and legal response mechanisms

Brazilian women journalists may be reluctant to report acts of online aggression because of the hostility or disregard they face at the police station. This study surfaced evidence of cases where the police and judiciary in Brazil have considered online violence, including threats, to be unimportant. Some legal progress has been made, notably the 2018 Lola Act, created as a result of feminist blogger Lola Aronovich's long fight against online violence since 2011. When the literature professor reported this to the Federal Police, she heard that it was not their job to investigate this type of crime (Declercq, 2018). But the Lola Act now obliges them to act on any reports of misogynistic online abuse ('Lei Lola' - Presidência da República, 2018).

However, journalists may still be discouraged from seeking legal redress for online attacks because some investigations are dropped due to pressure from different types of perpetrators. Another impediment is the long and costly nature of the legal process. In response, women journalists have formed collectives to support each other both psychologically and with practical advice as they endure legal battles. A number of civil society organisations in Brazil carry out research and take part in activism to counter online violence. The FENAJ annual report, 'Violence against journalists and freedom of the press', published for the past 30 years, is a notable example, compiling data from both public reports and journalists' complaints to unions.

Recommendations for action in Brazil

The Brazilian authorities and lawmakers should:

1. Recognise the serious nature of online violence in Brazil, including intersectional aspects. In particular, recognise that online violence can cause serious psychological damage, leading journalists to self-censor, and that it has created a mental health crisis among women journalists that necessitates discussions about both gender discrimination and structural inequality.

2. Commit to working to stop online violence rather than encouraging it. Crimes against journalists often go unpunished because the perpetrators enjoy a level of impunity, particularly in the case of public officials.

3. Work with civil society organisations and internet communications providers to combat attacks against journalists, including by non-legal measures, such as gender policy and advances in digital media literacy.

Brazilian news organisations should:

4. Change the work culture to become more gender- and race- sensitive, in response to many women journalists who do not feel safe or supported at work.

5. Make paid leave, as well as psychological and legal support, available to journalists who are experiencing harassment or are in the process of a legal case.

International human rights organisations should:

6. Take strong positions against the use of political office to systematically undermine press freedom, and demonise, attack and instigate violence against women journalists in Brazil, on- and offline.

7. Create spaces for training, information gathering, and awareness raising, both in the newsroom and the general population, with due consideration for cultural and linguistic diversity.

Big Tech providing services in Brazil should:

8. Provide greater transparency on the type of detection and reporting policies and practices they are implementing, the number of reports they receive, and the number and nature of content they remove or other steps taken.

9. Establish task forces and more active programmes to protect women journalists from common abuse types, such as the dissemination of intimate images without consent.

10. Improve algorithms to better detect misogyny, sexism and gender-based threats, and support research into online violence campaigns.

1.2 Kenya

Fiona Chawana and Julie Posetti

In a few years, we will not be having women journalists in this country...[and] around the world, because you get to a point, you're so scared for yourself. And also, you're scared for your family. When journalists self-censor...the fundamental right to freedom of information is threatened.

Cecilia Maundu, broadcast journalist and digital security trainer

This summary draws on a detailed case study about online violence against women journalists in Kenya, involving extensive desk research and in-depth interviews with eight journalists and experts.

Challenges faced by women journalists in Kenya

In 2016, a survey of 100 randomly selected journalists conducted by the Association of Media Women in Kenya (AMWIK) and ARTICLE 19 East Africa found that most women journalists rely on technology platforms for their everyday work and yet these were also the platforms where they were most exposed to online violence. During the pandemic, this problem was exacerbated in the region, with nearly two thirds of 100 East African women surveyed by African Women in Media (AWiM, 2020) indicating they had experienced worse online attacks in the context of COVID-19.[236] Beyond the usual difficulties of reporting abuse on these platforms, Kenyan women journalists also face the added struggle of social media companies' reporting mechanisms and algorithms not recognising local and vernacular languages, or the societal and cultural nuances of subtle threats.

Perpetrators of this violence also come in the form of well-organised groups funded by political actors who troll and attack women journalists in an effort to silence them and undercut their reporting. The Oxford Internet Institute has identified Kenya as one of the countries where the use of 'cyber troops' is prevalent in spreading disinformation, particularly during election periods (Bradshaw and Howard, 2019), a phenomenon associated with peaks in online violence against women journalists according to several interviewees. Reporters Without Borders (RSF) has also documented attacks on major media outlets, including the 2019 trolling of the Nation Media Group, East Africa's largest independent media house, by pro-government activists (RSF, 2021d).

---

236 The survey was supported by UNESCO.

Entrenched patriarchal norms underpin the online violence experienced by Kenyan women journalists, with body-shaming being a prominent form of attack against them on social media. Interviewees also pointed to a sexist perception that advancement in women's equality has disadvantaged men, and online violence against highly visible women TV journalists is part of that backlash. The 2016 study mentioned above also found that "digital harassment leads to women withdrawing from using the internet, and in many cases they have stopped working for some time." This trend was confirmed through this research, with several interviewees indicating that they had quit their jobs or left journalism altogether, as a result of online violence. An associated reluctance among women journalism students to pursue careers in the media was also observed.

A 2019 report titled 'Highlighting and Eliminating Harassment and Abuse in the Media', focusing on Kenyan newsrooms, revealed that at least 94% of the 70 participants in this research were aware of sexual harassment occurring at their media organisation. A low number of complaints had been made to employers out of fear of victimisation and being labelled as a troublemaker (Otieno, 2019). Poor capability among media employers to deal with the significant mental health impacts from online violence described by the Kenyan interviewees for this study is also a factor. Employers may also further exacerbate the climate of online violence by encouraging journalists to stoke attacks, while at the other end of the spectrum, some employers restrict the social media activity of their journalists. However, there are some positive existing practices such as by *The Standard* newspaper, which supports an internal group of women journalists from all levels to facilitate discussion of gender policies in the workplace and sexual harassment.

Institutional and legal response mechanisms

The Kenyan Constitution secures press freedom in Section 34, which also underlines the conditions for restrictions such as: "privacy, incitement, hate speech, and anti-government propaganda in times of war" (AMWIK, 2014). The Media Council Act, adopted in 2013, mandated the creation of the Media Council of Kenya, which governs the operation of local and international journalists and all media organisations operating within Kenya, and media personnel are expected to abide by its guidelines. It also mandated the Council to create a Code of Conduct for the Practice of Journalism in Kenya, allowing members of the public to lodge complaints in cases of misconduct by journalists and media companies. This is problematically prone to abuse, but the Code also lays out procedures for journalists and media organisations to lodge complaints when their constitutionally guaranteed right to freedom of expression is infringed upon. The Kenya Information and Communications Act also established a Communication and Multimedia Appeals Tribunal, which is meant to protect journalists and consumers. This enables an aggrieved member of the public to report a journalist or media organisation for alleged abuses of power, which may be legitimately used but is also at risk of abuse.

Although the abovementioned regulatory mechanisms do not deal specifically with online violence, the Computer Misuse and Cybercrimes Act of 2018 sets out provisions inter alia regarding cyber harassment, cybercrime and child abuse imagery. However, it lacks gender specific guidance. Kenya is party to the African Charter on Human and People's Rights, has ratified the International Covenant

on Civil and Political Rights (ICCPR), and is signatory to the Convention of the Elimination of Discrimination Against Women (CEDAW).

A 2020 report commissioned by the Media Council of Kenya identified digital and physical threats against journalists, and their implications for freedom of expression. 14% of the 52 cases reported examined between March and October 2020 involved threats and attacks on women journalists, mainly by security forces and the public, including bullying, trolling, insult and abuse (Media Council of Kenya, 2020). While the report concluded that there was poor follow through by the police, and low levels of reporting to the police, it promoted the need for: "Media literacy forums, publications of statements in support of victims/to condemn violations, and facilitating police involvement in securing justice for victims as effective countermeasures in dealing with online harassment".

Recommendations for action in Kenya

The Kenyan authorities and lawmakers should:

1. Take action against political actors initiating and stoking online violence against women journalists in Kenya, especially those engaged in recruiting 'troll armies'.

2. Invest in gender-sensitive online violence-response training for a range of Kenyan actors.

3. Empower the Media Council of Kenya to act as a liaison between civil society, industry and journalists to develop gender-sensitive policies for newsrooms.

4. Develop a holistic strategy to address the press freedom and journalism safety threats posed by gendered online violence against journalists, including legislative, regulatory and educative responses that respect international freedom of expression rights.

Kenyan news organisations should:

5. More effectively address mental health impacts of gendered online violence on their staff and create a culture where women journalists can feel safe to report such impacts.

6. Create safe spaces for women to share their experiences of abuse and find mutual support, such as gender equality networks.

Social media companies providing services in Kenya should:

7. Actively invest in abuse reporting tools and develop content moderation capability for reporting online violence on their platforms in local languages, and with reference to local cultural context to address the current climate of impunity for online violence.

8. Develop technological aids to help women report cases of online violence as they happen in Kenya, in local languages targeted to a Kenyan usership.

9. Support journalism education and training, and invest in reporting projects addressing gendered online violence through arms-length investment to ensure independence.

10. Establish a physical presence in Kenya, employing local staff.

11. Enforce existing policy with a view to online harassment and develop a framework on gender-based online violence in line with Kenya's Computer Misuse and Cyber Crimes Act and the Data Protection Act (ensuring international freedom of expression standards are upheld).

Kenyan civil society should:

12. Help facilitate multi stakeholder engagement to assess and address the problem in a culturally and linguistically relevant manner.

13. Provide holistic digital safety training for Kenyan women journalists, both inside news organisations and within freelance communities.

14. Assist women journalists suffering from online violence with local legal advice and counsel to help them to navigate the national judicial system, following the due process of law.

15. Facilitate constructive and inclusive conversations which also engage male leadership within journalism, politics and society to help improve their gender-awareness and sensitivity, while creating bridges between genders.

1.3 Lebanon

Nermine Aboulez

[Harassers] try to strip you of your 'honour', because they consider that a woman's honour is between her legs...so, if you're coming from a conservative upbringing, you'd most likely quit the profession and/ or be silenced.

Youmna Fawaz, investigative journalist

This summary draws on a detailed case study about online violence against women journalists in Lebanon, involving extensive desk research and in-depth interviews with 10 journalists and experts. This research provides qualitative insight (rather than generalisable data) which informs the country recommendations as proposed by the authors.

Challenges faced by women journalists

Taboos around sexual harassment contribute to the reluctance of women abused and harassed – on- and offline – to report such incidents, as does the severe underrepresentation of women in newsrooms in a country with a very large gender gap in the labour market (UN Women, 2021). Social media platforms have provided spaces for support and solidarity for women journalists dealing with sexual discrimination and harassment, but they are also major facilitators of online violence against them. This takes diverse forms, ranging from name-calling to death threats; varying from isolated incidents to orchestrated attacks.

According to interviewees for this case study, online abuse experienced by Lebanese women journalists is almost always sexualised (designed to injure them 'morally' as well as professionally). It usually involves digital security threats, and it is frequently graphic, with threats to kill or "blow up" the targeted journalists being commonplace. Freelance journalist and editor Zahra Hankir was threatened online multiple times with rape and death for voicing her support for Loujain Alhathloul, a Saudi woman's rights activist jailed for three years. The interviewees also described relentless hacking attempts, and threats that radiated to family members. TV presenter Dima Sadek's daughter, who has a disability, was targeted, and photos of family members with identifying location information were sent to former investigative journalist Myra Abdallah as a threat.

Intersectional abuse is also common. Investigative journalist Youmna Fawaz has experienced online violence for 10 years. She has been called a "whore" and a "bitch"; falsely accused of sleeping her way up the career ladder and taking bribes; and bullied for a lisp. Her family members have been threatened. Three-decade

journalism veteran Diana Moukalled said she was bullied because of her age and for being a Shi'ite. Similarly, Dima Sadek was targeted with smear campaigns for being a Shi'ite and criticising the Shi'ite Hezbollah. Freelance journalist and safety trainer Luna Safwan who is also a Shi'ite says she has been branded as a traitor to her religion and country, as well as falsely accused of being an Israeli agent for criticising Hezbollah.

Using 'honour'-based threats to intimidate women journalists is another common tactic deployed. Examples include smear campaigns against Dima Sadek, including false claims that she is involved in making pornographic films. Al Jazeera's Ghada Oueiss was similarly attacked in large-scale orchestrated attacks which misrepresented her. Myra Abdallah said she has faced blackmail attempts involving threats to distribute naked photos of her online. The online violence to which Lebanese women journalists are exposed is not contained by the internet. Frequent doxxing increases the offline risks associated with online attacks. Several of the Lebanese interviewees had been doxxed, including Oueiss, who faces very significant threats of physical violence. Abdallah described an incident where someone threatened to put her in a car trunk where no one would ever hear from her again, and she survived a kidnapping attempt. Other journalists described being chased and physically assaulted after being recognised while reporting on protests. Dima Sadek talked about a religious figure calling for her death in a public sermon for being an "enemy of God" in the context of online pile-ons.

The severity of online violence against Lebanese women journalists has driven some out of the profession. Myra Abdallah said she suffered PTSD and quit journalism in response to the online violence she experienced; "the day I decided to stop writing was the day I was sent a picture taken from inside my house, that's when I decided that it was not worth it."

Institutional and legal response mechanisms

A number of Lebanese journalists interviewed for this study described feeling that they had to absorb the online violence they experienced because of structural inequality within the news industry and society. Some women credited the support of their fellow journalists and solidarity from prominent international defenders like former UN Special Rapporteur Agnes Callamard, for their survival. Others are taking legal action against perpetrators (e.g., Ghada Oueiss) as an act of self-empowerment, demonstrating their determination to fight back.

The interviewees identified news organisations as part of the problem, describing media managers as often treating gender-based online violence with disdain. Some male colleagues were even cited as significant sources of online violence.

Social media platforms' responses to online harassment evoke mixed evaluations from the Lebanese women journalists interviewed. While some said the platforms were responsive and deleted abusive comments, closed fake accounts, and/or protected the targeted journalists' accounts (e.g., Dima Sadek), others described responses as non-existent or slow. Ghada Oueiss said she had a "price on her head" posted on Facebook, calling for her to be kidnapped or killed and when she reported the incident, the company told her she would need to create a Facebook account before they could delete the 165 fake Facebook pages in her name. The

platforms were also criticised for failing to recognise how damaging some content can be in different cultural contexts, especially in religiously conservative societies.

The Lebanese legal system contains laws with caveats and overly-broad definitions that could be used as loopholes against journalists (Amnesty International, 2020b; RSF, 2014). The constitution grants freedom of expression "within the limits of established laws", prohibiting insulting public officials and national or religious symbols. The Electronic Transaction and Personal Data Act allows the suspension of electronic service, closure of accounts, and blocking of websites for sharing content deemed threatening to "internal and external security" (Social Media Exchange, 2018). The interviewees underlined that the Ministry of Telecommunication (which is also responsible for online issues) or the Lebanese journalists' trade unions could be much more effective responders to online violence. However, civil society organisations such as the Samir Kassir Foundation (SKeyes), the Social Media Exchange (SMEX), and the Legal Agenda, play a significant role in assisting women journalists under attack, along with international organisations.

Recommendations for action in Lebanon

The Lebanese authorities and lawmakers should:

1. Ensure that mechanisms designed to address digital security and safety threats encompass attacks on women journalists specifically, recognising their particular risk.

2. Bring to justice political actors who threaten Lebanese women journalists online, and publicly defend those under severe attack.

Lebanese law enforcement and judicial actors should:

3. Participate in gender-sensitivity training, and digital skills development to help them better respond to instances of gendered online violence against journalists.

Lebanese news organisations should:

4. Recognise the structural inequalities that make it harder for women journalists to acknowledge online violence and seek help from employers and develop gender-sensitive support mechanisms for women journalists under attack online, also recognising the threat of offline harm.

5. Develop clearly defined sexual harassment and mental health policies, as well as digital, psychological, physical and legal support for their journalists.

6. Offer integrated training in defence against gender-based violence (on- and offline), blending physical safety, digital security, psychological support, along with instruction in documenting incidents in order to support potential legal action.

Social media companies providing services in Lebanon should:

7. Act more quickly and effectively to stop online violence against Lebanese journalists on their platforms, and be more responsive when journalists report incidents, recognising the risk of offline violence this may lead to.

8. Employ locally based, specialised contact points, with Lebanese cultural knowledge and language skills, to respond to women journalists when they are under attack on their platforms, especially in the context of orchestrated attacks.

1.4 Mexico

Yennué Zárate Valderrama, Luisa Ortiz Pérez, and Kennia Velázquez

[My harasser] used hateful words, misogynistic messages, he harassed me. Two years ago he claimed I was involved in organised crime matters. He calls me cow, hyena, that I disgust him; he is the son of an elected official...Out of five journalists who denounced his attacks, one was murdered - how could I not be worried?

Lidia Alejandra López Castañeda

This summary draws on a detailed case study about online violence against women journalists in Mexico, involving extensive desk research and in-depth interviews with 14 journalists and experts.

Challenges faced by women journalists

In 2020, 2,819 women were murdered in Mexico, with 976 of these cases identified as femicides – women who were killed because of their gender (Zárate, 2020). Violence against journalists in general is equally alarming, with 123 killings of journalists recorded in the country between 1997 and 2020 by UNESCO (UNESCO Observatory, 2021d). Mexican media workers operate in an unsafe environment in which anti-press violence is linked to various actors (such as drug cartels, armed forces, and officials) and is worsened by a climate of impunity for these crimes.

In this context, women journalists face a double risk: high rates of killings of journalists, and high levels of violence against women. They are subjected to systematic attacks rooted in structural causes, such as widespread domestic violence, political violence, limitations on access to justice, high levels of impunity, human rights violations, and a culture of "machismo". Online violence against women journalists reflects gendered politics of power and the prevalence of organised crime, with systemic violence transposed into the digital realm. In a country where extreme levels of physical violence demand constant vigilance, digital safety is often overlooked and considered as less urgent, despite established links between both modes of gender-based violence. In 2020, 30% of reported aggressions towards women journalists took place online (ARTICLE 19, 2020f). Interviews for this study revealed three clear trends in attacks against women journalists: brutality, incompetent investigations in relation to the attacks, and impunity.

Digital attacks in Mexico are mainly observed on social media platforms like Twitter, Facebook and WhatsApp, but also via email and mobile phone messages. Methods of attacks include threats of physical and sexual violence (including death threats), doxxing, trolling, deep fake videos, identity theft, surveillance, smear campaigns, criminal harassment, sexual harassment, and censorship via unjustified removal of content by the platforms. Harassment is often misogynistic, sometimes combining classism and racism (particularly against reporters of Indigenous descent), although the professional conduct of women journalists is rarely questioned. These threats and attacks aim at silencing journalists who report on sensitive political issues and on gender-based violence. Those responsible for abuse are often powerful men and public officials.

Many women journalists who have experienced online violence resort to self-censorship, as they fear for their lives and for the safety of their families. Many choose not to report attacks because they fear an escalation of violence against them, and having their private lives scrutinised and exposed. At the level of news organisations, online violence affects their very survival, as State-sponsored rhetoric (at federal, state or municipal levels) against specific women journalists can lead to the withdrawal of government advertising or defamation, resulting in a loss of credibility and revenue.

Institutional and legal response mechanisms

Mexico has an integral legislative framework to protect journalists and freedom of expression, yet in practice fails to implement it efficiently. According to Reporters Without Borders (RSF), collusion between different authorities and organised crime "cripples the judicial system at all levels" (RSF, 2021k). Structural issues in the Mexican criminal justice system lead to flawed and incomplete investigations into attacks against journalists, and in turn result in a climate of impunity. Due to the presence of public officials among the culprits of online violence and high levels of corruption among local police, women journalists seldom file complaints for fear of retaliation.

The right to freedom of expression including access to information is enshrined in the Mexican Constitution, as well as in the 2002 Law on Transparency and Access to Public Information. The 2010 Inter-Institutional Collaboration Agreement and the 2012 Law for the Protection of Human Right Defenders oblige authorities to provide protection for journalists. With regard to women's rights, the 2007 General Law on Women's Access to a Life Free of Violence categorises different forms of gender-based violence and establishes courses of action. The Mexican Penal Code provides for an 'Alert on Gender Violence against Women' mechanism, which provides protection orders and sets guidelines for the care that is to be provided to victims. In 2021, online violence was included in both sets of legislation (Reuters, 2021; Redacción Animal Político, 2021). The country's 'Olympia Law' on Online Safety covers crimes against privacy and cyberbullying, but does not cover other forms of online violence.

In 2010, a Special Prosecutor's Office for Crimes against Freedom of Expression (or "FEADLE") was created, with the authority of directing investigations and prosecuting crimes committed against media workers. Yet, the UN and OAS Rapporteurs on Freedom of Expression noted in 2018 that FEADLE had "not included [a] gender

perspective in its work to better deal with crimes against women journalists" (Inter-American Commission for Human Rights 2018, paragraph 46). On digital violence, they observed "ineffective investigations of threats and harassment of journalists online and offline". Furthermore, FEADLE statistics show that only 118 of the 790 investigations opened since 2010 reached the courts, and that no FEADLE-led murder investigation has resulted in a conviction.

Mexico is also equipped with a Mechanism for the Protection of Journalists and Human Rights Defenders, which provides police protection for journalists at risk. However, this mechanism does not directly address digital violence. Out of 11 states with laws that have established protection mechanisms, only two have links to this Mechanism. 11 States have unformalised protection initiatives, and seven have no proposed legislation (SinEmbargo.MX and democraciaAbierta 2020). Moreover, only 16 out of 32 Mexican states have "digital police", and there is limited coordination between these units and prosecutors' offices. According to the UN Special Rapporteur on Freedom of Expression in a 2017 report, the Mexican government has "failed to devote necessary resources and demonstrate political will. The mechanisms of protection have surprisingly limited resources given the context of a national crisis. The specialised mechanisms of accountability, at federal and state levels, have not delivered results".

Insufficient responses have also been identified within newsrooms, with interviewees assessing that the newsroom culture tends to disregard online violence as insignificant in relation to the physical threats journalists face, and few news organisations provide physical or digital protection for their staff. Social media companies have also failed to create a safe environment for women journalists with their inaction fuelling impunity for crimes against journalists.

There are several NGOs providing support for women journalists who have experienced online violence through initiatives such as providing workshops and toolkits/manuals on physical and digital security, education and specialised trainings on freedom of expression and gender, connecting victims of digital attacks with support services, as well as conducting research and analysis on digital violence against Mexican women, judicial responses and the persistence of impunity. These include ARTICLE 19, Luchadoras, R3D, SocialTIC, Internet es Nuestra, Derechos Digitales, APC, CIMAC, Sursiendo, or the Latin American feminist cyber-collective Ciberseguras.

Recommendations for action in Mexico

The Mexican authorities and lawmakers should:

1. Independently investigate FEADLE failures in homicide investigations and encourage local authorities to promptly investigate reports of online violence with appropriate digital forensics to aid cyber police units.

2. Consider introducing significant penalties for public officials and office holders implicated in online violence against women journalists.

3. Design improved mechanisms for holding social media platforms accountable for facilitating online violence in such a high-risk context.

4. Support coordination between national and local protection mechanisms and address online violence in tandem with physical violence.

5. Penalise public officials who instigate or amplify online violence against Mexican women journalists.

Mexican law enforcement and judicial actors should:

6. Allocate resources and digital infrastructure to train police officers responsible for investigating online violence reports and improve access to mechanisms for reporting and responding to violence against journalists, both locally and nationally.

Mexican media organisations could:

7. Develop integrated safety policies and mechanisms, recognising the interplay between different kinds of threats and the risk of online violence spilling offline in a country with extreme levels of violence against journalists and women.

8. Fund and/or provide gender-sensitive digital security training that integrates offline defences.

9. Provide or fund legal support and consider strategic litigation in serious online violence cases.

10. Allow women journalists experiencing online violence to take leave from work through financial support.

11. Address digital violence against Indigenous women who face intersectional risks.

Social media companies providing services in Mexico should:

12. Create local units to detect gendered online violence against journalists, aid official investigations, and deplatform repeat offenders.

13. Simplify reporting mechanisms; establish clear and transparent community guidelines and rules of enforcement that respond to specific cultural factors, such as high rates of femicide.

14. Be more transparent about strategies to address gender-based online violence.

Civil society organisations in Mexico should:

15. Include digital threats in monitoring and documenting cases of digital attacks on Mexican women journalists and strengthen alliances and collaborations between academia, civil society and the journalistic profession in order to produce research and advocacy.

16. Apply an intersectional lens when monitoring and responding to online violence, recognising that Indigenous women in Mexico and others who are additionally marginalised face more significant online violence threats.

1.5 Nigeria

Omega Douglas

I'm not someone who easily ignores things, so it was difficult ... expending my energy trying to respond to trolls. But I sort of got the hang of it, which is so sad... No one should get the hang of harassment.

Kiki Mordi, journalist and founding member of the Feminist Coalition

This summary draws on a detailed case study about online violence against women journalists in Nigeria involving extensive desk research and in-depth interviews with ten journalists and experts.

Challenges faced by women journalists

Even though Nigeria has one of the largest and most diverse media scenes on the African continent, repression by several actors, including officials, still occurs. Incidents have been reported of members of the police, the army and other governmental officials harassing, publicly criticising and detaining journalists (Amnesty International, 2019; Freedom House, 2019b; Paradigm Initiative, 2019). Moreover, Nigerian women journalists are affected by strong patriarchal norms, which spills over into their everyday work (Makama, 2013; Okonkwo, 2013; Okonkwo-Ogabu, 2020). Journalist Ruona Meyer, interviewed for this study, was prolifically targeted online in the context of her work and the abuse she experienced was focused on perceptions that she breached ideals of Nigerian womanhood.

In Nigeria, women are underrepresented in news media organisations, even more so in leadership roles, and sometimes do not have formal complaint mechanisms for the sexism and sexual harassment they face. They are also sometimes confronted with 'beat politics', in which news categorised as 'hard' brings more financial and career related benefits, and yet women who are assigned to 'hard' news beats may not be included by, or receive the same advantages as, their male counterparts (Tijani-Adenle, 2019). Furthermore, to be accepted as 'one of the boys', women journalists may be expected to deal with any difficulty the job might bring, without complaint (IPC, 2015).

Reporting on gender-related news or applying a gender lens to the issues they are covering can lead to women journalists being attacked or silenced online. When journalist Kiki Mordi, along with her colleagues, tried to draw attention to the gendered nature of brutality by members of the police in Nigeria - particularly regarding the #EndSARS protests - they encountered online abuse from various quarters.

Institutional and legal response mechanisms

Within 15 media organisations surveyed by the Wole Soyinka Centre for Investigative Journalism (WSCIJ) between 2018-2019, 53 out of 72 staff respondents said they had no internal policy on gender, and of those that did, none had sanctions for breaching gender policy requirements. Some media organisations run digital security training for their journalists and have lawyers and psychologists to support them in the context of online attacks. Other organisations, in an already economically precarious industry, struggle to find the funding. Yet, without investment in addressing the sexism and online violence they face, women journalists in Nigeria may feel compelled to leave the industry, placing news organisations in an even more precarious situation.

Nigeria is a member of the Group of Friends on the Safety of Journalists, an informal network of countries collaborating to support the UN Plan of Action on the Safety of Journalists (UNESCO, 2018), and has led the African region by sponsoring six of the eight resolutions addressing the safety of journalists adopted by UN bodies since 2012. Nigeria is a signatory to several international treaties on women's equality and its 1999 Constitution guarantees everyone the right to freedom from discrimination on the grounds of sex and recognises women's equal rights. However, laws are not always adequately implemented. There is a lack of clear legal frameworks for safeguarding women online and even with legal grounds, cases can be complicated and costly for the victim (World Wide Web Foundation, 2015).

The 2015 Cybercrime Act has also attracted controversy. Although the Act's Section 24[237] could, in theory, be viewed as a legal mechanism to be utilised by women journalists experiencing online harassment, this law is also able to be exploited to prosecute journalists, bloggers and activists.[238] The Digital Rights and Freedom Bill, drafted through a collaboration of lawmakers and the Paradigm Initiative, aimed to provide legal recourse for online harassment. The President, however, declined to sign the bill into law in March 2019. Even with clear policies on abuse, women journalists who have reported content to social media platforms say they have rarely received an adequate response, nor seen these policies upheld. Ruona Meyer reported the abuse she had experienced to Twitter but she said she never received a response from the company. Additionally, the interviewees proposed that the reporting process and algorithms used to detect and respond to hate speech need to be localised, as some offensive contextual situations for Nigerian society are not factored in.

Facing these obstacles, women journalists tend to eschew legal and institutional mechanisms and instead tell friends and family about digital harassment, block contacts on social networks, confront the perpetrators, or do nothing. In terms of civil society responses, organisations such as Wole Soyinka Centre for Investigative Journalism (WSCIJ) run programmes to equip women journalists with knowledge and skills to challenge the gender inequality in their industry, including mentoring, networking, and training opportunities to enable them to reach leadership positions

---

237 Section 24 penalises 'cyberstalking' or messages that are 'false, for the purpose of causing annoyance, inconvenience, danger, obstruction, insult, injury, criminal intimidation, enmity, hatred, ill will, or needless anxiety to another,' and is opposed by several civil society organisations.
238 https://freedomhouse.org/country/nigeria/freedom-net/2019

within their organisations, mobilise networks of women journalists, and build gender parity within Nigerian newsrooms.

Recommendations for action in Nigeria

Nigerian authorities and lawmakers should:

17. Support initiatives, like the Wole Soyinka Centre for Investigative Journalism's (WSCIJ) Female Reporters Leadership Programme, that seek to improve the number of women in leadership positions in Nigerian media organisations and better equip them to deal with online violence.

18. Inject resources into the news industry specifically targeted at improving digital defence capability and gender-based online violence awareness.

19. Enact specific laws and policies to protect women journalists from online harassment, including the long-debated Digital Rights and Freedom Bill.

International organisations should:

20. Support the establishment of a centralised emergency fund which could cover legal fees for cases of online harassment in situations where journalists and/or the organisation they work for can't afford to cover costs.

21. Advocate for an end to attacks against journalists by the authorities, particularly by members of the police and security forces, which go against national and international legal and normative frameworks that stipulate Nigeria's obligation to protect freedom of expression, and the rights of women online.

Media organisations operating in Nigeria should:

22. Work to ensure greater representation of women at every level of media institutions to help foster more opportunities for the implementation of gender responsive policies in addressing online violence.

23. Work collaboratively to share effective training and examples of good practice, such as providing digital security training as well as the services of psychologists and lawyers should a journalist come under attack.

24. Provide funds, time off work with pay, for those under attack.

Nigerian civil society organisations should:

25. Collaborate with lawyers, media houses and activists to improve effectiveness of existing frameworks in Nigeria.

26. Hold the government accountable by proactively supporting the Digital Rights and Freedom Bill, and other policies that could help to address online violence of women journalists in Nigeria.

1.6 Pakistan

Ayesha Jehangir and Fiona Martin

They attack us for 'honour', they attack us for our bodies, our looks, our femininity, our values.

Yusra Jabeen, journalist

This summary draws on a detailed case study about online violence against women journalists in Pakistan, involving extensive desk research and in-depth interviews with 10 journalists and experts.

Challenges faced by women journalists

Digital attacks against women journalists are a critical issue in Pakistan.[239] A 2019 survey of 110 women journalists by Media Matters for Democracy (MMfD) found that 90% of women journalist respondents (99 individuals) reported experiencing some form of online violence, mostly via Twitter, Facebook and WhatsApp (Kamran, 2019).

This phenomenon occurs against a backdrop of widespread gender-based discrimination and violence (D'Lima et al., 2020; Baig et al., 2020; Human Rights Watch, 2019a). Research shows that women who breach modesty norms, try to have social agency, or make decisions for themselves, are particularly exposed to reprisals (D'Lima et al., 2020). Online abuse is one method used to prevent women journalists from asserting their social position and trying to influence public discourse – especially on politics, religion, terrorism/extremism and war/conflict/militarisation (Khan, 2020; Rehmat, 2017). The murder of Baloch journalist Shaheena Shaheen, in September 2020, attracted attention to the prevalence of violence against women reporters and triggered countrywide discussions about how it could be addressed.

Several interviewees perceived that online violence was becoming more intense and coordinated, including that there was "state-orchestrated" online harassment triggered by critical reporting. Research shows that some digital attacks against Pakistani women journalists are committed by officials or political party supporters (CFWIJ, 2020b and 2020c), and increasingly associated with police and legal harassment (CFWIJ, 2021a and 2021b).

Online violence against women journalists manifests in myriad forms, including shaming them for their bodies, morals, looks, relationship status, and ideologies – a phenomenon that lawyer and non-profit executive Nighat Dad calls 'social surveillance'. Reputational attacks are widely used – e.g. accusing the targets of prostituting themselves, or committing blasphemy, a very dangerous allegation in a

---

239 Context includes physical attacks on women producing digital content. See: https://www.theguardian.com/world/2021/aug/19/hundreds-of-men-in-pakistan-investigated-over-sexual-assault-caught-on-tiktok

country where it carries a maximum penalty of death. On top of coordinated mass abuse, threats of murder or sexual violence, data or identity theft, blackmailing, and other forms of online violence, the interviewed women journalists noted the increasing use of fake social media profiles to stalk, harass and denigrate them, a trend also confirmed by other recent research (Kamran, 2019).

The psychological impacts of online violence are heightened by the existence of widespread physical violence against women in Pakistan, and knowledge that online attacks can lead to offline assaults. Many women journalists respond by disengaging from interactions with their abusers, deleting posts, or retreating from the digital space until the situation calms down. Some quit social media or journalism altogether. Many self-censor their reporting, steering away from covering issues such as politics, war, terrorism and religion (Jamil, 2020; Kamran, 2019; Khan, 2020). The 2019 MMfD report confirmed that 77% of respondents (85 individuals) said they had self-censored as a result of online harassment (Kamran, 2019). Another report found that over 21% of the 1,400 women respondents (294 individuals) said they had stopped using the internet altogether (DRF, 2017a). Considering that historically a high preponderance of Pakistan's Internet users have been male (DRF, 2016), this type of outcome is a major setback to freedom of expression.

95% of respondents to the MMfD survey (105 individuals) also indicated that online abuse had affected their personal relationships in the offline environment (Kamran, 2019). They experienced problems with partners, other family members and friends. Some journalists interviewed for this case study indicated that reputational attacks had undermined their credibility and their access to officials.

Institutional and legal responses

Women journalists can report acts of online violence to the Federal Investigation Authority (FIA) to pursue cases under the Prevention of Electronic Crimes Act (PECA) 2016. However, this legislation has been widely criticised for being open to abuse by law enforcement agencies and government officials seeking to suppress legitimate content (Khan, 2016; Ashraf and Asif, 2020). On the other hand, interviews for this case study and other research show that women journalists rarely take legal action in response to online violence. Sometimes this is because they lack knowledge of the existing law, but journalists more commonly refer to a lack of trust in the law enforcement system to effectively pursue cases. This is based on previous experiences when complaints were left unanswered, the police's perceived lack of sensitivity to women reporting online violence, and fear that complaints may backlash, especially if the journalists had made any controversial statements about religion or the State (DRF 2017a; Ghani and Khan, 2019). As a result, only 20% of women journalists (22 respondents) participating in the 2019 MMfD survey said that they had reported online violence to a relevant authority, with just a quarter of those saying that the perpetrator had been stopped (Kamran, 2019).

The Protection Against Harassment of Women at the Workplace Act 2010 provides a comprehensive framework for prosecuting workplace-based sexual harassment, and it is possible that women could invoke this law as a protection against "an intimidating, hostile or offensive work environment." However, at the time the research concluded no evidence could be found of any women journalists

having done so. In November 2021, Pakistan's Senate passed the long-awaited Protection of Journalists and Media Professionals Bill. While generally welcomed, journalists' unions and rights organisations, such as the International Federation of Journalists (IFJ) and Reporters Without Borders (RSF), have raised concerns that the law is a new censorship mechanism under the guise of protecting journalists. For example, Section 6 neutralises all protection that the law was intended to provide when it was first announced by the government in May 2021. This section prohibits all journalists and media professionals from spreading "false information" and producing material that "advocates hatred" or constitutes "incitement to discrimination, hostility or violence" – without clearly defining what any of these terms means (Ali, 2020; Ahmad, 2020).

70% of respondents (77 individuals) to the 2019 MMfD survey had reported online violence to social media companies at least once (Kamran, 2019). Those interviewed for this study expressed dissatisfaction with the platforms' regular rejection of their complaints, and with the fact that abusers' accounts were not removed, so they were left to block or ban them themselves. They said that reports of abuse in Urdu, one of Pakistan's two official languages, were not always actioned, and that the platforms were unable to moderate in minority languages and dialects. Similarly, they indicated that specific cultural implications of abusive language were hardly considered, in relation to universal platform community standards, resulting in delays or a lack of action.

Women journalists make up less than 5% of the profession in Pakistan (Nusrat, 2018; White, 2009). In this heavily male-dominated industry, the interviewees indicated that women journalists experiencing online violence usually receive little support from their employers, and that they are often told to tone down their reporting or retreat from the digital space. In turn, this leads to a reluctance to report incidents to employers, in order not to be labelled a "troublemaker" or to lose job opportunities. Very few media houses offer digital safety training to their employees, and even fewer provide counselling to help them cope with negative experiences (Solidarity Center, 2016). Interviewees said that women journalists would turn to friends and colleagues before ultimately contacting civil society groups for psychological support, and that most would do so discreetly because of stigmatisation associated with mental health issues.

The women journalists interviewed placed their hopes in civil society advocacy efforts to demand changes to relevant laws and policing. The main Pakistani NGOs working in this field are the Digital Rights Foundation (DRF); Media Matters for Democracy (MMfD); and the Coalition for Women in Journalism (CFWIJ). Journalists also acknowledged the efforts of Bytes for All and the Pakistan Journalists Safety Fund (PJSF). Overall, interviewees indicated that the major NGOs are having a useful impact in supporting women journalists to make complaints, to seek counselling and peer support, and to develop greater awareness of digital and physical safety and security strategies. As a result of their work in documenting these attacks and bringing them to public attention, these organisations have sparked debate about the need for political action and legal reform, and they have also inspired women to speak out about the issue. However, to date, they have had limited impact on convincing government actors, social media companies and the media industry to step up action against online violence.

A particularly high-profile campaign calling for State action was launched in August 2020, when 172 signatories joined a statement initiated by 20 women journalists with the Twitter hashtag #AttacksWontSilenceUs. The petition called attention to the coordinated online harassment these journalists had been subjected to, and stressed its far-reaching implications for freedom of expression and access to information in the country. The signatories asked the Pakistani government to restrain its members from targeting women journalists, to direct all supporters to desist from launching such attacks, and to take action against those responsible (Sarfraz et al., 2020). The petition initiators subsequently related their stories of abuse to the national parliament's Standing Committee on Human Rights. This was the first time Pakistani women journalists have run a national public campaign demanding government accountability for reducing politically motivated abuse.

Recommendations for action in Pakistan

Pakistan authorities and lawmakers should:

1. Publicly condemn/ act against gendered violence online, for example by establishing a roadmap to combat this problem. This could include an independent inquiry into the source of coordinated intimidation of women journalists, and a public awareness campaign.

2. In consultation with civil society and media industry bodies, call for a review of Section 6 of the Journalists & Media Professionals Bill by an independent commission to address concerns e.g., sub-section 3 says that journalists who fail to comply with these "obligations" will be subject to criminal prosecution.

Pakistan law enforcement and judicial actors should:

3. Identify online abuse and harassment against women (journalists) as a crime, with clear penalties.

4. In consultation with civil society and media industry bodies, political parties could call for a review of Section 6 of the Journalists & Media Professionals Bill by an independent commission to address concerns e.g., sub-section 3 says that journalists who fail to comply with these "obligations" will be subject to criminal prosecution.

5. Provide training for specialist, digitally trained officers in handling complaints of gender-based violence; hire more women officers considering the patriarchal context.

6. The Federal Investigative Agency could release an annual report detailing the progress it has made on progressing complaints about digital abuse against women journalists and pathways to prosecution.

Social media companies providing services in Pakistan should:

7. Establish more culturally sensitive protocols and local language capabilities for handling user reports of hate speech, abuse and harassment and direct internal contacts for media organisations to formally report harassment campaigns against journalists - partly by hiring specific language speakers.

Pakistan media organisations should:

8. Tackle discrimination in employment and workplace practices, hire more women as journalists and ensure they have peer and organisational support in mitigating digital risks, including through digital security and safety training, and policies and procedures for addressing online attacks.

1.7 Poland

Greta Gober

The worst in all this is the systemic consent to this kind of violence. There is a reason why we don't report this abuse. We know nothing will change, it's such a powerless feeling. We can organise workshops and talk about it, but it's still us (women) who have to come out as victims or have to prepare to potentially be victims.

Alicja Cembrowska, Editor-in-Chief

This summary draws on a detailed case study about online violence against women journalists in Poland, involving extensive desk research and in-depth interviews with 10 journalists and experts.

Challenges faced by women journalists

Existing research on online violence against women journalists in Poland is limited but interviewees shared the impression that it was a growing problem in the country. Polish women journalists are attacked online in different forms, including through insulting and hateful comments, attempts to undermine their work and credibility with audiences, threats of violence and even death threats. Often, both abusive commentary and threats of harm are sent by fake accounts or under pseudonyms. But there are also cases of public figures such as politicians, priests, lawyers, celebrities, or other reporters openly attacking women journalists, and sometimes they have engaged in dog-whistling (i.e., asking their followers to engage in abuse/discreditation of the journalist). Interviewees also mentioned sources or personal contacts being behind some of the attacks.

Online violence against women journalists threatens freedom of expression in Poland: multiple interviewees reported having engaged in self-censorship to avoid being targeted online. Acts of self-censorship ranged from shying away from certain stories, avoiding a certain beat or withdrawing to less publicly visible positions within journalism. Additionally, online attacks take a toll on the mental and physical health of journalists and have led to some of those targeted withdrawing from online activities.

Institutional and legal response mechanisms

While there is no single piece of legislation covering online violence in Poland, the Polish Criminal Code penalises online offences such as threats of physical

or sexual violence (Article 190), stalking or persistent harassment (Article 190a), 'revenge porn' (Article 191a) or insults (Article 216). 'Praising' (expressing a positive assessment of a crime) or incitement to commit a crime online (Article 255) are also prosecutable under the Criminal Code. Prosecuting some manifestations of online violence under civil law is possible under Art. 107 of the Code of Petty Offenses. More broadly, Poland also has an obligation to respond to gender-based online violence as a party to the Council of Europe's Convention on Preventing and Combating Violence against Women and Domestic Violence (also known as the Istanbul Convention).

However, as is common elsewhere, the anonymity of perpetrators limits enforcing civil protections, as challenging online attackers in civil proceedings is only possible when their personal details are known. Another barrier to seeking legal support is linked to the fact that sexist or misogynist speech is not covered under existing hate speech laws. Extending these laws to cover gender and sexual orientation was raised as a point of possible improvement in this study's interview with lawyer Zuzanna Warso, co-author of the Helsinki Foundation for Human Rights' study *Cyberviolence Against Women* (2017).

Law enforcement agencies are perceived to downplay the severity of the issue and often do not take the necessary steps to apply existing legal instruments to online violence (Smetak and Warso, 2017). Additionally, in what has been described by observers as an increasingly restrictive environment for women's rights in Poland, the government has attempted to revoke the Istanbul Convention and has labelled many of its provisions as 'gender ideology' (Ciobanu, 2020a). In an interview, Warso highlighted a context of eroding civil society space and a lack of institutional support in the current Polish political landscape which she viewed as necessary to take into account when considering measures against online violence. Recent cases of violence by members of the police against journalists and other repressive measures also complicate possibilities for taking action (Wiseman, 2021). Additionally, some existing legal provisions also risk being abused for restricting freedom of expression at large. Poland's Commissioner for Human Rights and the Helsinki Foundation for Human Rights have also called for the removal or significant adaptation of Article 212 of the Penal Code which deals with defamation, due to its propensity to be used to curb press freedom (Helsinki Foundation for Human Rights, 2011).

Tech platforms are liable to react to and/or remove illegal content under the Act on Providing Services by Electronic Means. However, in practice this law is rarely enforced. Removal of allegedly harmful content on Facebook, YouTube and Twitter has been shown to be less frequent in Poland than in many other countries (Nigdy Więcej, 2020a, 2020b). In 2018, Poland became the first country to offer Facebook users an external appeal mechanism for blocked and removed content through a service launched by the Ministry of Digitisation (Government of Poland, 2018). However, civil society organisation the Panopticon Foundation argues that this does not address systemic problems of lack of transparency and arbitrariness in moderation on the platforms (Obem and Głowacka, 2019b).

In this context, initiatives to combat online violence have not been numerous in Poland. Few newsrooms have developed protocols for dealing with gender-based online violence, with any existing measures largely limited to comment

moderation or ad-hoc support to targeted individuals. However, civil society organisations, most notably the International Press Institute, have conducted engaging awareness raising campaigns about online violence and countermeasures (Trionfi and Luque, 2019), while other NGOs have focused on the intersecting issue of disinformation. Lawyer Dorota Głowacka suggested that the rise of an anti-feminist movement presented an opportunity to mobilise against gender-based violence and discrimination online and offline:

We are in the midst of a big wave of demonstrations driven by women and we are only beginning to acknowledge how women's voices have been silenced...and perhaps this is the time when we can start changing this.

Recommendations for action in Poland

The Polish authorities and lawmakers should:

1. Support a debate about extending the application of Article 256 and 257 on hate speech to include sex, gender and sexual orientation as per the UN Strategy and Plan of Action on Hate Speech.

2. Consider amending or deleting Article 212 from the criminal code (pertaining to slander) to ensure it is not used to curb press freedom.

3. Ensure legislation regulating social media companies holds them accountable for swift removal of illegal content and adequately addresses their role in spreading and amplifying online violence against journalists

4. Dedicate resources to aid civil society actors in developing training, resources and policies targeting law enforcement, judicial actors, tech companies, legislators and policymakers, to support the development of monitoring mechanisms (independent from the State) that record violations against journalists disaggregated by gender, location (online, offline), nature, etc.

Law enforcement and judicial actors should:

5. Build knowledge in an effort to develop the understanding and skills required to deal effectively with cases of gender-based online violence against journalists.

6. Work actively to prevent premature rejection or dropping of harassment cases reported by journalists and media houses, and develop routines to thoroughly document both dropped and prosecuted cases for research and knowledge-building purposes.

Polish news organisations should:

7. Implement preventative strategies in the form of detailed policies, guidelines and response systems for dealing with online violence against journalists that include assessment of individual and gender risk factors and psycho-social support.

8. Invest in developing newsroom cultures that acknowledge, and are responsive to, online attacks.

9. Collaborate with a variety of media houses across the Polish political spectrum to develop a culture of solidarity within an extremely polarised society.

10. Engage with civil society actors on educational campaigns that sensitise the public to the negative effects of online violence and to facilitate systematic gender-disaggregated monitoring of cases.

11. Employ or establish collaborations with experts to assist newsrooms in dealing with the intersecting problems of disinformation, online violence and the line between 'free speech' and 'hate speech', including in efforts to develop successful communication avenues with social media platforms and improve internal incident reporting protocols.

Civil society should:

12. Develop training, resources and policies to assist news organisations, law enforcement, judicial actors, social media companies, legislators and policymakers to ensure systematic knowledge is built across the country on gender-sensitive and effective responses to online violence targeting journalists in the specific Polish context.

13. Engage in research to build evidence-based knowledge and develop a nuanced understanding of gendered online violence against journalists in Poland.

Social media companies providing services in Poland should:

14. Recognise that they have a responsibility to consider how gender-based online violence is facilitated and enabled by their own infrastructure and business models, keeping in mind that Poland's Act on Providing Services by Electronic Means obliges them to react prima facie to illegal content.

15. Work to establish swift and effective communication channels with media houses across the country to ensure press freedom-sensitive protocols for addressing content removal requests, and work more quickly and effectively to remove threatening, hateful and harassing content targeted at women journalists.

16. Develop capabilities for user incident reporting in Polish and develop capability to monitor and respond to requests in Polish, with appropriate cultural context.

1.8 Serbia

Greta Gober, Bojana Kostić, Nabeelah Shabbir, and Jennifer Adams

(I)f you are a victim of a smear campaign there are certain places you cannot go anymore because you won't be safe offline...which makes your job impossible.

Natalia Żaba, freelance journalist based in Serbia for a decade

This summary draws on a detailed case study about online violence against women journalists in Serbia, involving extensive desk research and in-depth interviews with 12 journalists and experts.

Challenges faced by women journalists

In Serbia, the context of EU accession negotiations creating a positive impetus for freedom of expression on the one hand, and deteriorating media freedom conditions on the other, serves as the backdrop for the online violence experienced by women journalists in the country. A number of media outlets aligned with the government are accused of frequently initiating disinformation campaigns and judicial harassment against opponents, with women journalists as common targets of misogynistic smear campaigns (Witkowski, 2016; N1 Belgrade, 2020; Djurić, 2019). One example from early 2021 involves a woman journalist from independent broadcaster N1, Žaklina Tatalović, who was targeted with sexually explicit abuse online. On one occasion, the Editor-in-Chief of a tabloid newspaper took a picture from the journalist's social media account without her consent and published it alongside sexist commentary (Commissioner for Protection of Equality, 2020a; Danas, 2020). In turn, this triggered an avalanche of hateful vitriol and bigotry online.

At the same time, regulatory authorities tasked with the role of promoting media freedom and the safety of journalists are called upon by some to be more effective and more independent (Vukasović, 2018). Combined with the fact that gender-based violence as a whole is on the rise (OSCE, 2019), this presents a complex context for addressing online violence against women journalists in Serbia.

A Serbian study from 2020 found that over a third of 80 women journalist survey respondents said they received threats through online channels or experienced safety risks in the previous five years (with an increase in 2020) and that just under one-fifth reported online threats to the police and prosecution services (Djurić & Jović, 2020). A 2021 study found that officials are among the biggest

instigators of pile-ons against journalists, and that government representatives often fail to respond to attacks against women journalists, particularly in smaller communities (NUNS, 2021). Journalists who cover beats such as corruption and organised crime may be targeted in general, but women journalists are more likely to face gendered and sexualised harassment and abuse because of their work (Djurić & Jović, 2020).

Online threats against women journalists in Serbia can turn into offline threats, including stalking, property damage, and assault according to interviewees participating in this study (Djurić & Jović, 2020). However, journalists interviewed perceived that they were not being taken seriously when trying to raise awareness and take action against online violence.

Institutional and legal response mechanisms

Serbian journalists interviewed in this study highlighted the vital role of international NGOs and intergovernmental organisations in monitoring and reporting cases of harassment and intimidation against women journalists in the country. Several local and regional organisations, including the Western Balkans Regional Platform, the Independent Journalists' Association of Serbia, and the South East Europe Media Organisation monitor threats against journalists and offer direct support to those in need. However, research for this case study established that the need to address the specifics of gendered online violence is still largely in the phase of awareness- and knowledge-building. Ad-hoc, often survivor-driven groups try to fill the gaps in support provision, but they face mounting challenges in the face of opaque social media companies, a complex political landscape, and legal systems fraught with barriers.

In the framework of an Action Plan designed for the purpose of EU accession negotiations, an Agreement on Cooperation and Measures for Increasing the Level of Safety for Journalists was signed by the Prosecutor's office, the Ministry of Interior and journalists and media associations in 2016. A Permanent Working Group for the Safety of Journalists was also established with contact points serving as interlocutors between journalists, state authorities, and the judiciary. While these contact points have been considered an important transitional step for building permanent procedures within the state system, awareness of them is still low among journalists (Djurić, 2019).

In 2020, the government adopted a new Media Strategy, acknowledging online and offline gender-based threats and recommending measures such as gender-disaggregated data collection, gender-specific training for the judiciary, law enforcement, media lawyers and journalists, and possible amendments of the Criminal Code (Republic of Serbia, 2020). It was followed up by an Action Plan and a Working Group addressing implementation of these measures. Many of this study's interviewees stressed that robust political support and financial commitment are necessary for these measures to be effective.

Legal experts interviewed for this case study agreed that significant institutions and mechanisms regulating the issues of journalists' safety and violence against women in Serbia do exist, but assessed that the implementation of legal provisions

is often inadequate or ineffective. The Criminal Code prohibits technology-facilitated offenses, and a comprehensive legal framework for addressing violence against women is in place, but neither of these addresses online gender-based violence. A special government department for combating cybercrime has been established under the Law on Organisation and Competences of Government Authorities in Fight Against High-Tech Crime Act (Djurić & Jović, 2020), and online attacks are now also investigated through the Special Prosecution for High-Tech Crime department in the Higher Public Prosecutor's Office in Belgrade. However, there is a perception that these bodies should do more to appropriately take gender-based crimes into consideration. Several studies have indicated that the lack of consistent implementation of the existing laws and provisions is the biggest missing puzzle piece for addressing online violence against women journalists in Serbia (Djurić & Jović; 2020, Slavko Ćuruvija Foundation, 2021b)

Article 138, para.3 of the Criminal Code (i.e., *the criminal act of endangerment of safety against a person who performs work in the public interest in connection with the performance of their duties*), is the most relevant article for criminalising online violence against journalists in Serbia and is currently the legal instrument most often applied to such cases (Djurić, 2019). However, Rade Djurić, author of the Slavko Ćuruvija Foundation report *Critical Points in the System of Safety of Journalists* (Djurić, 2019) assesses that complicated legal requirements in Article 138 and beyond pose a great challenge to journalists' safety. One example is the differentiation between acts defined as 'pressure' (which are neither criminal nor misdemeanour) and criminal acts. The wide range of behaviours that fall under the 'pressure' umbrella rather than the criminal code, including smear campaigns and other forms of online and offline violence that often affect women journalists, can contribute to impunity for attackers.

Women journalists rarely report online violence incidents to the police, as they often do not perceive these as criminal offences, or if they do, they are not always made aware of additional legal requirements for triggering the criminal law procedure (notably that the prosecution for sexual harassment and stalking offences requires that the injured party initiates the action). Many other such legal and social obstacles stand in the way for prosecuting other criminal offences covering online violence against women journalists, including a lack of understanding, lack of digital investigative capability, and a risk for experiencing a re-victimisation in law enforcement processes.

Recommendations for action in Serbia

The Serbian authorities and lawmakers should:

1. Follow more closely the Council of Europe standard-setting documents on the protection of journalists.

2. Provide more funding to, and ensure efficiency of, existing press freedom institutions, including the Standing Working Group for the Safety of Journalists.

3. Consider reviewing the criminal justice system (substantive and procedural elements) to ensure access to efficient redress mechanisms for women journalists facing attacks online.

4. Penalise politicians and government officials who orchestrate or instigate online smear campaigns against women journalists, or abuse, threaten or attack them directly.

5. Support cross-sectoral cooperation and inclusion of non-State actors in multi-stakeholder processes designed to cooperatively create preventive and responsive instruments.

Law enforcement agencies and the judiciary should:

6. Ensure the prosecutors and agencies tasked with implementing existing laws have the required digital skills and access to the appropriate technology.

7. Foster the creation of a gender-sensitive environment, through training and other capacity-building measures, in which both police and prosecution agencies recognise the gravity of harassment and offer agile and proactive protection to women journalists that report threats.

8. Collect disaggregated occupational data in the Republic's Bureau of Statistics to better track how many women journalists are affected in criminal proceedings.

News organisations should:

9. Address misogynistic newsroom cultures that deepen the impacts of gendered online violence against women journalists.

10. Foster and build a holistic and gender-sensitive safety culture within news organisations (e.g., safety protocols, guidelines, and access to support services) to tackle hybrid threats.

11. Work collaboratively to counter smear campaigns against women journalists through investigative journalism and solidarity initiatives.

12. Provide specific resources and mechanisms for inclusion of women journalists in vulnerable positions: freelance, investigative journalists, local reporters, covering sensitive topics, etc.

Social media platforms providing services in Serbia should:

13. Ensure that monitoring and reporting mechanisms are capable of operating in the Serbian language and ensure that appropriate in-country human responders with press freedom knowledge and gender-awareness are available to escalate online violence cases.

14. Monitor and openly report on online attacks (gender-disaggregated) and maintain open communication and cooperation with civil society and other relevant actors.

15. Develop culturally and linguistically intelligent content moderation expertise capability and train algorithms to respond to the particular issues in the Serbian context.

Civil society organisations should:

16. Develop and implement a set of tools and mechanisms to ensure immediate and efficient protection of women journalists against online violence, including a "one-stop-shop" for digital safety or a journalist solidarity fund in cooperation with media associations.

17. Continue to foster gender-sensitive expertise within media associations, lawyer networks, and other actors to create structured and evidence-based mechanisms to respond to hybrid safety threats facing Serbian women journalists.

1.9 South Africa

Glenda Daniels and Julie Posetti

I was the target of press statements from political parties, and rape threats; these were gendered misogynistic and violent. You won't find similar harassment of males in terms of rape and body shaming.

Qaanitah Hunter, Editor

This summary draws on a detailed case study about online violence against women journalists in South Africa involving extensive desk research and in-depth interviews with six journalists and experts.

Challenges faced by women journalists

Online threats to the safety of women journalists in South Africa have increased over the past five years, specifically targeting women journalists engaged in investigative and political journalism, with Twitter being the most heavily implicated social media space. Many instances of online violence in South Africa are linked to political issues and attempts to discredit women journalists' work in the sphere of political discourse.

Interviewees specifically described attacks following reporting on corruption and local and national politics. Former investigative political journalist and now editor Qaanitah Hunter was targeted online by political actors due to her reporting, including through death threats. She stated that social media platforms' business models "incentivise and reward trolling".

Journalists also reported difficulty getting social media companies to remove content and described apparent duplicity by the platforms, with one interviewee contrasting their "posturing" over free speech while failing to foreground other human rights, such as the right to be protected from gender- and race-based discrimination and hate speech, and being primarily focused on growing market share.

Doxxing and threats of sexual violence; the posting of sexualised and manipulated images and memes; death threats (e.g., being sent a picture of a gun), as well as threats of violence, and abusive and hurtful commentary are prominent types of online violence inflicted on South African women journalists. There are also signs of orchestrated disinformation campaigns and 'networked gaslighting',[240]

240 Coordinated or interlinked attacks that undermine the confidence and credibility of the abuse target by twisting perceptions of reality and sometimes falsely attributing the same abusive behaviour to the subject.

and instances of online violence spilling offline - including at political party press conferences in the case of Daily Maverick journalist Pauli Van Wyk.

Institutional and legal response mechanisms

The South African constitution protects freedom of expression, with limitations of this right in cases of incitement of violence and advocacy of identity-based hatred. Gender equality is also enshrined in the constitution, and discrimination on the basis of sex and gender is unlawful, however there is no specific legal protection for women journalists.

Some of the interviewees lamented what they see as a lack of capacity and understanding among the South African police, leading to insufficient action in tackling the problem. Editor and journalist Ferial Haffajee said that the courts were also lagging behind in comprehending gendered online violence as a threat to press freedom.

Further, interviewees reported unsuccessful attempts to hold political actors legally accountable for creating an environment that enables harassment of women journalists and called for legal reform in this area. There have been limited legal successes for women journalists seeking justice in online violence cases, however. For example, the late journalist Karima Brown successfully sued a populist political party leader in 2019 after he doxxed her on Twitter.

South African civil society has played an active role in advocating for effective policy responses to gender-based online violence while also recognising the impacts on freedom of expression. Organisations such as Media Monitoring Africa (MMA) and Gender Links argue for preventative action by social media platforms to act against gendered online violence, and remove the onus on individuals to flag offensive content, while the South African National Editors Forum interviewees also positively emphasised global efforts to study, discuss and combat the issue of online violence. Ferial Haffajee said that through attending international events she realised that others were experiencing the same types of attacks, and being heard was empowering and evoked a feeling of solidarity.

Recommendations for action in South Africa

The South African authorities and lawmakers should:

1. Recognise the role of a number of political actors and parties as perpetrators of significant targeted online violence campaigns against South African women journalists; note the chilling effects that these attacks have on freedom of expression rights protected under the South African constitution, and the flow-on effects for democracy; and call out such acts.

2. Consider proposing parliamentary conduct standards that preclude acts of gendered online violence by elected representatives against journalists and attach penalties.

3. Introduce legislative and regulatory responses to ensure transparency and accountability from social media platforms, in reference to gender-based

online violence, e.g., mandating incident monitoring and periodic reporting on incidence and responses to violations; requiring the companies to employ local staff with an emphasis on gender-aware, media-freedom capable, and multilingual moderation interventions (in South Africa's 11 national languages).

South African law enforcement and judicial actors should:

4. Train law enforcement officers to help them recognise and respond to online violence as a potentially criminal act, with serious offline impacts.

South African news organisations should:

5. Appoint dedicated point people within newsrooms, who have undergone gender-awareness training, to respond to online attacks on women journalists in real time, noting the potential for severe offline impacts in the South African context which includes an alarmingly high rate of sexual assault;

6. Provide access to trauma counselling to help the women journalists targeted deal with the impacts of gender-based online violence, regardless of financial constraints, recognising that it is a workplace safety issue, thus addressing the widespread absence of such programmes in South African newsrooms;

7. Support women journalists who wish to raise their voices, investigate and write about their experiences

South African civil society organisations should:

8. Consider strategic litigation and the pursuit of legal remedy as potentially effective responses to online violence against women journalists in the South African context.

9. Help create opportunities for South African women journalists to participate in international conversations and training to aid their empowerment.

1.10 Sri Lanka

Fiona Martin, Nirasha Piyawadani, and Jenna Price

It went viral in social media saying that I have been raped and killed. I found myself staring at some photoshopped pictures of my...murdered body, accompanied by news reports announcing my death by a fundamentalist group...it really disturbed me a lot because many of my friends and even my family, everyone thought it happened to me. So many people walked to my home for my funeral because people trusted and believed that news that much.

Sharmila Seyyid, senior journalist and human rights activist

This summary draws on a detailed case study about online violence against women journalists in Sri Lanka, involving extensive desk research and in-depth interviews with 10 journalists and experts.

Challenges faced by women journalists

Original research within the framework of this study and NGO reports suggest that online violence against women journalists is on the rise in Sri Lanka (Perera, 2020; Amnesty International, 2020f). Experts interviewed consider that online attacks against women journalists in the country are largely rooted in gender-based discrimination and misogyny. Sri Lanka's gender equality has been said to have "declined precipitously" in the past decade, while gender-based violence against women and girls "has increased in visibility in the public sphere and in the media" (International Finance Corporation, 2019). A 2015 report by the International Federation of Journalists (IFJ) found that, although slightly more women had been working in the media in the previous five years, they still struggled in unsafe environments, with poor working conditions and a lack of upward mobility, and with a strong male bias in newsroom employment (IFJ, 2015). According to several interviewees, there is frequent "victim-blaming" in the workplace, and in society more broadly, hindering acceptance of online violence as a crime.

Several journalists and experts interviewed suggested that some attacks were organised by political groups, and that a number of local politicians were paying people to harass journalists. Some argued that online harassment was generally politically motivated, but tended to be racial or religious in nature, as these factors were guaranteed to trigger greater moral offense and amplify the backlash against the target. Others indicated that some of the perpetrators were ordinary people,

caught up in a general backlash against women who advocate for social change. Most interviewees, however, agreed that the women most attacked are those engaged in contested public debates and working on social causes. Those from Tamil, Muslim and other minority backgrounds, as well as those who cover politics, sensitive national issues, human rights, or religion, are particularly vulnerable, along with women's rights advocates and those who write from a feminist perspective. Journalist and activist Sharmila Seyyid, who meets many of these criteria, described an extreme online hate campaign that moved offline, after which she gave up her journalism career and fled the country for three years with her infant son (Seyyid interview, 2020). Two interviewees said that they avoided writing about controversial topics after being harassed. Other interviewees indicated that they were warned by friends, family or work colleagues not to speak about the violence, or to tone down their comments online.

A challenge to effective content moderation on the social media platforms is that perpetrators often use images to avoid text-based detection or supposed humour to avoid triggers for removal. Women journalists are also harassed via private channels that are harder to police, such as Google Chat, Facebook Messenger, or direct messages on Twitter. Some are subject to 'cheap fakes', synthetic media attacks involving their headshots being digitally mapped onto pornographic body images, which are shared online with comments suggesting they are sexually promiscuous. Spoofing (i.e. impersonation) is also a growing concern. One journalist reported that a perpetrator wrote a controversial political article under her name and sent it to various publications, using her personal details.

Some journalists have withdrawn altogether from the main platforms because of the extensive harassment they have experienced there. But as with other countries in the region, Facebook's Free Basics or internet.org project means that to many Sri Lankans, Facebook is the internet, making it difficult to avoid.

Institutional and legal responses

The interviewees for this study shared the view that women journalists experience great difficulty reporting online violence to the authorities. Further, they noted that gender-based hate is not specifically addressed in the Penal Code (Sri Lanka Ministry of Justice, 2016a). Since January 2020, the Sri Lankan Defence Ministry has been finalising the draft of a Cyber Security Act to tackle social media defamation as well as ethnically and religiously sensitive posts inciting hatred and posing a threat to national security (Ministry of Defence, 2020). Digital attacks may, partly or fully, fall under several existing laws (Democracy Reporting International, 2021), but journalists' difficulty in prosecuting perpetrators is compounded by "lack of clarity as to how and where to report incidents of technology-related violence and institutional coherence on which state agency would follow each case" (Perera and Wickrematunge, 2019). As a result, several journalists indicated that they experienced little support from national agencies tasked with investigating such crimes.

The interviewees also stated that women journalists do not feel local police have the digital literacy to understand the crimes being reported, and they argued that officers do not handle the reports in a competent, timely or gender-sensitive manner. They described feeling uncomfortable talking to male police officers

about image-based abuse and salacious accusations, and noted there were few, if any, female police officers trained to understand the scope of online violence. For journalists from minority backgrounds, language barriers may further inhibit reporting, as all complaints must be handed in Sinhala.

The Sri Lankan government's key online violence countermeasure has been the short-term blocking of major social media platforms in 2018 after anti-Muslim riots, and again three times in 2019 after the Easter Sunday bombings.[241] The aim was to stop the circulation of hate speech and misinformation. Some observers say that these bans did not solve those problems (Wijeratne, 2018; Samarajiva, 2018), while others say they slowed the flow of legitimate news, disrupted family and social relations, and hurt businesses (Gross, 2018). Globally, concerns were raised about freedom of expression implications.

Following the 2018 social media ban, Facebook promised to work with the government to mitigate hate speech and platform misuse, improve automated moderation in Sinhala, and increase stakeholder engagement. It also commissioned an independent investigation into its role in the 2018 riots, which noted, amongst other issues, that the platform had been used against women for circulation of gender-based hate and image-based abuse (Article One, 2018).

The interviewees emphasised the need for social advocacy and awareness-raising considering Sri Lanka's gender, class and language divides. Several stressed that the gender inequity in the parliament may pose a challenge to the systemic change required to fight online violence against women journalists. In January 2021, only five percent of legislators were women, with Sri Lanka ranked 182nd out of 193 countries in female parliamentary representation (IPU Parline, 2021).

A number of NGOs have been working on projects to tackle gender-based violence and women's political participation including Women in Need, National Forum against Gender-Based Violence, Women's Development Centre, and Hashtag Generation. In March 2021, the latter launched a Media Gender Charter to guide media companies in taking steps to increase gender equity and inclusivity (Hashtag Generation, 2021). This is the latest in a series of moves over the past 15 years to increase female representation in Sri Lanka's media, to improve gender equity in pay and conditions, to eliminate sexual harassment and to embed gender sensitive reporting norms (Lanka Business Online, 2021; Media Reform, 2006). The country has also been implementing gender equity initiatives that may eventually affect gender relations in the workplace, such as the International Finance Corporation's (IFC) corporate project Governance for Women in Sri Lanka.

Recommendations for action in Sri Lanka

The Sri Lankan authorities and lawmakers should:

1. Reform the current laws to better enable the prosecution of digital attacks on women journalists while protecting the legitimate exercise of freedom of expression.

---

241 The "Easter Sunday bombings" describe a series of attacks by Islamic State militants on 21 April 2019 which killed nearly 300 persons and injured at least 500.

2. Set up a designated criminal policing unit to deal with online violence, supported by a public hotline, which offers legal and counselling advice, and a database for documenting cases.

3. Work with civil society groups to ensure greater female participation in journalism and broader awareness of how best to combat sexual and gender-based violence.

The Sri Lankan police force should:

4. Provide training for officers in receiving complaints of gendered violence and awareness of legal and social support avenues for women pursuing cases.

5. Develop capability for taking complaints involving minority language abuse.

Sri Lankan media companies should:

6. Provide all journalists with integrated (online-offline), gender-responsive digital security and safety training.

7. Develop policies and procedures for dealing with cases of online violence against women employees, appoint someone to collate data about attacks, and offer psychological support for the victims of abuse, along with time away from work when necessary.

8. Develop and implement gender equality programmes, with training provided to all staff.

Social media companies providing services in Sri Lanka should:

9. Introduce more effective content moderation and algorithmic pre-moderation capability in Sinhalese, Tamil and Malay languages.

10. Undertake outreach with women's groups to understand the issues facing women journalists and other women in public life.

1.11 Sweden

Sara Torsner, Greta Gober, and Julie Posetti

It is the threats from violent supporters of right-wing extremists that have made it necessary for me to have security windows and an alarm button next to my bed.

Åsa Linderborg, Aftonbladet journalist

This summary draws on a detailed case study about online violence against women journalists in Sweden involving extensive desk research and in-depth interviews with 10 journalists and experts.

Challenges faced by women journalists

In 2019, the Swedish Internet Foundation found that one out of seven internet users surveyed reported being exposed to 'net hate' (as online abuse is often called in Sweden) or to someone posting negatively about them online (Swedish Internet Foundation, 2019).[242] Online violence in Sweden should also be understood in the context of a high degree of digitisation in the country, with 95 percent of the Swedish population using the internet on a daily basis (Swedish Internet Foundation, 2019). Threats against journalists in Sweden have intensified in recent years (Stjärne, 2020). Hanna Stjärne, CEO of Swedish Public Television (SVT), reported that SVT deals with around 35 security incidents per day, including both digital and physical forms of attack (Stjärne 2020; Dagens Nyheter 2020, 2019). Several recent studies have confirmed that women journalists are particularly exposed to online violence and are targeted by gendered forms of harassment and abuse (Eriksson Almgren et al., 2017; Fernquist et al., 2020).

Additionally, and in line with findings from other countries, women journalists who report on certain politically charged topics, and those who experience multiple forms of marginalisation are particularly targeted online. These intersectional dynamics should be understood especially in the context of sustained far-right online threats in Sweden. Racist, xenophobic, anti-Semitic and nationalist groups are well-organised in the Swedish online sphere (Kaati et al., 2017), and frequently target journalists who are perceived to represent opposite political views (Mokhtari et al., 2017). Reporters Without Borders (RSF) Sweden reports that the rise of political populism and white supremacy movements has resulted in a growing online presence and surge of alternative media with a clear far-right political agenda (RSF, n.d.).

In this context, attacks against journalists with migrant backgrounds, journalists who report on migration issues, or journalists who carry out investigations into right-wing, populist and xenophobic political movements are prevalent (Löfgren

---

242 This is an annual demographically representative survey relying on a sample of 2,706 interviewees.

Nilsson, 2017). Journalists who report on gender issues or feminism are also routinely targeted (Tahir, 2021). These attacks are often carried out via public social media channels and websites that seek to discredit, 'name and shame', and instigate dogpiling on particular journalists. Online groups with names such as the 'Critique Authority' [Kritikverket], 'Audit Unit' [Tillsynsenheten], or 'Auditing Sweden' [Granskning Sverige] combine disinformation tactics with coordinated abuse and harassment campaigns to silence journalists on their 'red lists' (Ståhle, 2017; Mussa, 2019; Skarrie, 2020).

Institutional and legal responses

Workplace protections, which are relatively extensive in Sweden, can be invoked in the case of journalists experiencing online violence in the course of their work, but do not always cover all who are affected by the problem, particularly freelancers (Bohlin, 2017; Ministry of Employment, 2015). Research conducted for this study established that news outlets appear to prioritise the physical safety of their employees above their psycho-social wellbeing or digital security (at least in terms of resource allocations and interventions) and that there is a lack of gender-sensitive responses to online violence.

While positive examples of security experts getting involved early on in the process of responding to online violence (for example, by monitoring articles that may trigger attacks) were identified, experts and journalists also expressed that there is room for improvement. Major media houses tend to focus their security efforts on large and acute situations rather than the daily deluge of comments and private messages delivered to journalists. Smaller media houses also often lack the resources to offer adequate support to journalists dealing with online violence. However, acknowledging that news organisations can only do so much to address the crisis, Sveriges Radio CEO Cilla Benkö has called on social media companies to step up their efforts and develop more effective responses to online violence against Swedish journalists (Benkö, 2021).

Swedish Minister for Energy and Digital Development Anders Ygeman said that the EU's current framework was not sufficiently effective in holding the platforms accountable or removing unlawful content and wanted social media companies to be legally held partly liable for what is published on their platforms (Ministry of Infrastructure, 2020; Svärd, 2020). The media group Schibsted, which owns some of the largest media houses in Sweden, wanted social media companies made more clearly accountable under the law by combining EU-wide rules and national legislation; they issued recommendations on transparency, introducing national supervisory bodies, and more (Schibsted, 2019).

Online and offline, criminal activity is largely regulated by the same legal framework in Sweden. Legal provisions specifically designed to apply to the internet are the exception in Sweden - only a few such examples exist (Bladini, 2017; Polisen, n.d.; Posetti, 2017). According to Schultz (2020) it is important to point out that Swedish criminal law is applicable to the internet and that the Swedish legislation is not more lenient towards hateful statements made online than statements made in the physical world. Existing legal provisions with a particular bearing on online harassment, abuse and violence against women journalists include unlawful threat, defamation and insult, molestation and hate crimes.

The Act on Responsibility for Electronic Bulletin Boards (BBS-Act, 1998) was applied to a case of online hate speech in 2019 (Cyber Hate Crime Monitor, 2020a). The 'Enhanced penal protection of personal integrity bill', first adopted in 2017, came into force in 2019 and aims to address existing weaknesses in the law and to signal that these types of crimes, including online, should be taken more seriously (Ministry of Justice, 2017a; Svensson, 2016). With this legislative revision, "unlawful violation of privacy" was introduced as a new crime classification, making it unlawful to spread sensitive images and information. Importantly, the revisions also remove the requirement that online harassment should occur repeatedly over time to be punishable – even once is enough. Under this legislation, social media platforms and other providers of "interactive services" are obliged to remove unlawful threats and violations of privacy (Ministry of Justice, 2017b). In late 2021, a government investigation into legislative reforms around the need for strengthened criminal protection against crimes targeting someone exercising their freedom of expression as part of their professional journalistic activity was ongoing (Ministry of Justice, 2020).

Previous research has suggested that women facing online violence still lack adequate legal protection and support in Sweden (Rosenberg, 2017). The fact that gender is not listed as a motive in national hate crime legislation may be a contributing reason to this gap (Bladini 2017; Rosenberg, 2017). Another reason is grounded in the fact that criminal law is not sufficiently developed to understand the wider problem of the cumulative effects of longterm low range online harassment, making it difficult to apply to cases of online violence against women journalists (Bladini, 2021).

Legal experts interviewed for this study added that the lack of effective implementation of legal provisions already in place essentially leads to a state of impunity for perpetrators. This extends to issues within law enforcement. In a shadow report on the Swedish Police Authority's work with democracy and hate crime online, the Cyber Hate Crime Monitor identified a number of systemic deficiencies, including an uneven distribution of resources, knowledge, and skills among the police force when it comes to dealing with online harassment of journalists (Cyber Hate Crime Monitor, 2019). The process of reporting instances of online crime is also fraught with obstacles (Östensson, 2020). In 2015, the Swedish National Council for Crime Prevention showed that over 96% of online threats and harassment against private individuals reported to the police did not result in prosecution (Andersson et al., 2015).

In 2018, the Swedish Government adopted the National Action Plan Defending Free Speech – Measures to Protect Journalists, Elected Representatives and Artists from Exposure to Threats and Hatred. Grounded in a view that online violence and other threats against public figures constitute a threat to democracy, this Action Plan aims to safeguard actors who participate in democratic discourse and are thereby exposed to violence, online and offline. The Action Plan also specifically acknowledges and proposes measures to address gendered online violence against women journalists (Ministry of Culture, 2018). According to the Swedish Union of Journalists, the Action Plan has been an important tool in drawing attention to the issue of threats against journalists, but also needs to be accompanied by additional measures, notably involving law enforcement, to be truly effective (Carlén, 2020).

Recommendations for action in Sweden

The Swedish authorities and lawmakers should:

1. Ensure that the consequences and costs associated with committing crimes against women journalists are sufficiently significant to serve as a deterrent and investigate how formal legal recognition of crimes against journalists as harmful to democracy could be achieved.

2. Support a debate about the inclusion of sex- and gender-based hate, including misogyny, under hate crime legislation which currently covers race, ethnicity, religion and sexual orientation, and as per the UN Strategy and Plan of Action on Hate Speech.

3. Introduce systematic monitoring and documentation of online violence against women journalists in Sweden for the purpose of building a knowledge base that enables a more comprehensive understanding of the problem.

4. Legislate and regulate to ensure social media companies are held more accountable for protecting Swedish women journalists experiencing violence on their apps and platforms.

5. Commission a review of the legal chain for the purpose of identifying where interventions are needed to combat targeted online violence more effectively against women journalists through law enforcement, and conduct systematic follow-up on the effectiveness of the work undertaken by the Police Authority in the area of online crime.

6. Work to ensure that the 2018-2022 National Action Plan continues into the future, and feeds into an implementation strategy with systematic evidence-based follow-up, to point to areas where additional measures are needed to address online violence against women journalists through a gender-responsive approach.

Law enforcement and judicial actors should:

7. Strengthen coordination and redress regional gaps to ensure national consistency in responding to gendered online violence against journalists.

8. Collaborate more closely with affected journalists, relevant media houses and civil society expert organisations for the purpose of securing evidence gathering and act more proactively and systematically to ensure as much as possible that cases of online violence against Swedish women journalists reported to the police are not dropped, and that necessary evidence for prosecution is gathered.

Swedish media organisations should:

9. Work more methodically and rapidly to incorporate gender-sensitive organisational responses into the routine and often daily exposure of women journalists to online violence.

10. Take responsibility to escalate complaints to law enforcement and/or platforms to ensure that the onus for acting is not placed on the affected journalists.

11. Work together with industry bodies and civil society to share knowledge and develop collaborative strategies to fight back against online violence against women journalists - for example through collaborative investigation.

12. Collaboratively investigate coordinated online attacks against women journalists and the networks behind them.

Social media companies providing services in Sweden should:

13. Act accountably, responsively and transparently in reaction to online violence against women journalists – notably by recording data on incidents of online violence against women journalists, sharing it in privacy preserving ways with researchers, and publishing the results for the benefit of policy development and civil society responders.

14. Collaborate with law enforcement, the judicial system, civil society and academia to make such data broadly and readily available to support legal investigations, research and the continuous monitoring of the problem.

15. Hire Swedish content moderators who have linguistic and cultural expertise, and ensure there are local points of contact to respond to women journalists reporting incidents with capacity for escalation.

Swedish civil society actors should:

16. Facilitate multi-stakeholder collaborative efforts and develop training, resources and policies within industry and other relevant actors to ensure more gender-sensitive and effective responses to online violence targeting women journalists.

17. Support the building of evidence-based and research-derived knowledge resources that serve to establish a more in-depth understanding of online violence against women journalists and crimes against journalists more broadly.

1.12 The Philippines

Liana Barcia, Julie Posetti, and Fiona Martin

It's like I have no appetite to write. I've really lost the drive. I keep asking myself, why is that? What's happening to me? Am I afraid?...My fear is not just for myself, but also for my family, for my organisation.

Margarita Valle, reporter and columnist for Davao Today and Sun Star

This summary draws on a detailed case study about online violence against women journalists in the Philippines, involving extensive desk research and in-depth interviews with 10 journalists and experts.

Challenges faced by women journalists

Interviewees for this case study shared the view that a culture of sexism and misogyny dominates the Filipino media industry - both in the field, and in newsrooms. This is despite the fact that the Philippines is ranked 16th globally and first in Asia according to the Global Gender Gap Report 2020 (WEF, 2020), and women are also relatively visible in the news industry, representing more than 50% of the country's reporters, researchers, correspondents, and anchors (Tripon, 2015). Gendered online violence, harassment, physical attacks and killings continue. 86 killings of journalists were recorded between 2006 and 2020 by UNESCO's observatory of killed journalists.[243]

Women journalists in the Philippines are frequently subjected to slurs, insults, rape threats, and personal attacks regarding their physical appearance, their reputation, their personal relationships. And those belonging to multiple marginalised groups face compounding forms of violence, such as homophobic abuse. The word "presstitute" is often used by attackers trying to demean women journalists by sexualizing them, or making them appear incompetent or stupid. Another form of abusing journalists in the Philippines is "red-tagging" - the practice of publicly labelling individuals and organisations as "enemies of the state", "communist terrorists", or "members of communist front organisations". "Red-tagged" journalists become open targets for violence online and offline. Recent killings of human rights activists have also been linked to this practice, highlighting the risks facing women journalists who are targeted in this way.

The Philippines is a country with heavy internet and social media usage, and a recent history of coordinated online violence (Posetti et al., 2021). Facebook is the most-used platform in the country, and it serves as many users' portal to the internet as well as a primary source of news. This position of dominance in the

---

social media landscape also makes Facebook the main vector of online violence against women journalists in the country. The main perpetrators, however, are political actors, including with senior government officials, who become top instigators and amplifiers of online violence against women journalists. Several interviewees for this study linked the deterioration of the online environment for women journalists to a series of statements by officials that were hostile to independent media, and to women journalists in particular. Most interviewees stated that online attacks against Filipino journalists are primarily politically motivated, and that those who cover sensitive issues or report critically about governmental matters are the primary targets, with women journalists being disproportionately targeted.

Coordinated, State-linked trolling targeting critical independent journalists and news outlets like Rappler exacerbates general hostility towards the press, fanning public distrust in journalism. In spite of Filipinos' high level of interest in the news (95%°), trust remains low, at only 27% according to the Reuters Institute (Chua, 2020). 'Patriotic trolling',[244] abuse and reputational attacks are used against government critics. As a result of these types of abuse, interviewees reported that women journalists become increasingly cautious and selective about what they post online, as well as tending to self-censor, and at times avoiding social media altogether.

Institutional and legal response mechanisms

Several legal frameworks exist for the protection of women journalists in the Philippines, but gaps in their implementation remain. Legislation such as the Cybercrime Prevention Act of 2012 has been described as presenting a "double-edged sword" by many media practitioners worried about ways in which it can be weaponised to stifle press freedom. A case in point is the 2020 conviction of Maria Ressa and former Rappler journalist Rey Santos Jr. on criminal charges of 'cyber libel' under this law which could see them jailed for up to six years. Another law that media groups fear will be weaponised is the 2020 Anti-Terrorism Act, whose provisions create a presidentially appointed Anti-Terrorism Council with the power to designate "terrorists" who can be arrested without warrant and detained for up to 24 days without charges. Numerous petitions have been filed to challenge its constitutionality.

With regard to women's rights, the Philippines signed the Magna Carta of Women into law in 2009, a piece of legislation which signifies the country's adherence to international human rights frameworks on women's rights. However, it does not contain provisions to deter online violence. By contrast, the 2019 Safe Spaces Act expands the scope of the Anti-Sexual Harassment Act of 1995 to include cyberspace and it contains a definition of online sexual harassment. At governmental level, a Presidential Task Force on Media Security was created in 2016, which later included online threats in its operational guidelines and protocols. Interviewees for this study were pessimistic about the Task Force's ability to remedy online violence against women journalists in the political climate prevailing at the time this research was conducted.

---

244 Appropriation of notions of national loyalty in order to discredit other actors as being "traitors".

Outside of legislation, various initiatives and coalitions have been launched to advocate for Filipino journalists' safety, although few of them put a major focus on women or the online sphere. The first national Philippine Plan of Action on the Safety of Journalists was launched in 2019, within the framework of the UN Plan of Action on the Safety of Journalists and the Issue of Impunity. Its actions include developing a legal support mechanism and defence fund, as well as promoting the filing of cases related to attacks and harassment.

Interviewees for this country study identified several additional challenges standing in the way of effectively resolving the issue of online violence against women journalists in the Philippines. Among these are: the role of political actors in instigating online violence against women journalists; the rapidly evolving nature of both social media and online attacks; weak platform policing of communications involving gendered hate speech and incitement to violence; low levels of social media literacy among Filipino audiences; and the need to raise awareness among officials in cybercrimes units on social media literacy and gender sensitivity.

Recommendations for action in the Philippines

The Philippines authorities and lawmakers should:

1. Require accountability for tech firms.

2. Provide funding to roll out training initiatives for law enforcement, newsrooms, and all other actors implicated in addressing online violence against women journalists.

3. Penalise political actors who target women journalists online.

Filipino law enforcement and judicial actors should:

4. Ensure the prosecutors and agencies tasked with implementing existing laws have the required digital skills and access to the appropriate technology.

5. Acquire awareness about the phenomena of red-tagging and 'lawfare' associated with targeted online attacks against women journalists in the Philippines.

Filipino media organisations could:

6. Conduct training connecting physical, psychological and online risk management, also recognising the risk of online violence against women journalists transferring to offline violence in a country with a very high impunity rate for crimes against journalists. Prepare for worst case scenarios and how to deal with them.

7. Expand solidarity initiatives to ensure those women journalists under attack online are supported and defended in real time.

8. Use investigative journalism to expose threats to press freedom and networks behind coordinated online attacks against women journalists.

Social media companies providing services in the Philippines should:

9. Play a central role in tackling the abuse, disinformation and threats shared on their platforms: "Recognise that decisions made in Silicon Valley ricochet globally, with sometimes devastating negative consequences for women journalists and journalism".

10. Pursue takedowns of public officials / extremist actors and networks involved in gendered hate and harassment.

11. Ensure that monitoring and reporting mechanisms are capable of operating in Tagalog and the Filipino language and ensure that appropriate in-country human responders with press freedom knowledge and gender-awareness are available to escalate online violence cases.

12. Record gender-disaggregated data.

Civil society should:

13. Facilitate connections to international peer support networks, support initiatives to increase social media literacy and raise further awareness about the consequences of online violence and its links to offline attacks.

14. Support social media literacy and training programmes, acknowledging the particular threats posed by viral disinformation to independent journalism and democracy as it manifests as a feature of targeted online violence.

15. Provide specialised and regular training in the area of cybersecurity and gendered online violence, especially for freelance and community journalists.

1.13 Tunisia

Nermine Aboulez

It's normal to be yelled at and humiliated, and if you complain, they tell you that you should treat him [the harasser] like your father, or big brother and forgive him!

Khaoula Boukrim, Editor-in-Chief

This summary draws on a detailed case study about online violence against women journalists in Tunisia involving extensive desk research and in-depth interviews with eight journalists and experts.

Challenges faced by women journalists

Victim-blaming and inflicting reputational damage on women journalists normalise harassment and discourage victims from reporting incidents according to the Tunisian interviewees, who described a climate of impunity for online violence in Tunisia. Sometimes the police themselves tell women to "'ask for God's forgiveness' and calm down…you have no one but your husband, or boss, or whatever…just let it go" (El-Ghadamsy, 2021; Parker and Fahim, 2021; Halouani et al., 2019; Douki Dedieu, 2013). Tunisia is praised for its progressive social policies advocating for gender equality, however, patriarchy and misogyny coupled with particular interpretations of Islamic law are variables that might explain the challenges for an efficient implementation of these policies (Halouani et al., 2019). Interviewees also described the role of family pressure as worsening the impacts of online violence.

Four out of five Tunisian women have experienced online violence, and 51% have been abused on Facebook, according to a 2019 study by the Tunisian Center for Research, Studies, Documentation and Information on Women (CREDIF), released by the Tunisian Ministry of Women, Children and the Elderly.[245] And on the tenth anniversary of the 2011 Tunisian "Jasmine Revolution", nationwide protests exacerbated online and offline violence against women journalists.

The forms of online violence against Tunisian women journalists range from sexist slurs,[246] honour shaming, accusations of lack of patriotism, digital security breaches and blackmail, to violent rape and death threats which can either be random individual attacks or orchestrated campaigns designed to discredit them professionally and chill their reporting. One Tunisian journalist has written that the ultimate goal of online violence in the country is to silence and eliminate women

---

245 The number of respondents is unknown.
246 Comments such as "return to your kitchen", "go wash the dishes", "who is this blondie?" are examples of some such sexist slurs described by interviewees. (Landorf, 2014). In March 2020 Malak al-Bakari, a journalist at the privately-owned Attessia TV channel, was abused online for wearing a short dress (Jouini, 2020).

from public spaces (Boukhayatia, 2020). One example of the sexist body-shaming endured by Tunisian women journalists online is a Facebook-based campaign against broadcast journalist Fadoia Chtourou for not wearing makeup while reporting on TV "looking like that!".[247] Another incident involved using a woman radio host's name and photo to share obscene stories within Facebook groups.

Harassment campaigns against women journalists often involve 'patriotic trolling' narratives designed to question the target's loyalty to her country or culture. Former Nawaat and Shems FM radio journalist Henda Chennaoui was accused of "receiving money from foreign NGOs" and concocting "an act of conspiracy against Tunisia with the aid of Western countries". Khaoula Boukrim, Kashf media Editor-in-Chief, was branded a "leftist, an agent, pro-Emirati, [with] a secret agenda", "a mercenary" who was "not patriotic" and a traitor for criticising a political announcement on national television and for supporting the hashtag #FreeSolafa.[248] Another technique used against women journalists in Tunisia is feeding them false information to trap them into sharing false news and taint their credibility, according to investigative journalist Najoua Hammami. Recalling what she believes was an assassination attempt, when a van rammed into her car after she reported on the involvement of a member of parliament in a smuggling operation involving ancient artifacts (ARCS, 2014), Hammami also highlighted the tendency for online violence to spill offline, endangering women journalists' physical safety. She also described being kicked and beaten with electrifying sticks and knocked down by security forces while covering protests, causing injuries from which she still suffers.

Institutional and legal response mechanisms

Some Tunisian news organisations offer moral support and defend women journalists against harassment. For example, Fadoia Chtourou praised her Editor-in-Chief for defending her in response to abusive online comments. However, other interviewees shared examples of news organisations denying affiliation with journalists when they are experiencing attacks. Some news organisations are even a significant source of harassment for women journalists. One example came from Henda Chennaoui, who said a colleague tried to beat her for sharing pro-revolution views online - an act which ultimately silenced her. Tunisian women journalists also have longer working hours, lower pay and less access to permanent contracts and insurance compared to their male colleagues, which make them more vulnerable to harassment (Ghiloufi, 2021; Freedom House, 2021).

Tunisia has a strong civil society which received Nobel Peace Prize recognition in 2015. The National Syndicate of Tunisian Journalists (SNJT) is the most common refuge for women journalists seeking support, but Access Now, the Union of Arab Journalists, Union of African Journalists, UNESCO, the International Union for Journalists, ARTICLE 19, and Reporters Without Borders (RSF) were also mentioned. While most women journalists interviewed praised SNJT for their support, others felt let down. Khaoula Boukrim, for one, said "they didn't even comment" in response to the fierce campaign accusing her of anti-patriotic views. The intervention of international organisations is sometimes crucially needed to

247 Fadoia Chtourou works for the Tunisian public broadcaster Wataniya 1.
248 Solafa Magdy is a detained Egyptian journalist, who has testified about being sexually assaulted and severely abused in prison.

protect women journalists, since "nothing happens until international organisations intervene," according to Najoua Hammami.

Tunisian civil society organisations do not have direct contact with social media platforms and therefore, the onus of reporting abusive content to platforms rests on the individual, Faouzia Ghiloufi said. When women journalists do take legal action against online violence, the content is usually not removed without court orders. Conversely, Facebook, has been accused of censoring journalists' accounts and deplatforming them without providing any justifications, according to interviewees for this study.

Sexual harassment is punishable in Tunisia by two years in prison and a fine of 5,000 dinars (US $1837.90) (Halouani et al., 2019). However, there is a lack of awareness of this legislation and laws criminalising violence against women. Judicial institutions, which have been marked by a patriarchal system, and officials also often attempt to reconcile the parties and are reluctant to impose sanctions. This leaves women who choose to report vulnerable to pressure (Doelker and Scheicher, 2020).

Recommendations for action in the Tunisian context

The Tunisian authorities and lawmakers should:

1. Prohibit government officials and civil servants from perpetrating gendered online violence against women journalists, applying penalties as a deterrent.

2. Provide training for law enforcement, judges and prosecutors about the laws on gender-based violence, so that they possess the necessary understanding and digital capabilities to investigate gendered online violence and connect it to offline risks, and to allow them to be held accountable for inaction.

3. Create more effective legal mechanisms to remove offensive content more quickly when it is found to be misogynistic or otherwise hateful, or uses physical or sexual violence threats (while protecting freedom of expression rights according to international standards), and speed up redress in cases of online violence.

News organisations should:

4. Provide mandatory sexual harassment training (integrating online and offline issues) for all staff to ensure a safe working environment for women journalists and penalise for sexist or misogynistic conduct (online or offline) regardless of rank.

5. Ensure women journalists (including freelancers) receive equal pay, and parity with their male counterparts on working hours, insurance coverage and access to job security.

Civil society organisations could:

6. Fund independent media outlets and ensure the independence and professionalism of the outlets they fund, which could include an emphasis on coverage of gender and disinformation issues.

7. Develop contextual responses to intersectional forms of online violence affecting women journalists.

Social media companies providing services in Tunisia should:

8. Create an autonomous, accountable, transparent and gender-balanced regional council, composed of civil society representatives, editors and gender-sensitive freedom of expression experts.

9. Appoint staff in Tunisia, with cultural and social contextual knowledge, and linguistic expertise to improve complaints responses, content moderation, and algorithmic renovation in the interests of better addressing the problem of online violence against Tunisian women journalists.

10. Be more proactive in responding to reports of online violence from Tunisian women journalists under attack, providing channels for urgent escalation in recognition of offline risks associated with online attacks.

1.14 United Kingdom of Great Britain and Northern Ireland

Julie Posetti and Nabeelah Shabbir

I do believe that the online violence against me has created real life threats that are inciting criminals and very dangerous loyalist paramilitaries to issue threats against me. And that's put my life in grave danger.

Patricia Devlin, Sunday World

This summary draws on a detailed case study about online violence against women journalists in the UK involving extensive desk research and in-depth interviews with 22 journalists and experts.

Challenges faced by women journalists

Online violence against women journalists in the UK is frequently associated with populist politics, and polarising political debates - such as the 2016 Brexit referendum. In Northern Ireland, sectarian violence has caught the press in its sights and organised crime is also a significant factor. The pandemic has worsened the situation for women journalists in the UK, with digital conspiracy networks targeting reporters, and one member of the Conservative government at the time attacking two of this study's interviewees for their critical reporting of COVID-19 responses. More than half (51%) of 319 participants[249] in a 2020 National Union of Journalists (NUJ) survey had experienced online abuse in the previous year, while 94% of respondents agreed that the polarisation of democratic debate and public discourse in the UK had impacted adversely on the safety of journalists, and 78% agreed that "abuse and harassment has become normalised and seen as part of the job" within UK journalism (Lezard, 2020). The situation has become so serious that the UK's biggest commercial news publisher (Reach, PLC) created the position of Online Safety Editor in mid 2021.

Women journalists in the UK are acutely affected by online violence, especially at the nexus of misogyny, racism, xenophobia and religious bigotry. The discrimination that black and minority women journalists struggle with offline is amplified and exacerbated online, where they face a triple burden of intersectional abuse: they are attacked on the basis of their gender, their status as journalists, and because of the colour of their skin. In some cases, they are simultaneously abused because of their religion and their sexual orientation. So, while black and minority women journalists are largely absent in prominent and senior British media roles, they

249 An online questionnaire was completed by 319 UK-based NUJ members in September/October 2020, asking questions about different types of abuse and harassment, and how to tackle the problems. 62% identified as male and 35% as female (NUJ, 2020).

are among the most clearly identifiable targets of gendered online violence. This reality was underscored by interviewees' testimonies collected for this study. "It has become quite toxic, especially towards women, and especially towards women of colour," Vice UK Executive Editor-in-Chief Zing Tsjeng said.

The correlation of online violence and offline attacks against women journalists is significant in the UK, especially with regard to stalking. This study documented four cases of physical stalking associated with online harassment in the UK. In three cases, the perpetrators harassed the journalists at their newspapers,[250] and in the fourth case a stalker left a message for a BBC journalist on the noticeboard at the train station she uses to commute to work.[251] Northern Irish journalists covering sectarian paramilitary activity and organised crime are particularly vulnerable to offline attacks seeded online. Three cases of online-offline violence targeting women journalists in Belfast surfaced during the course of the research. The case of one interviewee, *Sunday World*'s Patricia Devlin is emblematic.[252] She received credible death threats online which were accompanied by graffiti on city walls featuring her name next to the crosshairs of a gun, and threats of sexual violence were also made against her children. The Police Ombudsman is investigating complaints made by Devlin about the Police Service of Northern Ireland's handling of online threats targeting the journalist and her family, while Facebook has been heavily criticised for failing to act effectively and with urgency to deal with perpetrators.

One interviewee for this study, BBC disinformation reporter Marianna Spring, documented her own experience of online violence in a documentary broadcast in late 2021, which also highlighted the ways in which Facebook and its Instagram app fuelled misogynistic hate speech online through algorithm-driven referrals (BBC Panorama, 2021b).

Women journalists working in the UK have been targeted by foreign State actors, including a number of BBC Persian service journalists who have been subjected to coordinated campaigns linked to Iranian authorities, according to a report from four UN Special Rapporteurs (OHCHR, 2020g). One of the BBC Persian service journalists received threatening messages attached to images depicting Iranian journalist Ruhollah Zam, who was kidnapped, renditioned, and executed in Iran 2020 (OHCHR, 2020g; BBC, 2021a). Patricia Devlin is another example of a UK woman journalist being issued an online death threat in the form of references to journalists murdered with impunity. She received a death threat warning her that she would "end up like" *Sunday World* investigative journalist Marty O'Hagan,[253] who was assassinated in Northern Ireland in 2001.

New sources and methods of online abuse described by interviewees include YouTubers, social media influencers, 'stan armies',[254] and fake blog posts written in the names of journalists. UK journalists interviewed described hiring private

---

250 See the cases of Rosamund Urwin, Lizzie Deardon, and Amy Fenton.
251 See the case of Marianna Spring.
252 Similar cases include the journalists Leona O'Neill and Allison Morris. See the case study featuring Patricia Devlin in Chapter 6.0 (Platforms and vectors: Assessing big tech responses to online violence).
253 Since the assassination of Martin O'Hagan near Belfast in 2001, the first killing of a journalist in the line of duty in the UK was Lyra McKee in Derry in 2019. No one has ever been convicted of O'Hagan's murder, which provides a chilling parallel to the online anonymous abuse that Patricia Devlin, employed at the time of writing at his newspaper, the *Sunday World*, gets (RSF, 2020a).
254 'Stan' is considered a portmanteau of 'stalker' and 'fan' - 'stan armies' are online communities dedicated to individual celebrities or groups. See: https://www.nbcnews.com/pop-culture/pop-culture-news/year-stan-how-internet-s-super-fans-went-pop-stars-n1252115

security, self censoring, and handing bylines over to other journalists as a means of trying to mitigate the risks they face in connection with online violence.

Institutional and legal response mechanisms

The murder of journalist Lyra McKee, who was killed in crossfire while reporting on paramilitary violence in 2019, signalled an urgent need for the development of more robust mechanisms to respond to digital and physical journalism safety threats (RSF, 2020i; 2021j). The UK National Committee for the Safety of Journalists was established in mid-2020. It produced a National Action Plan for the Safety of Journalists which was finally published as a 'living document' in March 2021 (DCMS, 2021a). It represents an attempt at connecting the UK government, the media industry, three separate policing and prosecution services (for England and Wales, Scotland, and Northern Ireland), and civil society actors to: "make sure journalists in the UK are able to operate free from threats and violence" (DCMS, 2021a).

Citing the aforementioned National Union of Journalists survey (NUJ, 2020), and preliminary research published from this study (Posetti et al., 2020c), the Plan[255] also recognised the surge in online violence against journalists in the UK. The document pays particular attention to the experiences of women and BAME[256] journalists, and describes online abuse as a "stark" risk: "...it is journalists leaving the profession, and a decline in new entrants; it is self-censorship of those that remain; it is anxiety, stress and depression; it is less challenge to those in power; it is a weaker democracy" (ibid.). A Draft Online Safety Bill[257] defined "lawful but potentially harmful" activity as including "online abuse, advocacy of self-harm, spreading disinformation and misinformation" and says it is "behaviour [which] may fall short of amounting to a criminal offence...[but] can have corrosive and damaging effects, creating toxic online environments and negatively impacting users' ability to express themselves online".

Legal pressure is a feature of the offline manifestations of online violence in the UK. Examples include SLAPP suits (Strategic Litigation Against Public Participation), such as those brought by a businessman and political donor associated with a campaign for Britain's withdrawal from the EU against investigative journalist Carole Cadwalladr. But UK journalists have also begun fighting back against online violence through recourse to litigation. In two cases (involving journalists Amy Fenton and Lizzy Dearden), perpetrators of online threats of physical and sexual violence, along with stalking and harassment, have been jailed in the UK in the past two years. In another case, a stalker was held in a mental health facility.[258]

---

255 The Plan builds on information from the NUJ, the National Council for the Training of Journalists (NCTJ), employers, prosecutors, police, and civil society organisations including the News Media Association (NMA), and the Society of Editors.

256 BAME is a UK acronym for 'Black, Asian, and Minority Ethnic' and is often used in the media, politics and government, but there is general discussion over whether it is the "correct" term to use when discussing UK citizens of different ethnic backgrounds. It does not, for example, encompass ethnic minorities such as Gypsy, Roma and Traveller of Irish Heritage groups.

257 This is intended to strengthen the National Action Plan for the Safety of Journalists and calls for protection for "journalism and democratic political debate" (Tobitt, 2021a).

258 See Chapter 6 (Legal and Normative Frameworks for Combating Online Violence Against Women Journalists) for further details.

Recommendations for action in the UK

The UK authorities and lawmakers should:

1. Call out instances when journalists are abused online by UK political actors and breaches of the ministerial code of conduct, especially when it fans intersectional abuse.

2. Establish protocols holding elected representatives to account for online conduct, regardless of whether or not they are using official government accounts to criticise journalists.

3. Support recognition of misogyny as a form of hate speech in law,[259] in line with racism, homophobia, religious bigotry and other forms of discrimination, and as per the UN Strategy and Plan of Action on Hate Speech and the UK House of Lords proposition.

4. Finalise, adopt and enforce the Draft Online Safety Bill, with reference to the assessment tool presented in chapter 7.0 to ensure freedom of expression rights are balanced against the need for legal responses to online violence.

5. Appoint a high-level official to have responsibility for media freedom and journalism safety, including implementation of the National Plan of Action for the Safety of Journalists and accountability mechanisms to ensure more effective and transparent responses to online violence against journalists from the platforms.

6. Pay particular attention to the high-risk situation in Northern Ireland, where there is significant evidence of online violence spilling offline with potentially deadly consequences.

UK media organisations should:

7. Expose targeted, intersectional and multi-platform online violence against their journalists, including where it is perpetrated by political actors, and pursue complaints with the UK's regulatory bodies[260] when other news outlets fuel pile-ons against journalists, as in the cases of Carole Cadwalladr (freelance) and Rianna Croxford (BBC).

8. Establish gender and diversity sensitive 'rapid response' units and employ Digital Safety Editors to oversee preventive and responsive measures.

9. Acknowledge and respond to intersectional threats and impacts, being aware of the role of structural racism and other forms of discrimination in British

---

259 The assassination of MP Jo Cox by a far-right extremist during the 2016 referendum, after she endured months of online threats, triggered a campaign to recognise misogyny as a hate crime in the UK (Saner, 2016; Cobain and Taylor, 2016). In mid-2022, legal reforms were being considered/implemented in England, Scotland and Wales to add sex and gender to the existing list of "protected characteristics" associated with hate crimes in the UK: race, religion, sexual orientation, disability or transgender identity https://www.gov.scot/publications/misogyny-human-rights-issue/; https://www.gov.uk/government/publications/police-crime-sentencing-and-courts-bill-2021-factsheets/police-crime-sentencing-and-courts-bill-2021-making-misogyny-a-hate-crime-factsheet (See also Duffy, 2021).
260 Ofcom regulates the broadcast press (including the BBC) and is to be the main regulator enforcing the Online Safety Bill. The Independent Press Standards Organisation (IPSO) oversees the majority of newspapers and magazines in the UK under a set of rules called the Editors' Code of Practice.

newsrooms, which can leave certain UK women journalists more exposed to online violence.

10. Ensure that women journalists have dedicated work devices to ensure they are able to minimise the impacts of doxxing and provide some separation between personal and professional lives.[261]

Social media companies providing services in the UK should:

11. Adapt protocols in response to threats against women journalists and establish rapid response units in the UK to deal with high-risk cases, such as Patricia Devlin in Northern Ireland.

12. Institute gender-sensitive online violence safety protocols and streamlined abuse-reporting processes as a feature of all partnerships with news organisations e.g., fact-checking initiatives, as well as arm's-length financial support.

13. Prevent their platforms from being used by foreign State actors or criminal forces to threaten and harass journalists based in the UK, recognising that these represent serious threats with potentially deadly consequences.

14. Work with independent academics to analyse data illustrating the extent and nature of gender-based online violence against women journalists and others using their platforms to share public interest information.

UK-based civil society organisations should:

15. Collaborate on efforts to address gendered online violence against women journalists, focusing on those experiencing intersectional threats and impacts, for example by supporting collaborative investigative responses to gender-based online violence.

16. Focus research efforts on data gathering to illustrate the particular experiences and risks faced by women journalists working in the UK, for example by investing in independent public interest journalism focused on addressing gender issues.

17. Hold the UK government to account with regard to implementation of the National Plan of Action for the Safety of Journalists, ensuring that it is implemented in a gender-aware way - it does not currently mention gender, women or girls.

---

261 Two BBC interviewees described having to operate through the pandemic using their personal phones to contact sources.

1.15 The United States
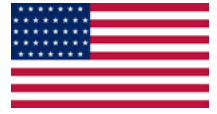
Julie Posetti, Silvio Waisbord, and Nabeelah Shabbir

The consequences of reporting on the far-right are real, and they take a toll... White nationalists think that I don't deserve to live in this country...Being Chinese, being Jewish, being a woman just becomes part of the content of the harassment.

Julia Carrie Wong, Guardian US

This summary draws on a detailed case study about online violence against women journalists in the US involving extensive desk research and in-depth interviews with 14 journalists and experts.

Challenges faced by women journalists

Online violence is the most serious safety threat faced by women journalists in the United States, according to the Committee to Protect Journalists (2019). It appears to have significantly worsened in recent years in the context of growing threats linked to rampant digital misogyny, viral disinformation, online conspiracy cultures, the rise of far-right extremism, and the demonisation of the press as "fake news" by former President Donald Trump. As indicated by various surveys (IWMF, 2018; Westcott, 2019; Women's Media Center, 2020; Posetti et al., 2020b) and reinforced by the US-based interviewees for this study, a majority of women journalists polled in the US have experienced online violence.

The methods and types of online violence experienced by women journalists in the US vary from rudimentary forms of abuse to high-tech attacks. They encompass numerous threat types - from sexualised online harassment and abuse, to vicious threats of sexual assault and murder, targeted disinformation campaigns amplified by partisan right-wing media and misogynistic journalistic actors, along with digital safety and privacy breaches. Examples include: threats of murder (e.g., using graphic images of dead bodies and credible bomb threats); threats of extreme sexual violence; doxxing and other forms of digital security attack, such as targeted surveillance and spoofing; coordinated/orchestrated attacks - 'pile-ons'/ 'brigading'/ 'dogpiling', including the use of misogynistic fora like 'KiwiFarms'[262] to identify targets and plot attacks; and using audio apps such as Google Voice and Clubhouse which, at the time of writing, did not have abuse reporting mechanisms.

---

262 'Kiwi Farms' is a forum for organising the group trolling of people and communities it holds up for ridicule and harassment at scale.

These assaults frequently radiate, with family members and colleagues of women journalists under attack also turned into targets.[263] Additionally, there are alarming examples of online violence against journalists in the US spilling offline, including the Capital Gazette newsroom massacre of 2018[264] (Robertson, 2018), assaults on journalists in the field, and a credible bomb threat against a White House correspondent.[265]

While all of the US interviewees identified misogyny, sexism and anti-journalism narratives as primary features of attacks against women journalists, they also highlighted the disproportionate exposure to online violence among women who face intersectional discrimination. At the nexus of sexism, misogyny, racism, homophobia, and religious bigotry, the abuse and harassment of women journalists is often more extreme; the impacts more severe.[266] Deep fakes, memes and other disinformation tactics used to shame and humiliate US women journalists at personal and professional levels often occur at the intersection of other forms of discrimination e.g., abusive memes melding misogyny, racism, and antisemitism.

In the US, political actors - especially on the political fringes - play a significant role as instigators and amplifiers of online attacks against women journalists. This was a particularly pronounced trend during the Trump administration, which saw the former president abuse women journalists online and offline, inciting pile-ons and fuelling the misogyny, racism, and anti-mainstream press sentiment which overlap in digital attacks. The role of the right-wing media ecosystem in fuelling politically motivated attacks on women journalists in the US - either triggering or fanning pile-ons - is also noteworthy (e.g., Fox News; OANN). Far-right and white supremacist groups, partisan media, networks of anonymous or pseudonymous misogynists, and tech entrepreneurs, along with sources, and toxic audiences, were also identified by the US-based interviewees as prime sources of online violence.

It affects women journalists in the US in very tangible ways. Psychological effects such as anxiety, depression, extreme fear and Post Traumatic Stress Disorder (PTSD) were some of the impacts that the women journalists interviewed were experiencing. The ripple effects of such injuries included forms of self-censorship; removal of their bylines; and their retreat to less exposed roles. The threat of online violence triggering offline attacks led a number of interviewees to move house, go into hiding, and increase physical security - at home and at work.

Institutional and legal responses

The role of the largely US-based social media companies in facilitating the global scourge of online violence against women journalists is highly relevant. Inaction by the companies serves in effect to privilege the free speech "rights" of abusers to attack women journalists over and above the rights of the targets to the safe use of freedom of expression on the very same services. They have used the First Amendment and other legislative mechanisms (e.g., Section 230 of the Communications Decency Act) to avoid responsibility for protecting women

---

263 See, for example, the case of Al Jazeera's White House Correspondent Kimberly Halkett.
264 The attack claimed the lives of five journalists, including two women, after they profiled a man who had spent a year cyberstalking a former school acquaintance he had found on Facebook, and was convicted of criminal harassment. The man railed against the journalists in tweets and sent death threats, which he eventually acted upon: https://www.nbcnews.com/news/us-news/woman-harassed-capital-gazette-shooting-suspect-says-he-tormented-her-n888196
265 White House Correspondent with The Grio and formerly CNN, April Ryan, was the target of a credible bomb threat.
266 See, for example, the cases of April Ryan, Julia Carrie Wong, and Apoorva Mandavilli.

journalists using their platforms. Journalists have been assumed to be "public figures" who should expect to be attacked, providing a high bar for acting against users perpetrating online violence against journalists. Under mounting pressure to deal with gendered online violence against journalists on the platform, Facebook declared in late 2021 that rape threats against journalists on its services would no longer be tolerated.

In the US, there is also the added burden of a misapprehension that targets of online violence have to absorb the abuse and have limited legal recourse because of the false binary notion that the first amendment protects the rights of their attackers above their right to practice journalism free from hate speech and other forms of violence. This results in a reluctance among news organisations and individual journalists to report incidents to the police or pursue perpetrators and facilitators of online violence through the courts.

Finally, the pattern of news organisations deploying victim-blaming and speech policing in response to gendered online violence against women journalists in their employ, is particularly evident in the US context. There are recorded examples of women journalists being suspended - or even sacked - in the midst of targeted attacks, often as a result of perceived breaches of restrictive social media policies. Online violence campaigns targeting the women's employers - seeking the journalists' silencing and dismissal - are part of this problem.

Recommendations for action in the US

US authorities and lawmakers should:

1. Recognise that US-based social media companies are primary enablers of online violence against women journalists, and that their business models and conduct not only chills legitimate free expression, it places women journalists' lives, well-being, and livelihoods at risk.

2. Conduct public hearings into the impacts of the platforms' systems on journalism safety, press freedom, media diversity, and the public's right to access public interest information; consider reforms to legal and regulatory regimes to ensure these companies practice transparency and accountability.

3. Require elected US officials to refrain from instigating or perpetrating online violence against journalists, using the example of the ongoing impacts of former President Trump's abuse of journalists with impunity.[267]

4. Implement the National Task Force on Online Harassment and Abuse campaign promise of November 2020, which would recommend how governments, social media companies, schools, and other public and private organisations should address online harassment in the upcoming National Action Plan to End Gender-Based Violence.

---

267 https://www.vice.com/en/article/4avqqn/reporters-survived-capitol-riot-struggling.

US news industry should:

5. Recognise that critiquing and reporting perpetrators (e.g., seeking their deplatforming), and taking defensive legal action in response to targeted hate speech against women journalists is not antithetical to "free speech" as protected under the US Constitution.

6. Disavow hate speech against women journalists - including that practiced by competing journalists and news organisations, recognising that counterspeech does not amount to editorial 'bias'.

7. Support women journalists under attack and recognise their right to push back on social media, and avoid disciplining them or ending their employment contracts in the context of such attacks.

8. Provide financial legal aid, and health benefits where needed, to freelance journalists when they come under attack while on assignment, or following the publication of a commissioned story, especially given the absence of universal health care in the US and employment precarity within journalism.

9. Hold Big Tech to account through critical, independent journalism - regardless of commercial arrangements between the news organisation and the companies, recognising the particular agency afforded to US-based news organisations in this regard, and the implications for international news publishers.

10. Advocate within the newsroom, and with the concerned platform/s, on behalf of a woman journalist who is especially at risk of being targeted with intensified online abuse because of their religious or ethnic background, gender identity or sexual orientation.

Social media companies providing services in the US should:

11. Establish human-centred, specialist rapid-response units to address online violence against women journalists, with gender-sensitive mechanisms and press freedom awareness and recognition of particular intersectional threats, and deliver timely action and reporting to US journalists, news organisations and civil society organisations.

12. Avoid using "free speech" as a justification to allow the continuation of attacks on journalists on their platforms which enable online violence, viral disinformation and hate speech.

13. Recognise that decisions made in Silicon Valley ricochet globally, with sometimes devastating negative consequences for women journalists and journalism.

14. When considering experimenting with algorithmic shifts and services like Facebook Free Basics in developing countries, ask not only "What are the implications for press freedom, democracy and the safety of journalists and human rights defenders?", but also: "Would we experiment like this in the US? If not, why not?"

US-based civil society organisations should:

15. Support collaborative investigative responses to gender-based online violence to expose and denounce perpetrators and instigators.

16. Advocate for news organisations to implement common protocols to address online violence and fund/aid the development of such protocols, along with training to embed them.

17. Lobby philanthropic foundations and State actors to make the issue of online violence a programmatic priority.

ABOUT THE EDITORS

Dr. Julie Posetti

is Deputy Vice President and Global Director of Research at the International Center for Journalists (ICFJ). A multi award-winning feminist journalist, researcher and journalism educator, Dr. Posetti is academically affiliated with the Reuters Institute for the Study of Journalism at the University of Oxford and the Centre for Freedom of the Media (CFOM) at the University of Sheffield. She was previously a news, current affairs and documentary reporter and editor for Australia's national broadcaster ABC. Dr. Posetti's journalism and commentary has also been published by *The Washington Post*, *The Guardian*, *The Atlantic*, BBC, *Foreign Policy* and *The Sydney Morning Herald*, among others. She is author of *Protecting Journalism Sources in the Digital Age* (UNESCO: 2017), and co-author of *Journalism, Fake News and Disinformation* (UNESCO: 2019) and Balancing Act: Countering Disinformation While Respecting Freedom of Expression (UNESCO-UN Broadband Commission: 2020).

Nabeelah Shabbir

is Senior Research Associate at the International Center for Journalists and a British-Pakistani freelance journalist. At *The Guardian*, she shared a British Journalism Prize with the 'Keep it in the Ground' team in 2015. She has co-authored a series of reports for the Reuters Institute for the Study of Journalism at the University of Oxford. She has also worked at The Correspondent, the *Financial Times*, and Twitter.

ABOUT ICFJ

ICFJ empowers an unparalleled global network of journalists to produce news reports that lead to better governments, stronger economies, more vibrant societies and healthier lives. ICFJ also helps improve the working lives of journalists through 'action research' collaborations on the safety of women journalists, disinformation and trust, and journalism in a post-pandemic world.

ABOUT UNESCO

UNESCO is the UN agency with the mandate to protect Freedom of Expression and spearheads the implementation of the UN Plan of Action on Safety of Journalists and the Issue of Impunity. UNESCO highlights the centrality of developing specific policies for the safety of women journalists and cooperates with partners to identify and implement good practices and share recommendations in countering violence against women journalists, online and offline.