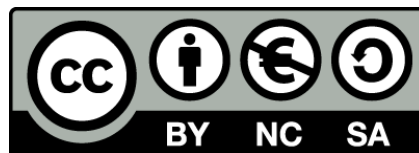




UNIVERSITAT DE
BARCELONA

Beyond the screen
Digital realities and embodied harm in the experiences
of gender-based online hate speech

Malin Roiha



Aquesta tesi doctoral està subjecta a la llicència **Reconeixement- NoComercial – Compartir Igual 4.0. Espanya de Creative Commons.**

Esta tesis doctoral está sujeta a la licencia **Reconocimiento - NoComercial – Compartir Igual 4.0. España de Creative Commons.**

This doctoral thesis is licensed under the **Creative Commons Attribution-NonCommercial-ShareAlike 4.0. Spain License.**



UNIVERSITAT DE
BARCELONA

Doctoral Thesis

Beyond the screen

Digital realities and embodied harm in the
experiences of gender-based online hate speech

Malin Roiha

Supervisor: Olga Jubany



UNIVERSITAT DE
BARCELONA

TESI DOCTORAL

Beyond the screen

Digital realities and embodied harm in the
experiences of gender-based online hate speech

Autora: Malin Roiha

Directora: Olga Jubany

Barcelona, 2024

Programa de doctorat Societat i Cultura: Història, Antropologia, Art i Patrimoni
Departament d'Antropologia Social
Facultat de Geografia i Història
Universitat de Barcelona

“The form is new, but the violence is old”
(DE-18 – Feminist journalist¹)

¹ This is a quote from the digital ethnography. As discussed in Chapter 4 on methodological and ethical considerations, codes have been used to protect the anonymity of the research participants.

Contents

Abstract	11
Resumen	13
Acknowledgements	15
1. Introduction	17
1.1. Setting the scene: Feminism and anti-feminism	18
1.2. Origins of the research and initial questions	25
1.3. A note on the vocabulary used in this thesis	26
1.4. At the intersection of hate crimes and gender-based violence	28
1.5. Objectives of the research	32
1.6. Explanation of the thesis structure	33
2. Constructing a Conceptual Framework: Gender, Technopolitics and Violence	39
2.1 Introduction	39
2.1. Identity and gender constructions in a globalised world	40
2.2. Technopolitics and digital social inequalities	46
2.3. Conceptualisations of violence	58
2.4. Conclusion: An intersectional technofeminist framework	70
3. From Definitions to Debates	71
3.1 Introduction	71
3.2 An overview of concepts in relation to online hostility: From flaming to hate speech	72
3.3. Gender-based online hate speech	86
3.4. Towards a working definition of gender-based online hate speech	103
3.5. Rationale of the thesis	105
4. A continuum of online – offline ethnography: Methodological and ethical considerations	109
4.1 Introduction	109
4.2 Reflexive ethnography and the intersectional feminist lens	111
4.3 Methods and sample	114
4.4. Data management and analysis	130
4.5 Ethical considerations	131
5. The legal framework: Overview, interpretations and gaps	137
5.1. Introduction	137
5.2. Freedom of expression versus the right to dignity	139
5.3. Hate crimes and hate speech in the Spanish criminal code	146
5.4. The legal framework on gender violence in Spain	153
5.5. Laws from autonomous communities	155

5.6. Other tools and mechanisms	156
5.7. Perceptions and interpretations of legal protection mechanisms for victims of gender-based online hate speech	159
5.8. Conclusion: A fragmented legal framework	169
6. Professionals' perspectives on online gender-based violence	173
6.1. Introduction	173
6.2. Perceptions and categorisations of gendered online violence	174
6.3. Preparedness to respond to online violence	179
6.4. Perceptions on reporting and responses by the police	181
6.5. Training needs and possible ways forward	187
6.6. Conclusion: A hierarchy of violence, challenges of technological development and protection gaps	189
7. Targeted voices: Experiencing gendered online violence	193
7.1. Introduction	193
7.2. Life online: Between work and activism - Between public and private	195
7.3. Being targeted as a feminist-, anti-racist or LGBT+ activist	199
7.4. Gaming, dick pics and unsolicited contacts: Other gendered experiences of violence	211
7.5. Intersectionality matters	215
7.6. Exploring the perpetrators through the eyes of the targets	218
7.7. "Real" violence vs "not real" violence	228
7.8. The normalisation of online violence	231
7.9. Conclusion: Unveiling a continuum of gendered violence across and beyond digital spaces	233
8. Digital harm, real impact: Unmasking the impact of gender-based digital harm	237
8.1. Introduction	237
8.2. Embodied experiences of online harm	238
8.3. Being on guard	241
8.4. Internalisation and insecurity	243
8.5. Self-censoring	246
8.6. Going offline - Leaving Twitter	248
8.7. The wider consequences of online violence for the participation of voices in the public debate	252
8.8. Conclusions: Lingering emotional impacts, ongoing vigilance, and the chilling effect of online abuse	257
9. Navigating Online Gender-Based Hate Speech: Responses and Contestations	261
9.1. Introduction	261
9.2. Digital self-protection strategies	262
9.3. Publicly speaking out	274

9.4. Building community	282
9.5. Interacting with the justice system	293
9.6. Conclusion: Digital protection, community and impunity	303
10. Recapitulation and concluding observations	305
10.1. Recapitulation	305
10.2. Along the continuum of violence, spaces, discourses, and self-protection mechanisms	307
10.3. Intersectional dynamics	308
10.4. Embodied and collective harm	309
10.5. Navigating self-protection, platform dynamics, and institutional challenges	311
10.6. Legal and institutional responses	314
10.7. Final reflections and future research paths	315
References	321

Abstract

This dissertation examines the phenomenon of gender-based online hate speech and its intersectional dynamics, focusing on online violence directed at feminists in Spain. The study is situated within the field of hate studies, aiming to contribute to the understanding of misogynist and intersectional violence within the hate crime framework.

Employing a “post digital” ethnographic methodology across digital and non-digital spaces, the study employs in-depth interviews with feminists and activists subjected to online violence, along with digital ethnography, to illuminate their lived experiences of continuous exposure to violence. Interviews with professionals and legal experts provide complementary perspectives from institutions and civil society organizations on available support.

The thesis highlights how digital spaces provide accessible avenues for opponents to express ridiculing discourses and dehumanising speech towards targeted groups with impunity, often through collective attacks. The findings underscore the continuum of gendered violence, where explicit verbal attacks blend with subtler forms of aggression, such as unsolicited sexualised contacts. This continuum serves as both a manifestation and perpetuation of gendered power relations. Furthermore, the study underscores the temporal dimension of online violence, emphasising its continuous nature rather than being isolated incidents. It demonstrates how the impact of online attacks on women extend beyond digital platforms, revealing how technology-facilitated verbal harms manifest as embodied harms, impacting both digital communication and non-digital bodies and interactions. This harm constitutes a linguistic injury that affects both individuals and collectives. It highlights the intersectional dynamics of technology, gender and other identity factors such as racialisation, sexual orientation, gender identity, and activism levels. Further, the thesis explores both informal individual and collective contestations, as well as institutional responses, revealing how targeted women rely on digital self-protection strategies and peer support, in the absence of efficient responses by institutions and social media platforms.

In conclusion, the dissertation argues for interpreting digital interactions as embedded within broader socio-political contexts, advocating for structural and collective responses to address online violence and protect vulnerabilised individuals and groups across online and offline spaces. The thesis advances our understanding of gendered violence in the digital age, promoting interdisciplinary and intersectional approaches to tackle this complex issue.

Resumen

Esta tesis doctoral examina el fenómeno del discurso de odio *online* por razón de género y sus dinámicas interseccionales, centrándose en la violencia dirigida a feministas en España. El estudio se sitúa dentro del campo de estudios sobre el odio, con el objetivo de contribuir a la comprensión de la violencia misógina e interseccional dentro del marco de los delitos de odio.

Empleando una metodología etnográfica "posdigital", el estudio se basa en entrevistas en profundidad con feministas y activistas sujetas a violencia online, junto con una etnografía digital, para iluminar sus experiencias vividas de exposición continua a la violencia. Las entrevistas con profesionales y expertos legales proporcionan perspectivas complementarias de instituciones y organizaciones de la sociedad civil sobre el apoyo disponible.

La tesis destaca cómo los espacios digitales proporcionan vías accesibles para que los oponentes expresen discursos de ridiculización o deshumanizantes hacia grupos específicos, con impunidad, y a menudo a través de ataques colectivos. Los hallazgos subrayan el continuum de violencia, donde los ataques verbales explícitos se mezclan con formas más sutiles de agresión, como contactos sexualizados no solicitados. Este continuum sirve como manifestación y perpetuación de las relaciones de poder de género. Además, el estudio enfatiza el aspecto temporal de la violencia digital, resaltando su naturaleza continua en lugar de incidentes aislados. La investigación demuestra cómo el impacto de los ataques online contra las mujeres se extiende más allá de las plataformas digitales, revelando cómo los daños verbales facilitados por la tecnología se manifiestan como daños encarnados, afectando tanto la comunicación digital como las interacciones no digitales. Destaca la interseccionalidad de la tecnología, el género y otros factores de identidad, como la racialización, la orientación sexual, la identidad de género y los niveles de activismo.

La tesis aboga por interpretar las interacciones digitales como integradas en contextos socio-políticos más amplios, argumentando por respuestas estructurales y colectivas para abordar la violencia digital y para proteger a individuos y grupos vulnerables en espacios tanto *online* como *offline*. En conclusión, esta tesis avanza nuestra comprensión de la violencia machista en la era digital, abogando por enfoques interdisciplinarios para abordar este complejo problema.

Acknowledgements

Part of this doctoral research was conducted within the framework of the project *Analysing and combatting online hate speech and gender-based discrimination from an intersectional perspective (COME-ON)*, funded by the I+D+i Programme of the Ministry of Science and Innovation, 2020-2023.

In terms of the people who have made this research possible, first of all, I want to express my deepest appreciation to all the participants who have generously dedicated their time and shared their experiences, particularly through their active participation in interviews. You have greatly enriched the depth of this study, contributing with personal reflections on an oftentimes unsettling topic.

I would have wished to extend my gratitude to Ulf Björklund, but unfortunately, he is no longer with us. Ulf played a crucial role as my master's thesis supervisor at Stockholm University many years ago, imparting invaluable writing wisdom: "Keep it concise. Say what you mean and nothing else".

My warmest gratitude goes to Olga Jubany, my doctoral supervisor. I couldn't (and wouldn't) have done this without you. You encouraged me to start this adventure, and I've always felt encouraged to keep going after our tutorías. I also want to acknowledge the wonderful dedication that you have for your doctoral students, despite your extremely busy work life.

I also want to thank Camila del Mármol for advising me in the process of formal registration at the PhD programme, as the coordinator of the PhD programme at the Department of Social Anthropology.

Thank you to all my co-workers at the European Social Research Unit over the years and particularly Alèxia, Nacho and Emma for being such great colleagues over the past few years. Further, I want to thank all the people of GENI, our research group, for inspiring discussions, exchanges, and support.

Thank you, Lisa, for being my personal guru for so many years. Thanks to Jonas for always being there, although far away. I know you advised me not to do a PhD, but I did it anyway. Thank you to all my girls in "Gruppen" for cheering me on, you know who you are. Thank you, Jen, for all our conversations over the years on anything music-, feminist- or Grrrl-related. You're my one friend who might actually read this thesis. Thanks, Luna, for all the hours in the parks with the kids, talking about anything but academic stuff.

Ibra, thank you for not letting me give up! And thanks for sending me off to write alone that weekend, I really needed it to get back on track with this thesis. Thank you, Roy and Finn for all the love and hugs, and for always getting my mind focused on something far from academia. ¡Sois los mejores! I do think there's hope for equality when I see you.

Thank you, mom and dad, for always believing in me and supporting me. Tack mamma och pappa för allt!

Finally, I couldn't have done this without Sci-Hub. I'm forever grateful to Alexandra Elbakyan for acknowledging in such a practical and groundbreaking way that knowledge belongs to everyone.

1. Introduction

I enter Twitter² and see the following tweet by a young feminist:

They have created porn pages with photos from my Instagram, threatened to rape me, impersonated me, harassed me, organized hate campaigns and insulted me every day since I'm on social media. Against this the message is clear: not one step back. For every misogynistic insult I receive, that I see directed at a fellow feminist or at a stranger, it makes me want to put my face and body on display for them and for Feminism. You will be facing us. Neither your insults nor your harassment will make us retreat.³

This tweet delineates the central theme of this thesis, namely gendered online abuse, with a particular emphasis on instances of digital violence directed towards feminists. Over the following ten chapters, I aim to elucidate the lived experiences of the targeted individuals, explore the ramifications of such violence, and examine the responses articulated not only at the individual level but also as interconnected, collective responses.

In this opening chapter, I aspire to familiarise the reader with the subject matter and the underlying motivations that propelled this research, as well as the research questions and objectives underpinning this thesis. First, however, I shall

² Although this social media platform is currently known as X, throughout this thesis, the name Twitter will be used, as this is how the platform has been known during most of the research that underlies the thesis.

³ As described in Chapter 4 on methodological considerations, the quotes collected from social media platforms have been translated to English, and are not entirely literal, to avoid identification of the persons behind them through a search engine.

provide a brief delineation of the social and political context where the violence that has motivated this thesis has unfolded; concurrently contextualising the genesis and development of the thesis itself. The next sections, thus, focus on setting the scene through briefly depicting the emergence of an anti-gender movement in Spain as well as the tensions within the feminist movement.

1.1. Setting the scene: Feminism and anti-feminism

It has been widely argued that online abuse is a new manifestation of old problems (e.g., Lewis et al, 2019; Powell and Henry, 2017, Siapera, 2019). Thus gender-based violence cannot be perceived as merely a remnant of the past but, following Eugenia Siapera (2019), it should rather be seen as a factor that plays an active role in determining women's position in society. An increase in gender-based violence and hate crimes often emerges in the context of what is perceived as a loss of relative position in terms of gender equality. Challenges to the collective heteropatriarchal hegemony are often met with aggressive attempts to reassert the “natural” dominance of men and “assaults, rapes, or homicides that are outside the bounds of an intimate relationship tend to be directed at individual women as proxies for the combined threats to masculine domination represented by women as a class” (Perry, 2001:86).

In light of this, online gender-based attacks against women need to be interpreted within a framework of patterned resistance to women's public voice (e.g., Sobieraj, 2018). Advances in rights and opportunities gained by women “have often been counteracted by reactive movements, not necessarily explicit, which have become obstacles to further progress” (Menéndez-Menéndez et al 2021: 10). From this approach we understand that digital spaces have become new arenas for this violence, manifesting through verbal expressions or images with a collective message threatening all women and non-gender conforming individuals. Implicit in the surge of online misogyny is the overarching transition to neoliberalism and a production paradigm tied to technologies and cognitive labour, characterised by employment precarity and a dearth of social welfare provisions. Concurrently, there is a diminishing gender gap in education, enabling women to compete for the same employment opportunities with men (e.g.,

Siapera, 2019). This can be understood as part of the backdrop to the current waves of online gender hate and anti-feminism.

To comprehend resistance to feminist ideas, it is crucial to consider how misogynist discourses are spread online not only by individual "angry white men" (Pease, 2020), but how resistance to feminism is expressed and enacted also in an organised and networked manner. Furthermore, to comprehend the surge of online misogyny directed specifically at feminists in the present social and political landscape it is essential to consider the rise of an anti-gender movement and the ensuing proliferation of anti-gender discourses. In the next section, I shall briefly outline the emergence of this anti-gender movement in the case of Spain.

1.1.2. The anti-gender movement

In October 2021, Judith Butler writes in an article in *The Guardian*⁴:

The attacks on so-called “gender ideology” have grown in recent years throughout the world, dominating public debate stoked by electronic networks and backed by extensive rightwing Catholic and evangelical organizations. Although not always in accord, these groups concur that the traditional family is under attack, that children in the classroom are being indoctrinated to become homosexuals, and that “gender” is a dangerous, if not diabolical, ideology threatening to destroy families, local cultures, civilization, and even “man” himself.

Research on opposition to gender and sexual rights in Europe frequently amalgamates diverse phenomena such as populism, far-right parties, religious fundamentalism, nationalism, racism, and neoliberalism (Paternotte and Kuhar, 2018). Numerous patterns connect mobilizations in different countries. These include interactions between local organizers and the international movement; use of hyperbolic language; the exportation and translation of terms; and shared mobilisation patterns. They also encompass common symbolism, and similar triggers, including same-sex marriage, abortion, recognition of gender identity for transgender individuals, and adoption by same-sex couples. Additionally, there are connections to ultraconservative social actors. (Cornejo-Valle & Pichardo,

⁴ Judith Butler, “Why is the idea of ‘gender’ provoking backlash the world over?”, *The Guardian*, (October 23, 2021), available at, <https://www.theguardian.com/us-news/commentisfree/2021/oct/23/judith-butler-gender-ideology-backlash>, last accessed on April 19, 2024.

2017). Similarly, anti-gender campaigns and right-wing populist parties have several aspects in common. Across the European Union these parties share several sets of discourses, such as scepticism against the EU, resistance to globalisation, as well as national and “racial” concerns, with common elements such as looking for scapegoats, turning the aggressors into victims and the creation of conspiracy theories (Igareda, 2022). “Gender” or “anti-gender” becomes an interpretative framework that allows for the connection of diverse actors who, despite their differences, find common ground in confronting an alleged conspiratorial alliance with gender as its central axis. Nevertheless, disentangling the anti-gender movement⁵ is a complex endeavour that also needs to consider context specific factors (Paternotte and Kuhar, 2018), in terms of how the emergence of this movement has played out in Spain over the last decades.

In the Spanish case, the anti-gender movement can be understood not just as a political or conservative religious reaction to feminist ideas, but rather as a coherent and organised political movement (e.g., Igareda, 2022). Whilst these movements to a certain extent have their roots in the “anti-gender stance” of the Catholic church expressed already in the 1990s, as opposition to the broadening of sexual and reproductive rights formulated in the UN Conferences of 1994 and 1995 (Cornejo-Valle and Pichardo, 2017), Spain was amongst the first countries where this new wave of anti-gender campaign became visible. This development accelerated around 2004, when the Catholic church and conservative political groups mobilised against the initiative of the Zapatero Government to legalise same-sex marriages (Igareda, 2022; Paternotte and Kuhar, 2018; Cornejo-Valle and Pichardo, 2017; Aguilar Fernández, 2013). That is, in Spain, the concept of “gender ideology” “serves to unite diverse actors and struggles and has proven to

⁵ Following scholars such as Paternotte and Kuhar (2018) or Cornejo-Valle and Pichardo (2017), I refer to these mobilisations as “anti-gender” to insist on the specificity of this wave of opposition. These movements share a critique of gender, labelled as “gender ideology”, “gender theory” or “(anti)genderism”. The focus of these movements is on questioning the cultural construction of gender, reducing sexuality and gender roles to a presumed biological sex. “Gender” in this sense is addressed as the root cause of their worries and the core of the reforms they aim to oppose (Ibid.).

be a powerful social mobilizer, an effective trigger for the ‘moral panics’ of cultural conservatism” (Alabao, 2021: 400, *own translation*). These anti-gender discourses occasionally encompass conflicting stances on gender and women. They may concurrently incorporate traditional and conservative discourses emphasizing a perceived natural order, advocacy for sexual abstinence, or support for pro-life policies, juxtaposed with hypersexualised references to women. Despite these inherent contradictions, a common thread among these perspectives is the notion that women bear responsibility for the biological and identity reproduction of nations, and thus, need to be controlled (Igareda, 2022). Feminist discourses are perceived as a threat to this “natural” order, and to the “natural” roles of men and women. The anti-gender movement, then, argues for a return to the heteronormative nuclear family (Siapera, 2019; Igareda, 2022). In contrast to other countries, such as Poland or Hungary, where the anti-gender movement driven by the Catholic Church has managed to halt the advancement in sexual and reproductive rights of women and LGBT+ persons, in Spain, the anti-gender campaigns have not had any greater success in penetrating the broad public discourse or in effectively stopping laws related to gender or sexual rights (Cornejo-Valle and Pichardo, 2017). This can at least in part be attributed to the pronounced gap between the broad public and the Catholic Church. This “divorce” between the public and the Church can be traced back to the role of the Church in the ideological repression during the dictatorship, as argued by Monica Cornejo-Valle and José Ignacio Pichardo (*Ibid.*).

Online anti-gender campaigns diverge from their offline antecedents through the manifestation of heightened misogyny and the deployment of personalised, often sexualised assaults targeting individual women. In contrast to pre-Web 2.0 eras, where campaigns typically focused on issues like divorce or custody, and used traditional mediums such as demonstrations or manifestos, contemporary campaigns employ considerably more individualised approaches. Consequently, discerning attacks on feminism from those directed at specific women becomes challenging. This phenomenon underscores the dual nature of anti-gender or gender-based hate speech, manifesting both against individual women and

women collectively as a group (Igareda, 2022: 102). Previous research has located a “manosphere” within online spaces in Spain, expressed as “a current of networked misogyny connected to an international anti-feminist and anti-gender movement driven through organizations and political parties of extreme right-wing ideology in charge of spreading and amplifying this type of attacks” (Villar-Aguilés and Pecourt-García, 2021: 43). While the manosphere may include a variety of groups (Marwick and Caplan, 2018), these groups tend to share a belief that feminine values dominate society, that this fact is suppressed by feminists and “political correctness,” and that men must fight back to protect their very existence (Marwick and Lewis, 2017). In this regard, “setting up feminism—and feminists—as villains, and men as victims, justifies the networked harassment that often emerges from the manosphere” (Marwick and Caplan, 2018: 5).

Having provided an overview of the anti-gender movement, it is also relevant to briefly examine how the internal tensions within the feminist movement have unfolded in Spain in recent years. This is with particular reference to the emergence of a split between trans inclusive feminists and groups that aim to exclude trans persons and transfeminism from the movement.⁶

1.1.1. Trans inclusive feminism versus trans exclusionary feminism

To understand the complex background to the topic of gender-based online hate speech in Spain, several entangled strands of tensions need to be considered. This includes those within the feminism movement itself, with the growing antagonism between trans inclusive feminists and persons who auto-define as “radical feminists”⁷, many of whom express transphobic discourses. The feminist agenda in Spain has always been diverse and, on occasions, has witnessed

⁶ Although I address this development separately, it should be kept in mind that there are many common elements between the anti-gender movement and the transexclusionary sectors of the feminist movement.

⁷ Whilst these persons often refer to themselves as “radical feminists”, their antagonists tend to refer to them as “TERF”, “Trans Exclusive Radical Feminist”. It has been acknowledged that the term TERF was first used by Australian blogger TigTog in 2008, as a neutral term to describe feminists who do not include trans women within their feminism. See Viv Smythe (TigTog), “I’m credited with having coined the word ‘Terf’: Here’s how it happened”, *The Guardian* (November 28, 2018), available at <https://www.theguardian.com/commentisfree/2018/nov/29/im-credited-with-having-coined-the-acronym-terf-heres-how-it-happened>, last accessed on April 19, 2024.

divergences between theoretical and activist stances marked by distinct priorities (e.g., Platero, 2020; Arce-García & Menéndez-Menéndez, 2023). Since 1993, trans women have actively engaged in Feminist State Conferences, contributing to discussions on the essence of womanhood, encounters with prostitution, and feminist collaborations, among other subjects (Platero and Ortega, 2016). The term "transfeminism" made its first appearance during the 2000 Feminist State Conference, asserting the feminist identity of trans women whilst aligning with global transfeminist dialogues (Willem et al, 2022: 186). This support of transfeminism within the broader feminist movement continued throughout the first decade of the 21st century. The inclusion of trans persons became consolidated in the National Feminist Congress that took place in Granada in 2009, as highlighted by Lucas Platero (2020: 35):

One characteristic of this plural third-wave feminist movement, regarding the question about the political subject, is that it not only considers the plurality of women's embodiment but also the participation of men. Topics such as feminine masculinity, trans men in feminism, or dissident sexuality of women were very present in the discussions.

Nevertheless, less than a decade later, this was no longer the case. Scholars situate the start of an increasing antagonism within the feminist movement in Spain to around 2018, with a decisive rupture following a discourse by trans feminist Sam Fernández at the *Escuela de Otoño*⁸ in October 2018, where Fernández argued for an expansion of the political subject of feminism (Platero, 2020; Romero, 2020). This discourse was filmed and became viral, leading to an array of discourses and attacks against trans feminism, from within the feminist movement. This rupture signified a departure from the previous integration of trans women within the feminist movement in Spain (Platero, 2020).

If the split had been decisive already following the 2018 event, it was even more fueled following the *XVI Escuela Feminista Rosario Acuña*⁹ in July 2019. During this event, several second-wave feminist academics, politicians, and activists

⁸ The "Fall School" is an event organised by the political party Podemos.

⁹ The Rosario Acuña Feminist School is an annual event organised by the Equality Policies Service of Gijón, Asturias.

derided a feminist discussion on trans rights by making disparaging comments about trans women (Willem et al, 2022). These “transphobic and anti-queer interventions” (Romero, 2020: 16, *own translation*) were quickly viralised on social media. The discourse employed at the Rosario Acuña event closely aligns with the rhetoric of the Spanish far-right party Vox and other ultra-conservative organizations. This unexpected convergence forms an alliance between feminists advocating trans exclusion and far-right entities. On social media platforms, resistance to these messages was expressed through the hashtag #hastaelcoñodetransfobia – “we’re fed up with transphobia” (Willem et al, 2022).

Further, in 2020 several transphobic manifestos were published by feminist collectives. The most impactful was a manifesto released by feminists active in the socialist party, which argued that the new Trans Law “would negatively impact women’s rights and safety” (Ibid.). Another transphobic manifesto was released by Confluencia Movimiento Feminista¹⁰ the same year, stating explicitly that it was “in favour of the feminist agenda, against trans laws” (Ibid.). With these transphobic manifestos, a significant anti-trans campaign was consolidated, engaging in an exercise of appropriating the term “women” in an exclusive manner (Romero, 2020). Many of these oppositions and tensions currently unfold on social media.

However, concurrently and somewhat paradoxically, both feminism and trans activism have been gaining influence in Spain. Notably, prominent voices are endorsing the rights of trans women within the broader feminist movement (Willem et al, 2022). The impact of feminist strikes on March 8th in 2018 and 2019, as well as public demonstrations addressing rape and gender-based violence, along with hashtag campaigns following the MeToo movement of 2017, highlights the growing influence of the feminist movement. This is notable despite

¹⁰ This organization defines itself as a network of feminist organizations that has joined forces “to achieve greater political strength for the advancement of the feminist agenda at this moment”, with an explicit focus “to prevent normative initiatives that seek to enshrine the concept of ‘gender self-identification’ in the Spanish legal system” (own translation, from <https://movimientofeminista.org/quienes-somos/> last accessed on April 19, 2024).

the internal tensions (Romero, 2020). This expansion of feminism has called for the recognition of various situations and voices. For example, feminist groups from Roma, Afro-descendant, racialised, and migrant communities have been increasingly visible in debates in recent years, offering a critical perspective on colonialism and racism, whilst also providing critique of the historically entrenched white ignorance sustained by privilege, including within feminism (Ortega, 2019; Romero, 2020).

It is at the core of the debate put forward that this thesis arises. The subsequent section will delineate my personal entry into this scholarly domain, and the initial questions that propelled the research.

1.2. Origins of the research and initial questions

Between the years 2014 and 2019, I was involved as a researcher in two cross European research and action projects funded by the Rights, Equality and Citizenship Programme of the European Union, *PRISM: Preventing, Redressing and Inhibiting Hate Speech in New Media (2014-2016)* and *REACT: Respect and Equality: Acting and Communicating Together (2017-2019)*. These projects addressed online hate speech on the grounds of racism and xenophobia. My role involved interviewing young people in relation to their experiences of online hate speech, either as direct targets or as bystanders, as well as exploring counternarrative initiatives in relation to racist and xenophobic hate speech. When speaking with girls and women, gendered aspects of hate speech tended to come across in their narratives. Many women and girls who had been targeted by racist hate speech online felt that their experiences differed from those of boys or men. The abuse was distinct, it was experienced differently, and discerning racism from misogyny was not always straightforward. It was clear, on the one hand, that girls and women were targeted to a higher extent, by gender-based and racist, xenophobic or Islamophobic abuse, and, on the other hand, that an intersectional perspective was needed to understand and address this violence. Further, not only did these projects open my eyes to this form of intersectional

violence, but they also initiated my path as a researcher within the field of hate studies, conceptually framing this abuse as *hate speech*.

In parallel to this, I observed the continuous increase in tension on social media between, on the one hand, trans inclusive feminists and transphobic persons, and on the other hand, continuous attacks coming from the increasingly active anti-gender movement. Considering the social and political context outlined in previous sections, feminists seemed to be under attack from several different angles. I already followed some of the women targeted by what seemed like a growing anti-feminist current on Twitter, and observing them whilst they dealt with waves of hate. This led me to wonder: How does encountering unwarranted hatred affect someone who tries to share a message, especially when they are unfairly labelled as unworthy of being heard?

At the onset of my research, there was a notable lack of scholarly exploration into the realm of gender-based online hate speech in Spain. Consequently, I opted to undertake a comprehensive examination of the issue, seeking to address several key questions: How do women navigate the experience of being targeted by gendered online hate? What are the impacts and consequences of these experiences? How do women devise strategies to shield themselves and/or engage in resistance? And, what role do digital technologies play in shaping this landscape? At a broader, theoretical level, my hope was to shed further light on experiences and conceptualisations of gendered violence in digital spaces.

I chose to focus specifically on online violence directed towards feminists and/or other activists (anti-racist and/or LGBT+) because, as I will argue in this thesis, it represents a distinct form of violence. This type of violence not only targets these women because they are (racialised, LGBT+) women, but also because of the opinions they articulate and the online spaces of resistance they are working to establish, advocating for their own rights or the rights of others.

1.3. A note on the vocabulary used in this thesis

I have chosen to use interchangeably the concepts “digital” and “online”, as opposed to “virtual” when referring to the technological context. This is because

the notion of “virtual”, as argued by Niels van Doorn (2011), tends to evoke associations with “virtual reality” or “cyberspace” as experientially distinct domains in contrast with the “real” world. As such, the term “virtuality” has often been used to represent either the opposite to or a lack of “reality”: a state of unreality or absence. Regarding the concept “digital”, I follow anthropologists Daniel Miller and Heather Horst’s definition (2012: 3), where the digital is defined as “all that which can be ultimately reduced to binary code but which produces a further proliferation of particularity and difference”. That is, the digital from this view can be understood as intensifying the dialectical nature of culture. The fundamental definition of digital binary offers the advantage of discerning a potential historical parallel; specifically, the impacts of the digital era can be analogised to the repercussions of the introduction of money in human history, signifying a novel stage in human abstraction (Miller and Horst, 2012). In this context, the evolution of binary code significantly streamlined information and communication, paving the way for increased convergence between technologies or content (Ibid). Nevertheless, humanity is not necessarily more mediated by the rise of the digital. Rather, as suggested by Miller and Horst (Ibid.: 4),

anthropology will progress to the degree that the digital enables us to understand and exposes the framed nature of analogue or predigital life as culture and fails when we fall victim to a broader and romanticized discourse that presupposes a greater authenticity or reality to the predigital.

Further, when discussing the experiences of the persons who are subjected to gendered online abuse, I have chosen to denominate these “targets”, instead of “victims”, “survivors” or similar concepts. This is to highlight the emic, lived perspective given that the women whom I have interviewed would not refer to themselves as “victims”. Neither do these women necessarily feel or define themselves as “victimised”, in part due to the very normalisation of the phenomenon of gendered abuse online. Most of these women are subjected to a continuous stream of abuse, much of which does not necessarily reach the threshold of a criminal offense. The term “target” speaks of a dynamic and continuous process (as opposed to a one-off or static event), which to a higher

degree than the term victim resonates with the experiences of being the objective of hate speech. Through using the term “targets” I follow the path of scholars such as Laura Beth Nielsen (2002), Danielle Keats Citron (2011), Emma Jane (2014; 2016), or Katharine Gelber and Luke McNamara (2015). Nevertheless, in the sections based on the data from interviews with professionals, in the specific context of support provision, the concept “victim” is used to a higher extent. This is based on the professionals’ own utilisation of this term to refer to the people they provide support to. In this context, refraining from labelling oneself or the people seeking support “victim” may pose an obstacle to accessing institutional assistance (Jägervi, 2014).

Finally, whilst I refrain from using entirely verbatim quotes from the online material to safeguard the anonymity of the women involved in the digital ethnography⁹, I still strive to adhere closely to the original expressions. This means that I have chosen not to censor sexually explicit or dehumanising language, following a deliberate strategy “to speak of the ostensibly unspeakable so as not to perpetrate—and thus perpetuate—the tyranny of silence about the sexually explicit nature of this material” (Jane, 2014).

1.4. At the intersection of hate crimes and gender-based violence

This thesis adopts a critical anthropological perspective on the lived experiences of violence in digital media, in the case of gender-based online hate speech in Spain. Simultaneously, the thesis contributes to the realm of hate studies, an inherently cross-disciplinary and international field (Schweppe and Perry, 2022). Furthermore, it also contributes to the conceptualisation of misogynistic hate speech by positioning this within the broader spectrum of hate incidents and hate crimes.

A significant part of the initial literature on digital media and the internet emphasized disruption and transformation, viewing technologies as paving the way for a "historically new reality" (Castells, 1996: 92). In the early days of the internet, there was an optimistic belief that online interaction would be liberated from oppressive structures, resulting in a gender-neutral environment. The web

was hailed as a potential space where one's identity or background could be bypassed and rendered irrelevant (e.g., van Zoonen, 2002). However, there were also early attempts to challenge this view of the digital as a new and wholly different realm, such as those by anthropologists Daniel Miller and Don Slater (2000). Miller and Slater noted that digital media must be analysed as “continuous with and embedded in other social spaces, that they happen within mundane social structures and relations that they may transform but that they cannot escape into a self-enclosed cyberian apartness” (2000: 4). In other words, this perspective underscores the importance of examining how digital technologies frequently support social reproduction rather than promote change.

Whilst online violence in the past has been debated both by scholars and media as gender-neutral, data from a range of sources shows that women and girls are the main targets of such digital violence, reflecting the violence against women that is pervasive in all spheres of life (e.g., European Women’s Lobby, 2017). Digital platforms provide spaces for hate speech, with attacks reaching targets anytime and anywhere, through mobile internet access (European Women’s Lobby, 2017). This online violence can be understood as having the function of silencing and scaring women from full participation in digital spaces (Jane, 2017); spaces that in our contemporary world make up an important part of the public sphere as platforms for political debate. That is, whilst online misogyny needs to be linked to the social structures surrounding it, it is also necessary to add the prefix “online” or “digital” to “hate speech”, to mark the internet and social media as new public spaces where women are vulnerable to systematic subordination (Megarry, 2014). At the collective level online, women as a group are targeted by different forms of verbal technology-facilitated violence: ridiculing gender theories, insulting or parodying women as a social group, trivialising or denying gender violence (Igareda, 2019). In the same way that a hate crime does, online verbal abuse has an impact not only on the individual woman but sends a message to all women that they need to watch out, that they shall not raise their voices, and that their freedom is restricted.

Gender-based online hate speech can be situated at the intersection between gender-based violence and hate crimes; however, in practical terms, this discourse – as I shall later discuss in-depth – is rarely treated as gender-based violence nor does it tend to be conceptualised within the hate crime framework. My positioning, as I will argue, is to conceptualise this violence as a verbal form of hate crime, that is, hate speech: This is within the specific context of Spain – where gender, sexual orientation and sexual identity are considered as aggravating circumstances of crime that negate the principle of equality. Hate speech is defined in EU law as the public incitement to violence or hatred on the basis of certain characteristics, including ‘race’, colour, religion, descent and national or ethnic origin.¹¹ While the EU Framework Decision that provides this definition explicitly covers only racist and xenophobic speech, there is wide agreement among different institutions that the list should be read as open-ended. In accordance, many Member States – including Spain – have extended their national laws to include other grounds such as sexual orientation, gender identity and disability. The European Commission against Racism and Intolerance (ECRI) also explicitly refers to sex, gender and gender identity in its General Policy Recommendation Nº 15 on combating hate speech. The Recommendation further addresses the role of intersectionality, stating that ECRI is “conscious of the particular problem and gravity of hate speech targeting women both on account of their sex, gender and/or gender identity and when this is coupled with one or more of their other characteristics” (ECRI, 2016).

Online gender-based hate speech can be interpreted as an old discourse reinforced by new technologies; a discourse shaped by the intersection of technological affordances with underlying, structural oppressions. To fully grasp the continuity and the lived experiences of oppressing structures, this must be explored along the interlacing continuum online-offline. This thesis argues that the “online” and the “offline” are not separate experiential realms; but rather that activities that take place in digital spaces are experienced as reality, with

¹¹ Framework Decision on combating certain forms and expressions of racism and xenophobia by means of criminal law, available at <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32008F0913>, last accessed on April 19, 2024.

material consequences (Lewis, 2016). This means that, in the ethnographic sense of the word, the field must also be defined along the continuum of the online and the offline, with ethnography carried out across different realms.

I further argue that the anthropological perspective, including the anthropological tradition of ethnography, is a crucial approach for a comprehensive analysis of the phenomenon of online gender-based hate speech. A grounded anthropological perspective on hate speech would start with victims' real-life experiences, as proposed by Sindre Bangstad (2017). This approach aims to offer distinctive insights into the functioning and consequences of hate speech on those it targets, as well as on the broader community to which they are somehow affiliated. Considering the viewpoints of individuals targeted by hate speech enhances the affirmation of their agency and ability to respond to or contest abuse. Simultaneously, this challenges the conceptual and ideological contexts that form the basis of online hate speech. In this regard, "the contribution of anthropology is to disturb comfortable understandings of the world by showing the simultaneous plausibility and arbitrariness of multiple ways of understanding and living in it", as argued by Hugh Gusterson (1996: 1).

Still, very little anthropological attention has been paid to online hostility, and particularly to online hate speech. As scholars such as Emma Jane (2014) point out, until recently, gender-based online hate speech has received scarce attention from scholars in general. As Jane suggests, part of the reason that scholars avoid this topic is that it is "heavily laced with expletives, profanity and explicit imagery of sexual violence; it is calculated to offend, it is often difficult and disturbing to read" (Jane, 2014: 558). However, perspectives such as that of Jane stem from placing a considerable focus on a textual approach, analysing the discourse, and not so much the experience from the point of view of those who are targeted by hate speech in their daily lives. Whilst a textual approach is valuable, without delving into the lived experiences of targets, we cannot reveal how these experiences intersect with other facets of life and identity. Furthermore, it is only through the anthropological lens, that we are able to analyse the integration — or lack thereof — of such experiences into daily

activities and political engagement, spanning both digital and non-digital domains (Lewis et al, 2016).

1.5. Objectives of the research

As argued so far, it is fundamental to address the analysis of online hate speech from an anthropological and intersectional approach, in order to explore and expose the daily, diverse experiences of the people who are targeted by it. This implies exploring the multiplicity of oppressions to which they are subjected, from a perspective that makes relevant not only the oppressions of gender reflected in the hate speech, but also the imbrication of these with other grounds of discrimination such as racialisation, sexual orientation, gender identity or religion.

As pointed out in previous sections, despite the generalisation of the phenomenon of hate speech based on gender on the internet, little research has focused on this topic from an ethnographic perspective. To address this important gap, this thesis seeks to analyse online gendered hostility expressed as online gender-based hate speech. This pursues a deeper understanding of the underlying mechanisms of gender-based hate speech, by analysing the structural relationships of domination, discrimination, power and control manifested in and underpinning this discourse.

Within this research exercise, a crucial aspect involves unravelling the intersection of gender with other categories in online hate speech, by delving into the embodied experiences of individuals subjected to gender-based online hate speech. To this aim this thesis aims to examine the real-life impacts and consequences of being a target of gendered online violence, both at an individual and collective level. In connection with this, this research scrutinises the broader societal implications concerning which voices can engage in public discourse, online and beyond.

Another crucial objective related to the lived experiences of online violence is to investigate the contestations and responses to this violence by individuals, collectives, and institutions. To conduct a comprehensive analysis of responses to this phenomenon, this thesis further seeks to analyse the construction of the legal

framework regarding gender-based online hate speech and its correlation with responses such as reporting to the police, along with the experiences of the victims. To provide further insights and contextualise targets' efforts in seeking institutional support, another objective has been to analyse the perceptions and perspectives of professionals in the fields of gender-based and anti-LGBT violence, including their readiness to offer support to those targeted, particularly concerning the online dimension.

Finally, at a theoretical and conceptual level, this doctoral research explores the relevance of analysing this abuse within a hate crime framework, aiming to contribute in a significant way to the field of hate studies, particularly in the context of gender and intersectionality.

1.6. Explanation of the thesis structure

Following this introduction, the theoretical framework for the thesis is laid out in Chapter 2: "Constructing a conceptual framework: Gender, technopolitics and violence". This is organised around three primary pillars: "Identity and gender constructions in a globalised world," "Technopolitics and digital social inequalities," and "Conceptualisations of violence." The first pillar delves into the concept of identity within the context of globalisation processes, narrowing in on normative gender constructions and intersectionality as crucial dimensions for comprehending the underlying structures of gender-based online hate speech. The second pillar concentrates on the digital realm, exploring its interconnection with the non-digital sphere and highlighting the multiple layers of interconnected and embedded inequalities within and across these spaces. Finally, the third pillar examines various conceptualisations of violence that gain significance when analysing gender-based online hate speech. These encompass theorisations surrounding gender-based violence, hate crimes, and speech that causes harm. The latter is explored through the concept of linguistic vulnerability (Butler, 1997), which covers particularly well the harm of being targeted by verbal attacks online. Altogether, these pillars make up what I refer to as an intersectional technofeminist framework.

Building on these debates, Chapter 3: “From definitions to debates: A scholarly framework for gendered online hate speech” addresses previous scholarly debates relevant to the topic of this thesis. The chapter explores the multifaceted landscape of online violence, particularly focusing on gender-based hate speech, drawing insights from diverse academic disciplines. It clarifies various conceptualisations of online hostility, including gendered online hostility. It introduces key debates surrounding violent behaviour in digital spaces, whilst exploring alternative concepts, such as "extreme speech", primarily originating from social anthropologists. Through this overview, I aim to highlight the challenges posed by the diverse array of concepts used to explore online hostility across disciplines, and argue for the relevance of the hate speech concept in relation to the topic of study. The chapter then provides an overview of studies related to gender-based online hate speech, often synthesised from broader categories such as "technology-facilitated violence" or "cyberhate", with a focus on findings relevant to the present research. Informed by this literature review, the chapter proposes a working definition of gender-based online hate speech for the doctoral thesis, whilst subsequently offering a rationale for the research focus.

Following this, Chapter 4: “A continuum of online – offline ethnography: Methodological and ethical considerations” introduces the methodology applied in the research, as well as the ethical considerations linked to this. The chapter first explores reflexivity and the application of a feminist intersectional approach. Following this, the chapter explores the main methodological tools utilised. These draw upon anthropological research traditions, emphasising in-depth interviews, along with digital ethnography as a form of participant observation conducted in digital spaces. I thus adopt a "postdigital" approach to digital ethnography, which transcends traditional online/offline divides, offering a critical lens to investigate the situatedness of digital media in everyday life. The final section of the chapter explores ethical implications and guidelines, considering aspects such as informed consent, anonymity, as well as specific ethical considerations related to research in online contexts.

At this point the thesis shifts focus from conceptual and methodological discussion to the analysis and discussion of the empirical findings., Chapter 5: “The legal framework: Overview, interpretations and gaps” starts with a general discussion of the fundamental rights at stake when addressing online hate speech with the help of legal tools, namely the right to freedom of expression and the right to non-discrimination and human dignity. It then presents a descriptive overview of the legal framework addressing gender-based online hate speech in Spain. Following this, the chapter further analyses this legal framework, introducing also empirical findings from the interviews with legal experts, highlighting the fragmented nature of the legal framework whilst revealing gaps and differing interpretations within the legal landscape. While not striving for an exhaustive legal analysis, the chapter provides an overview of existing legal responses to gender-based online violence.

This sets the stage for the subsequent Chapter 6: “Professionals’ perceptions of online gender-based violence and their preparation to provide support to targets”, which delves into the perspectives of professionals on gendered online violence. The analysis is based on ethnographic interviews with professionals who provide support to victims of gender-based violence and anti-LGBT violence from different perspectives, as well as professionals from the justice system and law enforcement agencies. The chapter focuses on these professionals’ perceptions of the online context, their categorisations of this type of gendered online violence, and their level of preparedness to respond to this violence and provide support to targets. This analysis ends with an exploration of existing training needs and possible ways forward to improve the support network provided by institutions and civil society organizations.

Subsequent to this exposition, the focus is turned to the lived experiences of those who are targeted by online violence, delineated across three distinct chapters. These chapters are based on interviews with targets and on the digital ethnography conducted, in a way that the findings and analysis overlap and intersect organically. Nevertheless, for analytical clarity, these domains are compartmentalised as follows: Chapter 7 predominantly analyses and chronicles

events from a situated standpoint of lived experiences, Chapter 8 explores the impact of these occurrences both at an individual and a collective level, and Chapter 9 delves into the responses, reactions, and contestations to this violence and the resultant harm.

Chapter 7: “Targeted voices: Experiencing gendered online violence”, delves into the online contexts of the targets, exploring their diverse uses of digital spaces. Subsequently, it examines their experiences within a "continuum of violence". The focus initially rests on experiences tied directly to the targets' feminism and activism from a gender perspective, considering not only gender, but also intersecting identity traits such as sexual orientation, racialisation, and gender identity. The exploration then extends to encompass various episodes of gendered online and offline violence in the narratives, signalling a continuum of violence that both reflects and perpetuates heteropatriarchal structures. Concluding sections scrutinise the women's perceptions of the perpetrators, followed by a reflection on the normalisation of online violence in the lives of feminists.

Moving forward, Chapter 8: “Digital harm, embodied impact: Unmasking the impact of gender-based digital harm” explores the impacts of gendered online violence on its specific targets and beyond. The focus extends beyond explicit threats or violence to encompass more subtle verbal abuse, highlighting how repeated instances in a public setting can significantly affect targets. The effects of online abuse are highlighted as multi-dimensional, influencing various aspects of targets' lives, including emotional, physical, and psychological well-being. The analysis highlights how technology-facilitated violence leads to tangible harm for real victims. Finally, the chapter examines the broader consequences at both individual and collective levels, including self-censorship, internalisation of messages, and the potential “chilling effect” on a collective scale.

The final empirical Chapter 9: “Navigating online gender-based hate speech: Responses and contestations” delves into resistance to online hate, which takes various forms, encompassing both individual and collective responses. The spectrum of responses is wide-ranging: some focus on minimising the immediate

impact of the violence to shield targets from further abuse, while others have broader goals, such as raising awareness or fostering a sense of community for sharing experiences and offering support among peers. Together with Chapter 8, this chapter provides evidence to how women facing gendered hate online navigate the digital sphere with an acute awareness of potential threats, incorporating a level of vigilance into their everyday lives. The perceived impunity of digital violence places a disproportionate burden on targets for self-protection, whilst the lack of institutional or platform support, may lead targets to lean towards strategies of seeking justice in more informal ways.

The concluding chapter of the thesis, Chapter 10: “Recapitulation and concluding observations” intends to intricately intertwine the thesis and its analysis. This final reflection aims to underscore not only the empirical conclusions but also the theoretical contributions arising from this doctoral research, thereby offering a comprehensive synthesis of the thesis’ findings and insights.

2. Constructing a Conceptual Framework: Gender, Technopolitics and Violence

2.1 Introduction

Theory is essential to interpretation and knowledge production and there is thus a “necessity of bringing theory to the field” (Burawoy, 2003). The present chapter aims to introduce the theoretical framework of this thesis, constructed around three main pillars: “Identity and gender constructions in a globalised world”, “Technopolitics and digital social inequalities” and “Conceptualisations of violence”. The first pillar introduces the concept of identity in relation to globalisation processes, before narrowing the focus on normative gender constructions and intersectionality, as critical dimensions to understanding the structures that underlie gender-based online hate speech. Afterward, the second conceptual cornerstone directs attention to the digital realm, exploring its interconnection with the non-digital sphere and highlighting the various layers of interconnected and embedded inequalities within and across these spaces. Lastly, the third pillar delves into distinct conceptualisations of violence, especially relevant in the analysis of gender-based online hate speech. These encompass gender-based violence, hate crimes, and linguistic injury. The latter explores diverse perspectives on how language shapes and perpetuates social structures, acknowledging the potential of speech to inflict harm.

2.1. Identity and gender constructions in a globalised world

2.1.1 Globalisation and identities

Globalisation can be defined as the manifold ways in which the world is increasingly interconnected (Hylland Eriksen, 2007). This interconnection has contributed to a shift in traditional ideals of identity as being naturally bound by place of birth, economic status, or gender (Bauman, 2000). The evolving reality brought about by globalizing processes, labeled by Zygmunt Bauman as "liquid modernity," is characterised by adaptability and lack of strict boundaries (Van Meijl, 2008). This encourages individuals more than before to actively shape their own identity traits, albeit within the confines of their social standing, influenced by factors such as class, racialisation, origin, and gender. In this liquid society, identity is seen more as a task, a performance—something transient, immediate, and dynamic—rather than a permanent fixture (Bauman, 2004).

The emergence of new technologies has been essential to globalisation over the last half century. In the so-called "age of techno-personal systems" (Gergen, 1991), technologies influence our relationships and interactions with other human beings. Nevertheless, being interconnected by technologies is not fully the same experience as being physically close, as "the experience afforded by these connections coexists with an undeniable, stubbornly enduring *physical distance* between places and people in the world, which the technological and social transformations of globalization have not conjured away" (Tomlinson, 1999: 4). Whilst new technologies can connect us over distance, that is, we can talk to people that we might never meet face-to-face, we are still physically separated, thus, the nature of interactions is still somewhat different. The flexibility of the liquid modernity makes people interact as "swarms" rather than groups as "they come together, scatter and gather again, from an occasion to another, everytime inevitably for a different reason, and are attracted to changeable aims" (Palese, 2013). Here the notion of "network" as characterising human interactions comes into play. As Bauman (2003) points out, "network" stands for a matrix and, at the same time, for connecting and disconnecting; these being equally legitimate choices, which enjoy the same status. "'Network' suggests moments of 'being in

touch' interspersed with periods of free roaming" (Bauman, 2003: xii). In a network, "connections", or "virtual relations", are entered on demand, and can be broken at will (Ibid.). This can be applied to how previously non-connected strangers come together in digital spaces, such as social media, connected by similar ideas and causes, or forming swarms prompted as reactions to ideas expressed by other people. Linked to this, Kenneth Gergen (1991: 180) refers to "technologies of social saturation" in what resembles an eerie prediction of contemporary social media use: "As the technology of saturation becomes more effective, the range of fractional relationships further dissipates family functions (...) Virtual strangers provide a personal service, often an emotionally significant one, matched to a particular facet of one's being". New technological advancements are critical as these create what Gergen refers to as "Ersatz Being", describing thus the capacity to jump from one identity to another, or one relationship to another, facilitated by the postmodern context of technological advancements. Related to this perceived flexibility and performativity in technological interactions, Erving Goffman's theory (1959) on the presentation of self easily comes to mind. As Goffman puts forward, people engage in different presentations of self to different audiences in different arenas of everyday life. Different presentations of self can be used to a higher extent online than in offline spaces; but different versions of the self can also be highlighted in different digital spaces; for example, the professional self, performed on the networking platform LinkedIn can be compared to the different versions of the self presented at social media platforms such as Instagram, or anonymous and less inhibited versions expressed in other types of online forums, such as Reddit or ForoCoches.

Whilst keeping in mind that both Goffman and Gergen document pre-internet examples of this identity performance, the performative nature of identity online seems almost unavoidable, as online participation enables the creation of multiple personae (Kendall, 2002). It can be argued that this new technological reality prepares us for a variety of roles and presents us with opportunities for performing identities to a much higher extent than previously. However, as I will argue further in section 2.2, this must be understood within the limits of social

structures that mutually interact with technologies. This is particularly relevant in the current context of "radical transparency" (Kirkpatrick, 2011) on social media platforms, which promotes the use of one's "authentic" identity online, encompassing the disclosure of gender, the use of photos, and real names. Such a trend fosters a growing integration of our online and offline existences, underscoring the crucial influence of our physical bodies in shaping our digital personas.

First, however, in the next section I will discuss briefly identity in terms of normative gender constructions.

2.1.2. Normative constructions of gender

As early as in the 1930s, anthropologists such as Margaret Mead (1935) showed how gender systems are social constructs, arguing that

any society that specializes its personality types by sex, which insists that any trait - love for children, interest in art, bravery in the face of danger, garrulity, lack of interest in personal relations, passiveness in sex-relations; there are hundreds of traits of very different kinds that have been so specialized - is inalienably bound of with sex, paves the way for a kind of maladjustment of a worse order (Mead 1935: 293).

Later on, Gayle Rubin highlighted that "every society has a sex/gender system - a set of arrangements by which the biological raw material of human sex and procreation is shaped by human, social intervention and satisfied in a conventional manner" (1975: 165). As Rubin points out, "sex/gender system" is a neutral concept - as opposed to concepts signalling a stratification of gender. This anthropological concept, thus, facilitates an understanding of oppression in relation to gender as not inevitable, but rather as the product of the social relations that organise it. However, as we shall see in the ethnography of this thesis, challenging the social relations of this system often comes with a price derived from the decidedly normative layer of gender constructions. As Rita Segato (2003) puts forward, patriarchy is at the same time a norm and a project of self-reproduction and, as such, to understand its complexity, as anthropologists, we need to be sensitive to power relations and their oftentimes immensely subtle discursive expression.

Gender can be understood as concerned with “managing situated conduct” (Perry, 2001: 81), according to society’s normative expectations of what constitutes essential “maleness” and “femaleness”. The cultural discourse on gender restricts, limits, and frames practices (Segato, 2003). Within the essentially dichotomous understanding of gender in our society, there is very little space for ambiguity, for blurring the line between masculinity and femininity, or for gender identities beyond these binary constructions. On the contrary, “a person engaged in virtually any activity may be held accountable for performance of that activity as a woman or a man, and their incumbency in one or the other sex category can be used to legitimate or discredit their other activities” (West and Zimmerman, 1987:136). Thus, if we perform gender “appropriately”, we simultaneously sustain, reproduce and render legitimate the institutional arrangements that are based on categories of sex (Ibid.: 146). However, if we fail to perform gender appropriately, we as individuals— rather than the institutional arrangements—may be asked to account for our character, motives, or predispositions. At the very core of this doctoral thesis are the responses and reactions to the questioning of normative expectations and constructions of gender.

As one of the pioneering post-structuralists, Judith Butler’s (1990), approach to the performativity of gender is particularly relevant. According to this theorisation, gender categories are not static or “natural”, but rather constitute a performative repetition of acts associated with the “male” or “female”. Through one’s body, one performs gender through repeated acts that are constantly being “renewed, revised and consolidated” (Butler, 1988: 526). The actions appropriate for women and men in different cultural contexts are transmitted to produce a social atmosphere that both maintains and legitimises a seemingly natural binary gender order, i.e., from a poststructuralist perspective, “the very subject of women is no longer understood in stable or abiding terms” (Butler 1990: 1). Following on from this, as Butler points out, the categories “women” and “men” are produced by the same structures of power through which emancipation is sought. The term “women” may denote a common homogeneous identity, which

is problematic, particularly from a feminist perspective (Ibid.). As I have highlighted in Chapter 1, the stability or fluidity of gender are core issues currently disputed and contested within the different polarised strands of the feminist movement. This debate is closely linked to and addressed by the concept of intersectionality, which problematises women as a homogeneous category. This is further discussed in the next section.

2.1.3. The intersectional lens

Intersectionality can be applied as both a methodological tool and a theoretical concept, which refers to the interaction between gender, racialisation and other categories of difference in individual lives, social practices, institutional arrangements, and cultural ideologies, *and* the outcomes of these interactions in terms of power (Davis, 2008). In other words, the concept of intersectionality can be understood as a tool to understand how social hierarchy creates the experiences that produce the categories that intersect (MacKinnon, 2013).

Intersectionality as a theory emerged from raising discussions about the impact of experiences of 'race'¹², and gender merged in the lives of Black women in the United States, led by Kimberlé Crenshaw (1989). Her analysis showed how the outcomes of the intersection of these factors "is greater than the sum of racism and sexism" (Crenshaw 1989: 140). Intersectionality, thus, does not merely introduce additional variables but instead adopts a specific perspective, embodying a dynamic approach to comprehending the fundamental dynamics of the reality it aims to understand. However, at the same time it remains grounded in the experiences of people within hierarchical relations (MacKinnon, 2013). With

¹² In relation to this, it should be highlighted that "race" is a highly contested concept, and it will not be used as such in this thesis. Critiques of the fixity of the concept of "race" have led to other concepts that highlight "race" as a socially constructed category, rather than something biologically fixed and objective. In this regard, the term "racialisation" has been taken up by many scholars (see e.g., Miles and Brown, 2003; Murji and Solomos, 2005), referring to the ways in which racial ideas are constructed to define differentiated social collectives and applied as the basis of exclusionary practices. The thesis refers to racialised identities as the bio-racialisation that persists in the policies, practices and social discourses that people experience individually and collectively. Racialisation is a construction with historical genesis in the synthesis of multiple discriminations and reductions, determining social positions, social prestige, privileged relationships, and diffusing the systems of representation, meanings, behaviours and subjective practices.

its roots in law studies, intersectionality in this view is also a method in the sense of an operative approach to law, society, and their symbiotic relation, or “a method for fitting law to reality, rather than reality into law” (MacKinnon 2013: 1024). It captures not just the static outcomes of the problems it brings into view, but also their dynamics.

Thus, intersectionality addresses not classifications in themselves but rather power relations, and particularly the “white supremacy and male dominance, or white male dominance” (MacKinnon, 2013: 1023), at the expense of women and particularly racialised women. This approach thus exposes what can be referred to as “the dominant framework of discrimination” (Crenshaw 1989: 152). Hence, adopting an intersectional approach means targeting the forces that create the outcomes, and not just their static products. These dynamics are not recognised by conventional discrimination analysis, which rather than challenging the power relations that underlie inequalities, tends to mirror them. As MacKinnon (2013: 2023) concludes:

Categories and stereotypes and classifications are authentic instruments of inequality. And they are static and hard to move. But they are the ossified outcomes of the dynamic intersection of multiple hierarchies, not the dynamic that creates them. They are there, but they are not the reason they are there.

Still, decades after the incorporation of an intersectional perspective into much research and the uptake of intersectionality across disciplines, there is no consensus about what a critical intersectional analysis entail (Rajani, 2022). Several scholars have referred to how intersectionality has been used as a buzzword, devoid of meaning (Davis, 2008; Hill Collins, 2015), and how it has been depoliticised and neutralised in much contemporary academic scholarship (Bilge, 2013). As a response to this lack of consensus and the perceived depoliticisation of the concept, Patricia Hill Collins and Sirma Bilge (2016) propose six core ideas that should be present in a critical intersectional inquiry. These include: first, that social inequality needs to be understood as interactions between socially constructed categories; second, that power is mutually constructed by many diverse power relationships, such as sexism and racism;

three, that relationality and connections between gender and race should be prioritised over differences between categories of social identities; four, that there must be an awareness of the social context; five, that there must be a recognition of complexity of analysis; and six, that critical analysis must contain social justice motivations that challenge the status quo as an outcome.

For this thesis, adopting an intersectional approach means to propose a theoretical and methodological approach from which it is possible to understand that these multiple categories of difference play a specific role and must be equally addressed, understanding that the intersections of these categories are more than the sum of their parts (Jubany and Guasch, 2020). According to this understanding, each person is subject to and at the same time interprets their personal experience, creating an order to the narratives that build their multiple social positions and their self-definition. These experiences and self-representations must therefore also be situated and contextualised (Anthias, 2013). Further, intersectionality also helps shape the way that violence is defined, as a social problem that “is shaped by and helps structure intersecting power relationships” (Hill Collins, 2017).

In the next section, I will look at theoretical constructions of one aspect of the context of violence relevant to this thesis, namely the role of technologies and the mutual shaping of technologies and society.

2.2. Technopolitics and digital social inequalities

2.2.1. Technopolitics and technosociality

The public sphere can be perceived as a space where citizens act out and critically debate political life, through an open dialogue in which everyone can participate and raise their concerns (Habermas, 1989). Much of the scholarly work which draws upon Jürgen Habermas’ public sphere theory – and applies it on the internet as a public sphere – considers the act of public debate to be a process distinct from the socio-economic realities that both enable and constrain individuals. More recent theorisations, however, have expanded and nuanced our understanding of online participation by delving into various aspects related to

the structure and lineage of commercial participatory platforms (Gillespie, 2010; 2015), the governance of peer-to-peer projects (O'Neil, 2009), their gendered construction (Collier and Bear, 2012) or non-participation as a form of mediated political action (Casemajor et al, 2015). Following this more nuanced understanding of the possibilities of online participation, as well as the work of scholars such as anthropologists Miller and Slater (2000) or sociologist Saskia Sassen (2002), this thesis argues that we must recognise the embeddedness and the variable outcomes of new technologies for different social orders. New technologies can be constitutive of new social dynamics, but also merely reproduce older structures and conditions. In other words, technologies are inscribed by power, contestation, inequality, and hierarchies, in complex imbrications of technology and society (Sassen, 2002). Technologies, thus, can be read in their social character, which is inevitably also a political character (Latour, 2007). This can be applied, for example, on social media platforms: Far from being straightforward private companies or marketplaces, they represent a type of non-neutral digital infrastructures that contribute to determining social, political, and economic relationships.

In view of this, “technopolitics” is the strategic design or use of “technology to constitute, embody, or enact political goals” (Hecht, 2001: 257). The concept of technopolitics delves into the long-term impact of technology, aiming to unveil the trajectories that culminate in a particular phenomenon rather than merely describing an awakening to realities perceived as new or sudden (Udupa et al, 2021). In this way, the concept of technopolitics accounts for how the “same” technology can be captured by competing actors and discourses to profoundly affect the way it is used and the shape it takes (Gagliardone, 2019). A related concept is “technosociality”, defined by anthropologists Mizuko Ito and Daisuke Okabe (2005) to refer to the integration of technology, social practice and place, resulting in technologically mediated social orders. Following this understanding, the technical and the social dimensions are “inseparable outcomes of ongoing and historically contextualized practice” (Ito and Okabe, 2005: 259). Technologies are not seen as neutral - a mere addition to a social system - or as

determinative, but rather as an embedded and co-constituting aspect of society, its structures, cultures, and practices (Powell and Henry, 2017). These concepts facilitate a reflection on how technologies *mediate* interaction and inequalities online at different levels, including coding, content production, content moderation, social interaction, and political debate. This is further addressed in the next section.

2.2.2 Living in media, mediated interactions and polymedia

Due to the expansion of the media system in this increasingly globalised world, also the societal institutions have undergone a shift towards self-reflective commentary and positioning, particularly towards the media (Jameson, 1991). The media system has in this sense taken over the former role of authorities, leading our attention away from power imbalances and towards self-identity as a life project (Giddens, 1991). As proposed by Mark Deuze et al (2012), media has become so inseparable from us that we can no longer say that we live with media, but we rather live *in* media. At the same time, media has become pervasive and embedded in our lives to the extent that it has become largely invisible. Media and technology from this perspective can be described as this thing that fades to the background of our world, but which also creates that world (Arthur, 2009: 10). The elements inherent in the relationships between media and the human connection, can be observed in the following broader societal trends: First, a primacy of self-governance and self-reliance in contrast to a previous deference to authorities; second, an extension of community premised on simultaneous co-presence and telepresence as directed by the individual and their concerns; and third, the emergence of mass self-communication next to mass communication, which signifies the shift from survival values to self-expression values (Deuze et al, 2012).

As media become invisible and omnipresent, at the same time, they become all-powerful:

In this abundantly mediated and progressively mobile lifestyle media are such an augmented, automated, indispensable and altogether inalienable part of one's activities, attitudes and social arrangements that they

disappear – they essentially become the life that people are experiencing on a day to day basis (Deuze et al, 2012: 3).

In this sense, digital technologies evolve faster than our cultural, legal and educational institutions can keep up with them (Bolter and Grusin, 1996). At the same time, digital spaces open windows to other worlds, where local idioms and traditions are fused and hybridised with global media brands and genres (de Block and Buckingham, 2007). Nevertheless, as Deuze et al (2012) highlight, media activities and practices can only be understood in a context that includes both spatial and material consideration, reflecting how the social arrangements both stretch existing ways of doing things and making sense of the world across cultural and spatial boundaries, while at the same time functioning to articulate and demarcate local communities and identities.

Further, life in media can be understood at the same time as connected and isolated (Deuze et al, 2012). This is because, on the one hand, it allows people to retreat into their own “personal information space” (Deuze, 2007); and on the other hand, this individualised immersion connects people with others, “turning their very own bubbles of space into fully mediated spaces of global coexistence” (Deuze et al, 2012: 5). Different types of computer-mediated communication may have positive effects on community interaction, involvement and social capital, enabling people to keep in touch with old friends, colleagues and acquaintances, this way deploying media largely in the service of connections (Purcell, 2006).

Closely connected to the concept of living in media, is that of *mediation*. Applied to the area of new communication technologies and interpersonal relationships, the concept of mediation draws on debates related to the social shaping of technology (Wajcman, 2002) from a dialectical approach. This can be exemplified for instance, when we choose a certain medium for an interaction, as this interaction always becomes partly constituted by the medium through which it is expressed (Madianou and Miller, 2012). From the perspective of mediation, email is not just email; it is defined relationally also as “not a letter”, “not a text message” (Ibid.) or “not a Zoom conversation”. This perspective is thus different from perspectives on media technologies that focus on each individual medium

and its affordances and can be understood as rooted in a structuralist perspective, as developed by Claude Lévi-Strauss (1963). That is, an entity is determined as much by its relationship to that which it is *not* as to that which it appears to be (Madianou and Miller, 2012).

Finally, regarding relations between persons and media, the concept “polymedia” (Madianou and Miller, 2012) is particularly relevant to describe the emerging environment of proliferating communication opportunities. This concept relates to how the evolving use of a progressively integrated range of communication technologies affects the way interpersonal communication is carried out and experienced. This phenomenon is relatively new, emerging in the past decade due to the proliferation of platforms that offer users a diverse range of communication opportunities, unrestricted by the specific features of each medium. In polymedia conditions, the focus transitions from the individual characteristics of each medium as a distinct technology to viewing new media as an environment with various capabilities. Polymedia is not just about the environment; it is about how users leverage these capabilities to navigate their interactions, relationships, and emotions. Thus, the emphasis shifts from the limitations of each medium to the social and emotional repercussions of choosing among these different media (Ibid). In this context, the concept of polymedia contributes to a re-socialization of technology, as the responsibility of choice moves from technical and economic considerations to moral, social, and emotional concerns.

2.2.3 Digital technologies, culture and “the other”

As pointed out in Miller and Horst’s principles of digital anthropology (2012), it is important to note how the digital itself intensifies the dialectical nature of culture. Dialectic in this regard refers to “the relationship between this growth in universality and particularity and the intrinsic connections between their positive and negative effects” (Miller and Horst, 2012: 3). Contrary to some popular discourses claiming that cultures and identities have been homogenised through the emergence of the internet, the digital and its globalising characteristic is not necessarily homogenising (Miller and Horst, 2012). Instead, local identities are

often strengthened by the connection through globalised technologies, as people begin to emphasise their uniqueness overtly only when it appears to be threatened (see e.g., Hylland Eriksen, 2007). From this view, identity obtains its meaning primarily from being contrasted with the identity of the other (Barth 1969; Van Meijl 2008). Or, as Zygmunt Bauman (1991: 14) puts it,

in dichotomies crucial for the practice and vision of social order the differentiating power hides as a rule behind one of the members of the opposition. The second member is but the other of the first, the opposite (degraded, suppressed, exiled) side of the first and its creation. Thus, abnormality is the other of the norm ... woman the other of man, stranger the other of the native, enemy the other of friend, 'them' the other of 'us'.

This leads to the concept of “othering”, which refers to a “process of differentiation and demarcation, by which the line is drawn between “us” and “them” (Lister 2004: 101). Whilst the concept can be traced back to Simone de Beauvoir (1949) and her analysis of woman as socially constructed as “the other” in relation to man, many theorists have since then developed this concept further. In Gayatri Chakravorty Spivak’s (1985) notion of othering, this concept concerns the consequences of racism, sexism, class - or a combination - in terms of “symbolic degradation” as well as the processes of identity formation related to this degradation. As put forward by Spivak, three dimensions can be defined in processes of othering: first, power, making the subordinate aware of who holds the power, and hence about the powerful producing the other as subordinate; second, construction of the other as pathological and morally inferior; and third, depriving the other of knowledge and technology, which is instead defined as the property of those holding the power. As Sune Jensen (2011) highlights, this notion has a lot in common with intersectionality and “interlocking systems of oppression” (Hill Collins, 1989), as othering concerns “the consequences of racism, sexism, class (or a combination hereof) in terms of symbolic degradation as well as the processes of identity formation related to this degradation” (Jensen, 2011: 65). This means that to speak of othering is not an alternative to speaking of racism(s)/sexism or class, but rather a way of addressing an aspect thereof (Wren, 2001: 144).

Through processes of othering, categories formed along a hierarchy of racial, religious, and gender-based oppression, stemming from the era of modernity, serve as tags by which individuals are perceived and comprehended, shaping how they are treated, marginalised, excluded, and discriminated against. It is a dynamic grounded on the multi-layered processes of domination and subordination that determines “us” versus “them”, in which the construction and denial of “the other” emerges as the ultimate expression of social control (Jubany, 2017). Conclusively, “the others” are reduced to stereotypical characters and are ultimately dehumanised (Lister, 2004: 102). Such processes imply reduction and essentialisation in the sense that those who are othered are reduced to a few negative characteristics (Jensen, 2011).

In relation to the inscription of inequalities in digital spaces, the concept of “online othering”, as proposed by Emily Harmer and Karen Lumsden (2019) facilitates an analysis of

the myriad behaviours, interactions and discourses which seek to (re)draw boundaries in, around, and between virtual spaces, and shape the rules and norms concerning which individuals and groups are endowed with status and legitimated to participate in these spaces, and those who are not (Harmer and Lumsden, 2019:15).

This concept allows for an examination of the justifications and motivations of those who perpetrate or enact discrimination, prejudice, hate or abuse in online spaces. In the same way, the concept of othering applied on interactions in and affordances of digital spaces can help understand and conceptualise “how exclusion from the internet and its spaces operates across various groups, individuals and contexts” (Ibid.). This facilitates a view of a range of behaviours and interactions in digital spaces as based upon the same mechanism of creating, recreating and performing “we” and “them”, whether based on gender, sexual orientation, gender identity, racialisation, or any other factor or intersection of factors. Hate speech is one way through which othering takes place in digital spaces and beyond. In the next section, I shall discuss theories surrounding another aspect of digital spaces, namely digital embodiment.

2.2.4 Digital embodiment

The body - as something we both have and are - is how we exist, the very act of existing (Turner, 2000: 489). It is a certain setting in relation to the world (Merleau-Ponty, 1962), as the point where perception begins. Or, as argued by Pierre Bourdieu, the body can be understood as a reservoir of social experience. In this sense, the principle that generates and unifies all practices,

is nothing other than the socially informed body, with its tastes and distastes, its compulsions and repulsions, with, in a word, all its senses, that is to say, not only the traditional five senses-which never escape the structuring action of social determinisms-but also the sense of necessity and the sense of duty, the sense of direction and the sense of reality, the sense of balance and the sense of beauty, common sense and the sense of the sacred, tactical sense and the sense of responsibility, business sense and the sense of propriety, the sense of humor and the sense of absurdity, moral sense and the sense of practicality, and so on (Bourdieu, 1977: 124).

Embodiment refers to the idea of the body as not only a physical and metaphysical, but also a cultural phenomenon (Powell and Henry, 2017) or, as put forward by Michel Foucault (1990), the body as a social construction and a site for governance. In this discussion of embodiment as a paradigm for anthropology, the body is the fundamental condition for the existence of culture and the subject and, thus, for the analysis of these (Csordas, 1990). Embodiment, from this perspective, overcomes a dichotomic analysis of body and mind, deconstructing this dualism, in a similar way that *habitus* is used by Bourdieu to collapse these dualities (Ibid.: 11).

Theories of embodiment have been central to feminist understandings of gendered relations of power, control and subordination. The body is a useful way of conceiving the construction of the self vis-a-vis the social world (Powell and Henry, 2017), as “embodied subjectivity” (Merleau-Ponty, 1968), which can be understood in the same way as “lived experience” (Powell and Henry, 2017). This perspective places consciousness and perception in the body itself. Building on theories of embodiment from a feminist approach, the body is “the threshold through which the subject’s lived experience of the world is incorporated and realized, and as such, is neither pure object nor pure subject” (McNay, 1999: 98,

cited in Powell and Henry 2017: 55). This also resonates with Butler's theory on performing gender:

Just as a script may be enacted in various ways, and just as the play requires both text and interpretation, so the gendered body acts its part in a culturally restricted corporeal space and enacts interpretations within the confines of already existing directives (Butler, 1988: 526).

Nevertheless, the myth of *disembodiment* in gender relations in digital spaces is still highly present in the public mind (Powell and Henry, 2017). The web is still largely perceived as an alternate, disembodied space, where one's "real" identity does not influence interactions. Cyberfeminists such as Donna Haraway (1987) or Liz van Zoonen viewed the new technologies as possibilities to "escape from bodily gender definitions and construct new identities" (van Zoonen, 2002: 260). However, as Anastasia Powell and Nicola Henry (2017) argue, this perception can be linked to a constructed dichotomy between the online-offline. Breaking with this dichotomy, later strands of cyberfeminisms have focused less on a utopian vision of technology, adopting a more nuanced perspective on the relationship between the body, gender and technologies, where technology is viewed as sometimes liberating and sometimes oppressive (Wajcman, 2013)

Building on these theories of embodiment, digital spaces cannot be analysed as disembodied, but we must analyse also digital life and any engagement with technology as lived through and mediated by our bodies. Or as noted by Silvia Plant (1995: 60), there is "no escape from the meat, the flesh". Through this understanding it becomes clear how harms inflicted through digital technologies produce embodied effects, because they are lived through our bodies and thus affect all aspects of our bodies, including the physiological, relational, intersubjective, and affective aspects (Powell and Henry, 2017: 62).

2.2.5 Digital social inequalities and the gendering of digital spaces

Following this notion of digital spaces as embodied, embedded, and co-constituting aspects of society, it becomes apparent that these spaces are also characterised by *digital social inequalities* (e.g., Halford and Savage 2010;

Ragnedda and Muschert 2013), at several levels. That is, digital spaces are not neutral spaces.

Analysing the internet's roots in historically male-dominated institutions such as the military and academia, the prevailing coding culture and internet norms can be seen as products of the hyper-masculine influence defining the medium since its inception (e.g., Megarry, 2014; Citron, 2009; Barak, 2005). In digital spaces, the default, dominant group is still a white man (Everett, 2009). An increasingly important level where digital social inequalities have an impact is in the very coding of these spaces. Humans design technological systems that mirror existing power structures, embedding pre-existing biases about social differences into the fundamental architecture of the internet (Noble, 2018). Additionally, the economic framework of the internet relies on the commodification of its users, as offering free access to platforms transforms users into targets for advertisers (Fuchs, 2009). Furthermore, the design of social media platforms itself may promote specific behaviours, potentially hindering equal participation (e.g., Massanari, 2017). This combination of private interests in promoting certain sites and biased coding, in addition to the monopoly status of a relatively small number of internet search engines, leads to a biased set of algorithms. This bias tends to privilege whiteness and discriminate against racialised persons, and especially racialised women (Noble, 2018). Thus, digital technologies can be understood as facilitating the “re-embedding” (Adkins, 1999) of social inequalities in relation to gender, racialisation, class, disability, and sexuality. As a result, even though digital technology has the potential to democratise communication processes, existing social, political, and economic inequalities significantly influence people's ability to engage online and the way such participation unfolds (e.g., Harmer and Lumsden 2019).

In the analysis of this thesis, the most salient aspect of this re-embedding of social categories and structures is the gendering of digital spaces; how gender intersects with other categories in this regard and the implication of this for the use of internet as a public space. Whilst the embeddedness of the web is undoubtedly gendered, however, this gendering is also coupled with a great

variability by place, age, class, racialisation, nationality, issue-orientation (Sassen, 2002). Thus, intersectionality must also be considered in analysing the embeddedness of digital spaces, as “digital technologies reproduce white, heterosexual, masculine cultures and hierarchies of power” (Kendall 2002, cited in Daniels, 2017: 81). In this regard, a key point of analysis of this thesis is to explore to what extent the intersections of digital social inequalities across these categories are reflected in women's experiences online, and particularly in relation to feminist activism and gender-based online hate speech. For example, as argued by Jessica Megarry (2017), the encoding of male bias in platforms, practices, algorithms, and bots may pose a problem for feminist activists as these cannot ensure that materials reach their target audiences.

Thus, any examination of digital space needs to consider both its embeddedness in the larger societal, cultural, subjective, economic structures, and its entanglement with everyday practices and embodied knowledge (Bareither, 2019). This perspective on digital spaces allows for a perspective that goes beyond the duality between utopian and dystopian understandings of the internet (Sassen, 2002), and beyond the duality of the online and offline. That is, the internet cannot be treated as a kind of “placeless place” but needs to be analysed along the same structures that impact on our experiences in non-digital spaces (Miller and Slater, 2000). And because our virtual bodies bring physical world manifestations into digital spaces – gender, racialisation, class, sexuality, religion, education, etc. – we replicate “real world” inequalities into virtual space (Gray, 2013).

2.2.6. Platform moderation and embedded bias

The responsibility and mechanisms of social media platforms in terms of protecting users from harm tends to be linked primarily to their capacity to handle reports and thus moderate content. In this context, content moderation cannot be perceived as just an occasional act that platforms must engage in. Rather, it needs to be understood as a fundamental aspect of their place in public discourse: essential, constitutional, and definitional (Gillespie, 2018). Platforms not only depend on moderation for survival but also derive their essence from it.

Moderation is inherent from the beginning and persists throughout, albeit mostly concealed. This concealment is necessary, partly to sustain the illusion of an open platform and partly to evade legal and cultural responsibilities: “Looking at moderation in this way shifts our view of what social media platforms really do: from transmitting what we post, to constituting what we see. There is no position of impartiality” (Ibid.: 21). That is, “there is growing recognition that online content moderation is not merely a matter of technical capacity or corporate will but also a serious issue for governance” (Udupa et al, 2022: 7). Moderation, thus, is a highly complex task – for example, platform policies often go much further than just illegal content since there are categories that are legal, but still cause harm to users¹³. As Tarleton Gillespie (2018: 9) reminds us,

Moderation is hard because it is resource intensive and relentless; because it requires making difficult and often untenable distinctions; because it is wholly unclear what the standards should be; and because one failure can incur enough public outrage to overshadow a million quiet successes.

Moderation of harmful expressions is also complex because of the heterogeneous and dynamic nature of online hate speech: it takes many different shapes, and those shapes can shift and expand over quite short spaces of time (Brown, 2018; Citron, 2014; Delgado and Stefancic, 2014). In addition to this, content does not only include text, but also memes, multimedia materials, hashtags, tagging and other forms (Siapera, 2019).

Different platforms have different moderation mechanisms. Caplan (2018) distinguishes between “artisanal” approaches, where case-by-case governance is normally performed by between 5 and 200 workers (platforms such as Vimeo, Medium and Discord); “community-reliant” approaches, which tend to combine formal policy elaborated at the company level with volunteer moderators (platforms such as Wikipedia and Reddit); and “industrial-sized operations”, where tens of thousands of workers are employed to enforce rules produced by a policy team (large platforms such as Google and Facebook).

¹³ The Markup, Understanding the Digital Services Act, available at, <https://themarkup.org/newsletter/hello-world/understanding-the-digital-services-act> last accessed on April 19, 2024.

In the realm of content moderation, Artificial Intelligence (AI) is often considered a solution, using methods ranging from simple database scans to sophisticated machine learning models (Udupa et al, 2022). However, AI detection mechanisms face challenges in uncovering complex forms of online hate speech, primarily due to limitations in the quality, scope, and inclusivity of training datasets (Ibid.). The bias inherent in classification algorithms and training data, stemming from a homogenous tech workforce, further exacerbates these limitations (e.g., Noble, 2018). Furthermore, cultural contextualisation is crucial, as the lack thereof results in false positives and over-moderation (Udupa et al, 2022).

Addressing these issues, initiatives like Google's Perspective API and Twitter's Birdwatch incorporate user experiences through crowdsourcing models. While these models show promise in detecting evolving extreme content, they are susceptible to false positives and racial bias (Ibid.). Additionally, whilst racist and LGBT-phobic tweets are more likely to be identified as hate speech in the North American context, gender-related comments are often treated as merely offensive speech (Davidson et al, 2017). Attempts to involve humans in the annotation process are met with challenges, as this approach assumes that human involvement can offset the limitations of machine detection, but humans bring their own biases. The definitional problem of hate speech further complicates these efforts, as at an international level, there is no consensus on what constitutes hate speech legally or culturally (Udupa et al, 2022).

2.3. Conceptualisations of violence

Categories serve both to perceive, explain and understand reality and to guide action (Mancinelli et al, 2023). That is, the way we label any phenomenon is a means of classifying and identifying it, influencing both the conception of the problem and how to approach it (Comas, 2011). The meaning of violence is a historical and cultural construction; that is, the way violence is constructed, for what purposes, how it manifests, and how it is recognised is shaped by society and social relationships (Zurita, 2011). Society generates violence through intricate processes involving negotiations (determining what qualifies as violent or not), struggles (challenging specific forms of violence while accepting others),

actions, and representations (Ibid.:5). In conceptual terms, this thesis locates gender-based online hate speech at the intersection of gender-based violence and hate crimes. This framework has also been proposed by scholars such as Lewis et al (2019). In this section, I will discuss the theoretical frameworks related to different conceptualisations or typologies of violence that come into play when analysing gender-based online hate speech, namely gender-based violence, hate crimes, and speech as violence.

2.3.1 The continuum of gender-based violence

Any discussion of technology-facilitated gender-based harm must begin on the one hand - as I have exposed in previous sections - with an analysis of the ways in which sex, gender, and power in intersection with other factors both shape and are shaped by technologies. On the other hand, it needs to comprise a comprehension of the power relations of sex and gender, which underlie gender-based violence in general (Powell and Henry, 2017). Thus, another crucial theoretical cornerstone that comes into play when analysing online gender-based hate speech is related to debates surrounding gender-based violence.

A gender perspective allows us to account for the way in which the configuration of a gendered body order takes place through various practices, representations, and experiences in both public and private spaces. From this perspective, gender-based violence can be understood as constitutive and at the same time a product of a framework of gendered power relations (Biglia, 2007: 13). Violence is thus not a random or individual response to concrete situations, but it is based on a collective construction and its function is to uphold gendered social structures rather than to attack these (Juliano, 2004).

Physical violence tends to be treated as the prototype of violence (Powell and Henry, 2017: 66), which becomes evident in relation to discourses both on gender-based violence and on sexual violence. This perception renders invisible other types of violence. The concept of *symbolic violence*, thus, becomes useful, to explore the hierarchy of violence, where physical violence tends to be conceived as the main form of violence. In contrast to this, symbolic violence, as defined by Bourdieu (2002), is power and domination that is naturalised,

universalised and internalised, both by the dominator and the dominated, such as in the case of gender relations. In fact, Bourdieu's later work defines gender domination as "the paradigmatic form of symbolic violence" (Bourdieu and Wacquant, 1992: 170). This hierarchy of violence, in turn, influences how online harms are perceived (Powell and Henry, 2017) - as often not even characterised as acts of violence, although they contribute to social subordination and exclusion.

A critical concept to go beyond the view of physical violence as the only form of violence, is the *continuum of sexual violence*. This concept was coined by Liz Kelly (1987: 48), to refer to the various forms of abuse, coercion, and force that men use to control women. From this perspective, sexual violence, whether in the shape of gender-based hate speech or a physical transgression, is understood as both a manifestation of and a tool for gendered power relations. Subsequently, recognizing a diverse range of experiences of gender-based harms allows for an understanding of the individual impacts, cumulative effects, and collective implications of these harms (Powell and Henry, 2017). As highlighted by Stanko (1985), this continuum of violations results in women being "continually on guard" in relation to the possible threat of male violence. I would add to this perspective also the experiences of LGBT+ persons and especially trans persons or non-binary persons, who also need to constantly relate to the threat of violence.

Further, in relation to this last point, the concept of surveillance becomes relevant. In recent years, social media platforms have been portrayed as providing distinctive opportunities for women to challenge the prevailing social and political order. However, this increased visibility also makes their activism easily traceable and susceptible to dismissive interjections (Megarry, 2017). Surveillance is here understood as "the coupling of information collection and the use of power" (Andrejevic 2015). Studies such as that of Elisabeth Stoycheff (2016) and Rodrigo Gomez Garcia and Emiliano Treré (2014) show that fear of surveillance has resulted in activist social media users feeling hesitant to openly express dissenting opinions online. In relation to gender-based harassment,

surveillance may range from spyware and stalking apps (Powell and Henry, 2017), to anti-feminists surveilling feminists on social media and coordinating attacks. Following this, a crucial question for this thesis is how male surveillance - or techniques of watching (Megarry, 2017) - may contribute to shaping the type of activism emerging from social media platforms, such as Twitter.

In the next section, I shall discuss the other conceptual pillar of violence where the online violence explored in this thesis can be located, namely the hate crime framework.

2.3.2. Hate crimes

As argued by Neil Chakraborti and Jon Garland (2015: 2),

‘hate’ is a slippery, emotive and conceptually ambiguous label. Hate can mean different things to different people, and that has important implications for the way in which we conceive of the offences that fall under its umbrella framework.

Hate as understood in the context of the hate crime framework is not simply an emotion; rather, it represents structured patterns of oppression: for example, race and racism; sex and sexism and heterosexism (Schweppe and Perry, 2022). In this respect, this framework highlights how biases are socially reinforced and not simply personal (Wang, 1999: 818).

In terms of hate crime theories, Barbara Perry’s work (2001) is particularly useful for analysing bias-motivated crimes as an arena where predominantly white males reaffirm “their place in a complex hierarchy and respond to perceived threats from challengers of the structure” (Perry, 2001:2). As Perry highlights, hate crimes must be understood as processes rather than time-constrained one-off events. This specific aspect of hate crimes fits well with the characteristics of online hate speech. That is, the crime does not occur in a cultural or social vacuum, nor is it “over” when the perpetrator moves on (Ibid: 8). Because of this, hate crime must be defined as a socially situated, dynamic process involving context, actors, structure and agency, and must therefore be understood as perpetrated within a framework of specific relations of power (Ibid.). This means acknowledging that bias-motivated violence is not “abnormal” or “anomalous”

in relation to the social structures, but that this violence needs to be analysed as an extension of the racism, sexism, and LGBT-phobia that allocates privilege along racial and gender lines:

in a generally racist, sexist, homophobic culture, violence motivated by hatred is not deviant behavior. In fact, it conforms to what is a normatively unjust value system. It is an affirmation of the gendered and racialized hierarchy that constitutes the 'legitimate' social order (Perry 2001: 35)

Within this conceptual framework, hate crime can be seen as a mechanism of power and oppression, aimed to reaffirm the precarious hierarchies that characterise a given social order. As such, it is a means of marking both the Self and the Other in a way that reestablishes their "proper" relative positions, as established and reproduced by broader patterns of social and political inequality (Perry, 2001: 9). Hate violence reveals that the personal is political; that "such violence is not a series of isolated incidents but rather the consequence of a political culture which allocates rights, privileges and prestige according to biological or social characteristics" (Sheffield, 1995: 438). In this regard, the specific victim is often almost immaterial, as the victims of these symbolic acts are interchangeable. The target audience, then, is not so much the victim as it is others like them.

Misogynistic violence as hate crime

Looking more in-depth at the gendered aspects of hate crimes, much of the violence expressed against women is indistinguishable from other hate crimes, in that it is intended to intimidate and control the wider group of women. This violence is "meant to teach women, both individually and collectively, a lesson about remaining accountable to their femininity" (Perry 2001: 83). Violence against women in this regard can be defined as a "'classic' form of hate crime" (Ibid.) as it terrorises the collective by victimising the individual, reaffirming the superiority of the male perpetrator in relation to the female victim (or victim not conforming to normative gender constructions). Victims are chosen because of their gender and because of the assumptions about how they should enact their gender. Drawing upon the work by Crenshaw (1989) in relation to

intersectionality, Black women are uniquely vulnerable to gendered violence because of their multiply determined structural disempowerment.

Gendered hate crimes can be linked to the wider continuum of gender-based violence. Following Perry's approach, all these different types of violence can be analysed as having the same function, of putting women into place, whether it is within the limits of their own homes or in the public sphere. As argued by Jane Caputi (1993: 7), regardless of where a particular act lies on the continuum of violence, it can be analysed as a "ritual enactment of domination, a form of terror that functions to maintain the status quo".

Also sexual violence and harassment fall within Perry's wide umbrella of gender-based hate crime, as it draws upon and recreates power relations. As highlighted by MacKinnon (1991:1302), sexual assault can be analysed as an institutionalised rather than aberrant means by which (heterosexual) men can perform their masculinity while "symbolizing and actualizing women's subordinate position". This becomes relevant in relation to the language used in online hate against women, which is oftentimes sexualised and containing rape threats. Scholars such as Jane (2018) has referred to this discourse - which shares many characteristics across countries and online cultures - as "rapeglish".

When applying an analysis of these types of violence as having the function of subordinating women and "putting them in their place", it is not far-fetched to draw the conclusion that an increase in gender-based violence and hate crimes often emerge in the context of what is perceived by men as a loss of relative position. As I have argued in the Introductory chapter of this thesis, this can be analysed as the backdrop to the new online gender hate and anti-feminism (Siapera, 2019). Within this context, attempts to challenge the collective dominance of men are frequently met with aggressive efforts to reestablish the "natural" superiority of men: "Assaults, rapes, or homicides that are outside the bounds of an intimate relationship tend to be directed at individual women as proxies for the combined threats to masculine domination represented by women as a class" (Perry 2001:86). Within this framework, the digital space can be analysed as a new venue for this already existing type of violence, expressed

verbally or through images, but bearing the same collective message of threat to all women and non-gender conforming individuals.

Nevertheless, there are challenges to conceptualising misogynist online abuse as hate crime. On the one hand, there has been a certain institutional and scholarly reluctance to recognising violence against women as a hate crime. This has been interpreted as related to institutional sexism and patriarchal ideology that does not acknowledge gender-based prejudice (Gill and Mason-Bish, 2013). On the other hand, there are also challenges to the conceptualisation as such. As put forward by Lewis et al (2019: 123) hate crimes are characterised as being “signal crimes”, that is, crimes aimed at communicating to wider communities that they are not welcome, that they are inferior or at risk. This requires that the offence either occurs in a public domain, or that the news of the crime is spread to a public domain. As Lewis et al argue, a part of online abuse is experienced in an online environment where the distinction between public and private space is complex. Another challenge, at least from a legal perspective, as argued by Víctor Gómez (2016) is the view of hate crimes and hate speech as aimed at attacking minorities or vulnerable groups and that since the safety or dignity of *all* women is not likely to be threatened by this type of speech, it should not be prosecuted legally as hate speech. Whilst I shall further discuss the legal ramifications and interpretations in chapter 5 of this thesis, in the next section I will explore the notion of how speech can enact violence.

2.3.3. Words that hurt: The performativity of speech

As Joao Pina-Cabral (2019) highlights, people are entangled also through language. As anthropologists we must reject the assumption that “persons are entities that pre-exist their own engagement with each other (their entanglement); that they exist outside signification” (Pina-Cabral 2019: 3). Sociality and communication are intrinsically related, as “from the moment that there is meaning there are nothing but signs” (Derrida, 1976: 50). Language, in this regard, is imbued with power relations that language itself both expresses and reproduces.

Language and symbolic power

In the same way that digital spaces, as discussed above, are inscribed by wider societal structures, so is language. This has been conceptualised, for example, in the work of Pierre Bourdieu in relation to language and symbolic power. As Bourdieu argues (1991), common linguistic exchanges can be analysed as situated encounters where every linguistic interaction bears the traces of the social structure that it expresses and at the same time reproduces. In this sense, this is a critique against the semiotic analyses developed by structuralists in that they are purely internal to a text and ignore the socio-historical context in which the text was produced. Rather, language itself is a social historical phenomenon and linguistic exchanges are mundane, practical activities like many others (Ibid.). Thus, speaking cannot be thought of as the mere execution of a pre-existing linguistic system but must be understood as a much more complex and creative activity.

Bourdieu's notion of "linguistic markets" is particularly relevant here as it explores how some expressions produced in specific contexts or "markets" - on a given linguistic market, are more valued than others (Ibid.). Speakers need practical competence to produce expressions that are highly valued in the market concerned. This can be related to forms of verbal abuse in social media, and how hate speech evolves as perpetrators learn how to best express hate whilst avoiding being caught by systems of automatic machine detection, e.g., through "dog whistles".¹⁴

Further, the idea of symbolic power – related to the concept of symbolic violence briefly exposed above - is also key to this framework. As has been previously exposed, the term "symbolic power" can be used to refer to an aspect of most forms of power as they are deployed in social life. In day-to-day life, power is usually not expressed as overt physical force, but is instead transmuted into a

¹⁴ A dog whistle can be defined as "a coded message communicated through words or phrases commonly understood by a particular group of people, but not by others", primarily used to describe political speech. Definition taken from Merriam-Webster, *What's the Political Meaning of Dog-Whistle*, available at <https://www.merriam-webster.com/words-at-play/dog-whistle-political-meaning>, last accessed on April 19, 2024.

symbolic form, and is thus endowed with a kind of legitimacy that it would not otherwise have. In this way, symbolic power is an “invisible” power, however it is “misrecognised” as such and thereby “recognised” as legitimate:

the efficacy of symbolic power presupposes certain forms of belief, in such a way that even those who benefit least from the exercise of power participate, to some extent, in their own subjection. They recognize or tacitly acknowledge the legitimacy of power, or of the hierarchical relations of power in which they are embedded; and hence they fail to see that the hierarchy is, after all, an arbitrary social construction which serves the interests of some groups more than others (Bourdieu, 1991: 23).

This concept of symbolic power highlights how the exercise of power through symbolic exchange always builds upon shared belief, that is, it is based on common social and cultural norms, and also imbued in the very language itself.

Speech acts and illocution

Linked to the above, John Langshaw Austin’s (1962) theory on speech acts and illocution also becomes relevant for this thesis. Austin distinguishes between perlocutionary acts, which are acts performed by saying something, and illocutionary acts, as acts performed in the moment of saying something. The illocutionary speech act is itself the deed that it effects; whilst the perlocutionary leads to certain effects that are not the same as the speech act itself (Butler, 1997: 3). Perlocution can be illustrated by how words function when someone shouts, “the building is on fire” and because of hearing those words, people start exiting the building. Illocution, on the other hand, can be exemplified with how words are used in the act of marriage. However, as Austin argues, it is not the words themselves that constitute the illocutionary force, but rather the action of uttering the words, *the speech act*. For the words to perform an illocutionary action they must be spoken in a particular context, and it is also essential that uptake is secured, that is, “bringing about the understanding of the meaning and of the force of the locution” (Austin, 1962: 116). That is, to a certain extent this theory also considers the social conditions of communication. Nevertheless, Austin has been criticised for not fully considering what it might mean to treat what he refers to as “conventional means” as social phenomena integrated in social relations imbued with power and authority (e.g., Bourdieu, 1991). Bringing

the concept of illocutionary force to the analysis, however, still facilitates an understanding of how a speech act in a specific social setting not just constitutes a reflection of unequal social relations, but how this speech act can perform and reproduce social subordination.

Linguistic vulnerability and hate speech

As Butler (1997) suggests, linguistic vulnerability, that is, the possibility of words to hurt, can be linked to the very recognition of one's existence through words. Following the concept of "interpellation" (Althusser, 1971), there is no inherent meaning in the individual, but there are only subjects who come into being when they are "hailed" or interpellated by ideology. That is, from this view, the subject exists only as they are recognised in a specific way that has a social structure as its referent. The subject is thus preceded by social forces, or "always-already interpellated" (Ibid.). An example put forward by Louis Althusser of how a subject is constituted in language and thus interpellated, is being called a name. Building upon the concept of interpellation, as Butler highlights, "to be addressed is not merely to be recognised for what one already is, but to have the very term conferred by which the recognition of existence becomes possible (1997: 5). Then, if language can sustain the body; language can also threaten its existence. From this view our very condition as linguistic beings also makes it possible for us to be harmed by words.

Nevertheless, the specific speech situation impacts on the power of the speech (Ibid.). The speech situation is not a simple kind of context, easily defined by spatial and temporal boundaries. To be injured by speech is in fact often to suffer a loss of context, that is, to not know where you are, to be put out of control, to suffer the disorientation of one's situation as the effect of injurious speech (Ibid.: 4). In this sense, hate speech in digital spaces can be viewed as putting women and those not conforming to heteronormative gender constructions into their place, whilst, often even literally, taking away their place in the digital space as they are pushed to abandon social media platforms or lock their accounts.

Useful to understanding how words can hurt are also the contributions of legal scholars such as Catherine MacKinnon (1993) and Mari Matsuda (1993), who

have theorised further on speech as illocutionary, particularly in relation to speech as not only reflecting, but also recreating inequality. As MacKinnon (1993: 13) argues, certain words are “not seen as saying anything (although they do) but as doing something”. In this way, social inequality is created and enforced - performed - through words and images: “Segregation cannot happen without someone *saying* ‘get out’ or ‘you don’t belong here’ at some point. Elevation and denigration are all accomplished through meaningful symbols and communicative acts in which saying it is doing it” (Ibid.). At the same time, a victim is traumatised by a speech act not necessarily only because of the content of the words themselves, but also because of the experiences embodied by the words (Ibid.). Words, in this regard, can be understood as carrying lived reality from one place to another. The historicity of speech needs to be considered within this framework of analysis. This is also highlighted in the work of Matsuda (1993: 36), expressing that a certain kind of speech is harmful “because it is a mechanism of subordination, reinforcing a historical vertical relationship”.

In this illocutionary model, speech can be understood as constituting its target in the moment of its utterance, meaning that it is a performance of the injury as such. Going beyond the mere reflection of social domination, speech enacts domination, reinstating a social structure, with the utterance of such speech performing an injury, understood as social subordination, building on historicity. Hate speech can thus be understood not only to act upon its listener -a perlocutionary scene- but to contribute to the social constitution of the one addressed, thus becoming part of a process of social interpellation (Butler, 1997).

Following this reasoning, what hate speech does is to constitute the subject in a subordinate position (Ibid.). Nevertheless, we must question the presumption that acts of speech - such as hate speech - always work, leaving open, “the possibility that its failure is the condition of a critical response” (Butler, 1997: 19). Introducing a more dynamic view of social structures, it is possible that “if such a structure is dependent upon its enunciation for its continuation, then it is at the site of enunciation that the question of its continuity is to be posed” (Ibid.). For example, in the case of a threat, “the threat states the impending certitude of

another, forthcoming act, but the statement itself cannot produce that forthcoming act as one of its necessary effects” (Ibid.: 12). Such speech is vulnerable to failure, and it is that vulnerability that must be countered to counter the threat.

Within this context, hate speech must be understood as a complex social, cultural and psychological phenomenon (KhosraviNik and Esposito, 2018), which does not appear out of nowhere, but needs to be addressed as part of the social and cultural order, reflecting and reinforcing values from the surrounding society. Attacks feed from the interrelation between stereotypes, prejudices and structural racism, reflecting the strategies of polarisation and social exclusion used to erect new frontiers between a perceived “us” and “them”, perpetuating and normalising a social stigma. As Perry (2001) has illustrated through her research on hate crimes, these, like hate speech, are better understood not as an aberration, but rather as a by-product of a society struggling to face changes in power relations in the context of shrinking opportunities. Hate speech, in this regard, conveys two types of parallel messages: one directed to the targeted group, aiming to dehumanise and diminish them, attacking the very core of human dignity (Waldron, 2012). The other message targets those with similar views, reinforcing a sense of a group as opposed to and threatened by “the other” and at the same time uniting this group of like-minded persons. In this sense, hate speech can be understood as dividing and uniting at the same time (Gagliardone et al., 2015). An image that can capture this quality is that of a swarm, of a connected network of individuals responding to similar stimuli and showing emerging behaviours in ways that often appear unpredictable. However, swarming behaviours can progressively turn into more coordinated efforts to act against specific targets (Gagliardone, 2019). For example, in contemporary Europe, political forces on the far right have taken advantage of sentiments of fear toward immigrants, leading to more frequent and coordinated violent attacks towards minorities, such as in the UK following the Brexit vote (Weaver, 2016). With this background in mind, further scholarly debates surrounding the topic of online hate speech will be highlighted and discussed in Chapter 3.

2.4. Conclusion: An intersectional technofeminist framework

The foundational conceptual principles of this thesis form an intersectional technofeminist framework, where gender and intersectionality serve as pervasive and interwoven themes. The interdependence between violence and gender norms, as outlined by Rita Segato (2003), and the reciprocal influence of gender and technology are essential considerations within this framework.

Technofeminism, following Judy Wajcman (2002: 356), conceives of "a two-way mutually shaping relationship between gender and technology in which technology is both a source and a consequence of gender relations and vice versa". Drawing upon an understanding of the concept of gender itself as a performance, the notion of performativity also comes into play in the mutual shaping of gender identities and technology. This means that both technology and gender can be seen as products of a moving relational process, emerging from collective and individual acts of interpretation.

Yet, the analysis cannot only consider gender; rather, it needs to be intersectional. This approach brings into the analysis an understanding of how gender interlaced with other factors such as racialisation, sexual orientation or gender identity, shapes and is shaped by technologies; in the case of this thesis, the digital context of Internet and social media. Thus, digital performances of identity need to be interpreted within a wider socio-political context, which embeds the digital mediation, analysing meanings expressed through digital spaces as negotiated at the intersection of individuals, culture, and media technologies (KhosraviNik and Esposito, 2018). This allows the analysis to go beyond an understanding of technologies as either driving changes in gender-based harms, or as mere tools used by individuals. Instead, it moves towards a more complete account of technology and gender relations as mutually shaping digital (inter)actions - shaping also the experiences of online harms, such as hate speech. Further, drawing upon theories of linguistic vulnerability and symbolic violence helps explain how technology-facilitated verbal harms may constitute "embodied harms" (Powell and Henry, 2017).

3. From Definitions to Debates

3.1 Introduction

Examinations of online violence in research draw from various disciplines and have classified such violence using a range of different labels. One of the most encompassing concepts is “technology-facilitated violence”, which has become a commonplace label in social science discourses, particularly in relation to gendered violence. However, scholars have not reached a consensus on the definition of this term or the nature of the relationship it implies between technology and violence (Mitchell et al, 2022). The same is true for many of the terms used to describe gendered online violence. As I shall expose in this chapter, when online hate speech is based on gender, this speech is often addressed as one of several subcategories to the broader category of “technology-facilitated gender-based violence” (e.g., Powell and Harris, 2017).

This chapter reviews the existing debates that have approached online hate speech, and particularly gender-based hate speech from different fields of study; most notably media and communication studies, law and criminology. In this framework, studies from social anthropology are scarce. Nevertheless, social anthropologists have investigated online hostility under other concepts and labels. Primarily, anthropological perspectives on online hate have, until recently,

concentrated on the examination of hostility manifested within distinct online cultures or subcultures. Given this, it is relevant to offer an overview of the concepts employed in discussing online hostility and to introduce key debates surrounding the manifestation of violent behaviour in digital spaces. The first section of this chapter, thus, aims to clarify the various conceptualisations of online hostility on the one hand and gendered online hostility on the other. This overview holds significance, as one of the initial challenges encountered during a literature review on the subject of online hostility, including hate speech, lies in the extensive range of concepts used to explore this phenomenon from various perspectives and disciplines. The section also covers concepts proposed as alternatives or contestations to the concept of hate speech, most notably the notion of “extreme speech”, which is a concept that stems largely from research conducted by social anthropologists.

Following the discussion of concepts, I provide an overview of studies specifically related to gender-based online hate speech, often synthesised from studies that address broader categories such as “technology-facilitated violence”, or “cyberhate”. In this overview, I briefly highlight findings with relevance to the present research. Finally, and based on this literature review, I propose a working definition of “gender-based online hate speech” for the purpose of this doctoral thesis, as well as a rationale for the research focus and the scope of the study.

3.2 An overview of concepts in relation to online hostility: From flaming to hate speech

Over the past few decades, scholars from a range of disciplines have researched and debated online hostility under concepts such as “cyber-bullying”, “cyber-stalking”, “cyber-violence”, “trolling” or “flaming” (Jane, 2015). As put forward by legal hate speech scholar Jeremy Waldron (2012: 34), “what we call a thing tells us something about our attitude toward it, why we see it as a problem, what our response to it might be, what difficulties our response might cause, and so on”. In this regard, “hate speech” is a term that is often interpreted as having strong political connotations. Analytical efforts to define what may or may not be labelled as “hate speech” have therefore tended to go hand in hand with more

practical impulses to act, contain, and punish (Gagliardone, 2019). On the other hand, terms such as “flaming” and “trolling” tend to be used for more in-depth cultural analyses. Nevertheless, oftentimes when academics —and in particular anthropologists— have conducted in-depth analyses of communities engaging in flaming or trolling, in attempts to analyse the online cultures that breed hostility, they have been accused of siding with the “haters” or perpetrators (Ibid.). That is, harmful behaviour has been linked more to specific online cultures and not so much to structural factors that act across the online and the offline.

In the following sections, I will discuss the different concepts predominantly used to address online hostility, to unveil the variety of scholarly debates in relation to this topic, from the discipline of social anthropology and beyond.

3.2.1 Flaming

Early digital scholars tended to discuss online hostility under the concept of “flaming” (e.g., Lea et al, 1992; Reinig et al, 1997; Kayani, 1998; O’Sullivan and Flanagin, 2003; Lange, 2006). This concept emerged from popular discourse surrounding online communities to describe aggressive, hostile, or profanity-laced interactions via email and discussion groups (Jane, 2015). However, whether the offensive aspect of flaming is necessarily intentional has been subject to scholarly debate. On the one hand, according to Patrick O’Sullivan and Andrew Flanagin’s definition of flaming (2003) “true flames” are “intentional (whether successful or unsuccessful) negative violations of (negotiated, evolving, and situated) interactional norms”. Peter Moor et al. (2010: 1537), on the other hand, define flaming as “displaying hostility by insulting, swearing or using otherwise offensive language”. That is, this perspective on flaming argues that some flammers intentionally offend others, but that most of what is considered flaming is simply meant to express disagreement, an alternative opinion or humorous play.

In roughly three decades of research on flaming, the academic discussion has been generated largely around the issue of how to define flaming, while the ramifications of such speech acts have often been ignored in the debate (e.g.,

Jane, 2014). In this regard, the use of flaming as a concept to analyse online hostility has received some criticism. The criticism most relevant to this thesis, as argued by social anthropologist Patricia Lange (2006), is that the term “flaming” tends to be used to refer exclusively to *online* behaviour. This has the effect of predisposing researchers to perceive flaming as a category that is inherently different from the kinds of hostility that take place offline (Ibid.). That is, using the term flaming may make researchers reinforce differences between the online and the offline, where there may be many similarities.

3.2.2 Trolling

Whilst the concept of flaming is currently being used to a very limited extent, the terms “trolling” or “hating” have become more commonly used (McCosker 2014:5), both in academic debates and by users to denominate problematic behaviour, especially in the context of social media sites. Trolling often refers to the same type of discourse as flaming; however, it tends to define the behaviour as deliberate provocation by posting inflammatory or off-topic materials in online debates, to elicit reactions. Trolling, in its original meaning, entailed luring others into pointless and time-consuming discussions (Herring, 2002). The name derives from the practice used in fishing, where a baited line is dragged behind a boat, although internet discourse tends to refer to the concept “troll” as stemming from the fictional monster waiting under the bridge to snare innocent bystanders (Ibid.).

Judith Donath (1999) was among the first scholars to discuss trolling, which he describes as an identity game where “the troll attempts to pass as a legitimate participant, sharing the group’s common interests and concerns”. According to Susan Herring (2002), trolling starts “with a message that is intentionally incorrect but not overly controversial”. Nevertheless,

when using the term troll to describe behaviour online, it often brings with it certain ideological baggage. To troll is to have negative intents, to wish harm or at least discomfort upon one’s audience. To be trolled is to be made a victim, to be caught along in the undertow and be the butt of someone else’s joke (Bergstrom, 2011:1).

Contemporary scholars, such as digital ethnographer Whitney Phillips (2015), have come to criticise the concept of trolling. This critique mainly argues that over the past few years, trolling has been turned into a vague, behavioural “catch-all” concept on the internet (Ibid.). That is, having lost its original connotation of deliberate provocation of particularly naïve internet users, it is now being used to describe almost any type of heated discussion online. With this in mind, Phillips proposes to use the term “trolling” only to describe very specific types of trolling communities, such as those that can be found in web forums such as 4chan¹⁵ or Reddit.¹⁶ Even if only this specific type of behaviour is included - sometimes described as “subcultural trolling” - trolling entails a wide spectrum of behaviours: Some trolling is aggressive and meets the legal threshold for harassment; other forms of trolling are comparatively innocuous. Some trolling is persistent, continuing for weeks or even months, and some occurs once and then never again (Ibid.). Notwithstanding, trolls of the subcultural variety tend to self-identify as such. That is, simply saying provocative things online, or engaging in racism, misogyny or LGBT-phobia does not necessarily make someone a subcultural troll. Instead, trolling in the subcultural sense should be understood as something that a self-identifying troll sets out to do, as an expression of their online identity (Phillips, 2015).

3.2.3 Online hate speech as a general concept

The excitement surrounding newly discovered communication possibilities, both among users and researchers, has often diverted academic discussions from the critical social, ethical, and political dimensions of online hostility. Thus, replacing terms like “flaming” and “trolling” with alternatives such as “cyberhate” or “hate speech” could serve as a meaningful starting point. This shift may encourage a more thoughtful approach to the violent and exclusionary nature of online

¹⁵ 4chan is a mostly image-based discussion-board, where users tend to post anonymously. The site has been linked to Internet subcultures and movements, such as Anonymous and has also been instrumental in the production and popularisation of some memes.

¹⁶ Reddit is a social media platform and online community where registered users, known as “Redditors,” can submit content, such as text posts, links, images, and videos, and participate in discussions. The platform allows users to create accounts and participate in discussions anonymously.

hostility, rather than downplaying the issue by emphasising its playful and communal aspects (KhosraviNik and Esposito, 2018: 51).

None of the concepts such as flaming, trolling or e-bile adequately address an important factor, namely the oppressive grounds, aggravating circumstances, or structural roots; in the present thesis, gender interlaced with other factors such as sexual orientation or racialisation. What is missing, thus, in these concepts, is the specific ground for the expressed hatred. Whilst online hate speech often converges with flaming, trolling and e-bile, i.e., they often attack the same factors of vulnerability or oppression, the latter concepts can be designated for any online person-to-person hostility, regardless of structural factors, whilst the structural oppressing factors or vulnerabilities need to be present in the hate expressed for a discourse to be defined as hate speech. Highlighting these elements, thus, brings important structural and interactional factors into the analysis.

Most studies on online hate speech stem from the field of Law, either debating freedom of expression online (e.g., Waldron, 2012; Phillipson, 2015; Heinze, 2016), or posing solutions for the regulation of online hate speech (e.g., Banks, 2010; Citron and Norton, 2011); cCiminology (e.g., Lewis et al, 2016); Communication and Media studies (e.g., Jane, 2014, 2015, 2017; Megarry, 2014; KhosraviniK and Esposito, 2018); Information Technologies (e.g., Burnap and Williams, 2016); or from the field of Sociology, analysing the mechanisms underlying and surrounding racist online hate speech (e.g., Daniels, 2008) or aiming to quantify online hate speech (e.g., Oksanen, 2014).

Whilst, as previously mentioned, anthropologists have tended to focus on the (sub)cultural dimensions of online hostility through concepts such as flaming and trolling, the field of anthropology has increasingly entered also the field of broader hate studies in the last few years, proposing alternative concepts such as that of “extreme speech”, as I shall discuss further below. Anthropological endeavours that address the concepts of hate speech, dangerous speech and other related concepts, have predominantly focused on speech emerging in critical moments in the political life of a national community, particularly

elections (Gagliardone, 2019). This emphasis on critical moments and incidents suggests how hate speech, particularly in its online form, has often been treated as a phenomenon that quietly and progressively permeates societies, but whose significance can be fully understood only around specific events (Ibid.).

In more specific terms regarding the state of the art in Spain, there is a predominance of legal studies on the general topic of online hate speech. In this regard, some relevant studies include Laia Serra (2018a, 2018b) on the legal context in Spain in terms of hate speech, and Víctor Gómez (2020), who has analysed jurisprudence related to online hate speech. Further, Fernando Miró Llinares (2016) has constructed a taxonomy of violent communication and hate speech on the Internet, based on an analysis of Spanish tweets following Charlie Hebdo. A few studies have analysed young people's experiences and perceptions of online hate speech (e.g., Megías, 2021; Jubany and Roiha, 2018). Overall, however, qualitative and/or ethnographic research on online hate speech, is still very scarce, and social anthropology is almost completely absent from the debate in Spain.

Further debates surrounding the online component of online hate speech

There is also a need to distinguish between hate speech expressed online and offline. Whilst both forms are based on the same structural grounds and functions, online hate discourses are amplified by the new affordances of the participatory web (KhosraviNik and Esposito 2018). In this regard, around 2006, there was an acceleration in the transformation of the static web into the interactive and user-driven sphere known as web 2.0 (Jane, 2015: 81). As McGranahan (2021) highlights,

On social media, people no longer consume news and entertainment as they did with newspapers, radio, and television. Instead, they generate media, critically interacting with others online to share information, learn, and validate ideas—in other words, to build community. Coming to know others online involves building affiliative, positive community by recognizing kindred types by, for example, hobby or profession or politics. But something about social media also encourages the darkest elements of exclusion, such that the performance of community at times includes negative abuse and violence—flaming, trolling, lying, mobbing, doxing—

often in registers deemed humorous or clever within the group and cruel or unethical outside it.

That is, whilst social media has empowered ordinary users to participate in the production and distribution of online content, the dramatic spike in user-generated texts around this time also means that an important time period started in relation to the circulation of online hate speech (Jane, 2015). The global, immediate, and participatory nature of social media communication has made the online sphere a breeding ground for the expression and dissemination of a range of exclusionary, intolerant, and extremist discourses, practices and beliefs (Kopytowska and Chilton 2018). As James Banks (2010: 234) has noted, the internet can be defined as “the new frontier for spreading hate”.

In relation to this, a common scholarly argument - especially from the fields of Social Psychology, Criminology, and Media and Communication Studies - is that social media affordances act as multipliers, both in terms of quantity and vitriolic quality of interactions (KhosraviNik and Esposito 2018). Some of these digital features and affordances have been especially highlighted as contributing to the spread of digital hate discourses. Analysing the affordances of digital media as “environments of affordances” (Madianou and Miller, 2012) allows for a perspective that analyses the particular action potentials and action-restrictions of such media and how they relate to everyday practices (Costa, 2018).

First, anonymity, or perceived anonymity, is one of the factors most widely recognised as contributing to the increase in online hostility. It is regarded to play a major role in liberating people from following social norms and conventions as they do not feel the threat of a sanction or accountability for their acts, emboldening people to be more outrageous or hateful in what they say (Wallace, 2016; Brown, 2018). Online anonymity, from this view, is related to the assumption that interactions online are intrinsically different from interactions face-to-face, i.e., less important, freer and differently valued (KhosraviNik and Esposito, 2018). The perceived anonymity of the Internet may remove fear of being held accountable for the opinions expressed and may give a sense that the normal rules of conduct do not apply (Brown, 2018; Citron, 2014; Delgado and

Stefancic, 2014). Concern about the effects of anonymity (and the use of pseudonyms) has led to something of a backlash against anonymity, particularly on social media platforms (Barendt, 2016). What was taken only a few years ago as a distinctive advantage of online communication —the freedom conferred by anonymity —is now increasingly perceived to carry a major disadvantage: the absence of accountability to those harmed by anonymous trolling (Ibid.). On the other hand, as Alexander Brown (2018) signals, the perceived anonymity of the internet may also liberate victims of online hate speech, as well as bystanders, to engage in counter-speech. That is, the fact that the hate speakers against whom counter-speakers are speaking back do not know who the counter-speakers are, could reduce the fear that counter-speakers might otherwise have about being identified and targeted.

Part of the perception that anonymity protects those spreading hate speech is due to physical separation or invisibility. Because the online hate speaker is not physically present, they do not have to worry about an immediate physical backlash (Brown, 2018). Nevertheless, the technical affordances that allow people to engage in anonymous cyberhate are not bulletproof; there is always the risk of being identified and digitally outed (Ibid.). However, the physical separation may also in itself be regarded as a factor that exacerbates violent behaviour online (Lewandowska-Tomaszczyk, 2017). The lack of a face-to face context, and the related lack of acknowledgement that the person on the receiving end is a feeling human being, may play a role in the increase of aggression online (Weisband and Atwater, 1999). Online communication often means that the immediate impacts of speech acts are unseen by the perpetrator. If one cannot see the emotional harm created by one's online hate speech, one may be more likely to downplay its significance (Brown, 2018). The lack of non-verbal cues may also have a role in this (Epley and Kruger, 2005), as those expressing hate speech online cannot see the faces of other people who might disapprove of what they are saying (Brown, 2018). As suggested by Danielle Keats Citron, "people are quicker to resort to invective when there are no social

cues, such as facial expressions, to remind them to keep their behaviour in check” (Citron, 2014: 59). However, it might be that

a person who is committed to engaging in offline, face-to-face hate speech is unlikely to be deterred by seeing the anguish on the face of the victim, and may even take pleasure in it. It could be that hate speakers’ own consumption of hate speech—such as reading articles telling them that certain others ‘aren’t really like us’ or ‘are animals’—has already undermined or created deficits in natural empathy and sympathy for the suffering of members of the out-groups (Brown, 2018: 301).

That is, a person with extreme views may have dehumanised their targets to such an extent that it does not matter whether they are anonymous and invisible or not.

Nevertheless, when discussing online anonymity, it is important to note that contemporary social media sites increasingly encourage users not only to use their real name and photo, but also to define their gender. This shift away from anonymity, sometimes denominated “radical transparency” (Kirkpatrick 2011) encourages people to have only one identity online, as today’s digital media platforms have self-disclosure affordances that induce users to identify themselves. The applications allow users to determine the type of information exposed, giving them the power to self-present (Katiambo 2021). Whilst the network of real-life friendships in certain ways limits anonymity, it may also motivate users to create an “idealized projection of the real-life ‘actual self’” (Krasnova, et al. 2009, 42). These networks of friendships, combined with other self-disclosure features, prevent users from engaging in misrepresentation, unlike in purely anonymous online forums (Katiambo, 2021). That is, the rise of social media “has resulted in an increased blurring between our online and offline lives rendering our bodies more central to our online personas” (Megarry, 2014). And wherever bodies are central, gender becomes equally central.

In contrast to the identity model used by the larger social media platforms, on platforms such as 4chan, the absence of user accounts means that markers of “offline” identity are largely absent, and everybody appears more or less the same (Tuters and Hagen, 2021). Some of these platforms also use an ephemeral

strategy for posting, meaning that posts are deleted after a few days or even minutes.

The community aspect is another factor that may make online hate speech different from that expressed offline. People who might or might not otherwise engage in offline hate speech are given pathways into cyberhate by the Internet. For example, people who lead hate groups may have a desire to grow their memberships and the internet is invaluable for the purpose of attracting new members and ensuring that their existing members feel connected (Brown, 2018; Citron, 2014). So from this perspective, hate speech is addressed, not so much to the victims, but to like-minded people (Brown, 2018). As Cass Sunstein maintains, it is “clear that the internet is playing a crucial role in permitting people who would otherwise feel isolated, or move on to something else, to band together and spread rumors, many of them paranoid and hateful” (Sunstein, 2007: 58).

Further affordances specific to the online environment and influencing the dissemination of hate speech include permanence or persistence, referring to the durability of online expressions and content (Gagliardone et al. 2015); searchability, relating to the possibility for content to be found; and itinerancy or “spreadability” (Boyd 2014), alluding to the ease with which content can be shared, meaning that even when hate speech is removed from one place, it may be revived somewhere else (Gagliardone et al. 2015).

Finally, as Brown (2018) proposes, another factor with relevance to the expression of hate speech online is instantaneousness. That is, “as compared to offline modes of communication, the Internet encourages forms of hate speech that are spontaneous in the sense of being instant responses, gut reactions, unconsidered judgments, off-the-cuff remarks, unfiltered commentary, and first thoughts” (Brown, 2018: 304). It can be argued that the internet not merely facilitates, but also encourages, instant responses that are by their nature more spontaneous (Ibid.). This does not mean that online hate speech never takes the form of careful, considered, and extensively planned statements – as is the case for example with explicit hate sites - but that some parts of the Internet encourage gut reactions, unconsidered judgments, and unfiltered commentary, as

these spaces encourage instant responses (Ibid.). All these affordances combined - anonymity, lack of physical presence, permanence, itinerancy, searchability, instantaneousness, as well as being relatively cheap and easy to use - may drive the spontaneity of some online hate speech.

3.2.4 Concepts put forward as alternatives to hate speech

Despite multiple efforts to define the contours of hate speech, in everyday discourse the term has tended to be used well beyond the boundaries set by international bodies, national bodies, and Internet companies. Hate speech is a multifaceted concept intricately connected to freedom of expression, individual, group, and minority rights, as well as the principles of dignity, liberty, and equality (Gagliardone, 2019). Labelling an expression as a "hate speech act" has the potential to amplify the reach of its speakers, particularly if they can portray themselves as "martyrs" of censorship or depict unsuccessful censorship attempts as a validation of their views (Ibid.). For these reasons, alternative, more narrowly defined concepts, such as "dangerous speech", "fear speech" or "extreme speech", have been proposed, focusing more on the propensity of expression to cause widespread violence. Below, I briefly examine these concepts, which can be seen as both subcategories and contestations to the concept of hate speech.

Dangerous speech

The concept of "dangerous speech" was coined by Susan Benesch (2013) as a subset of hate speech, based on the argument that the category of hate speech is too broad to serve as a concept of early warning of mass atrocities. This can be attributed to two main factors: Firstly, hate speech is prevalent in numerous societies, including those with a minimal risk of genocide. Secondly, certain instances of hate speech may not significantly heighten the likelihood of mass violence, although they can inflict serious emotional and psychological harm. In essence, speech has the potential to cause harm directly, indirectly, or both. On the one hand, it can directly offend, denigrate, humiliate, or instill fear in the targeted individuals. On the other hand, speech can lead to indirect harm, often with equal or even greater severity, by inspiring others to harbour negative

thoughts and take actions against members of the targeted group (Benesch, 2013). In the definition of “dangerous speech”, violence means direct physical (or bodily) harm inflicted on people, not other forms of harm such as incitement to self-harm, discrimination, or social exclusion (Benesch et al, 2018). When an act of speech has a reasonable chance of catalysing or amplifying violence by one group against another, given the circumstances in which it was made or disseminated, this speech act should be referred to as “dangerous speech” (Ibid.).

Dangerous speech may include various expressions, such as oral or written speech, traditional print and broadcast media pieces, blogs and social media posts, images, symbols, or even music and poetry (Gagliardone, 2019). Its message may be interpreted as a call to violence against a specific group or may portray this group in a way that legitimises or justifies violence against it. NGOs have notably employed the framework of dangerous speech to analyse patterns and potential outbreaks of violence during critical events, such as elections (Gagliardone, 2019). In this context, the framework of dangerous speech is contextually embedded and therefore responsive to the atmosphere in which a speech act occurs. It encourages practitioners, researchers, and policymakers to consider various factors, including the influence of speakers, the receptiveness of the audience, the content of the speech, the medium of dissemination (including language and the platform for distribution, which may serve as the primary source of news for the relevant audience), and the social or historical context of the speech (Benesch, 2012).

Fear speech

The term "fear speech" (Buyse, 2014; George, 2015) has been introduced to highlight speech acts that could instill fear in a particular group, suggesting that "the other group" intends to use violence or even annihilate them in the near future (Gagliardone, 2019). Consequently, the concept of fear speech provides a way to assess whether conditions for violence might be developing gradually, as the cultivation of intense anxiety serves as a more accurate predictor for the

escalation of violence compared to adopting broader analytical frameworks, such as that of hate speech. (Ibid.).

Antoine Buyse (2014) outlines various factors that can escalate violence triggered by speech acts, where both the message's content and context play crucial roles. Significantly, a key predictor involves examining how past instances of potential violence between two groups have been interpreted or justified (Ibid.). For instance, whether a violent incident is explained as an isolated crime or part of a broader threat of violence from one group to another; whether blame is assigned to a group or an individual, and if this involves the use of stereotypes. The more dehumanising the stereotypes employed, the lower the threshold for resorting to violence. This language can progressively foster a siege mentality, eventually legitimising violent acts as a defense of a group's safety or integrity (Gagliardone, 2019).

Extreme speech

A main critique against the concept of hate speech stems from within anthropology, where in recent years a growing body of literature has been produced in relation to so called “extreme speech”, most notably through work by media anthropologists such as Sahana Udupa and Matti Pohjonen (Pohjonen and Udupa 2017; Udupa et al 2018; Udupa and Pohjonen 2019; Udupa et al 2021). As a concept, “extreme speech” aims to be an anthropological critique of the hate speech discourse, going away from binary classifications between “hate speech” and “acceptable speech”. Instead, the concept extreme speech emphasises the need to contextualise online debate with an attention to user practices and particular histories of speech cultures, analysing online vitriol in situated contexts (Udupa et al, 2018). According to Udupa et al (2018), extreme speech can be defined as speech acts that push the boundaries of acceptable speech along the twin axes of civility/incivility and truth/falsity, and as a concept, extreme speech is open-ended to acknowledging its potential both as a resource for agonistic politics as well as dominance via exclusionary logics. Such a perspective can help historicising online vitriol, rather than seeing the actors involved in this type of behaviour as antagonists who have suddenly appeared on

the scene. That is, this perspective implies a methodological and ethical stance that seeks to understand people who participate in online abusive exchange, avoiding a tendency where critique precedes knowledge, or where a moral-evaluative framework predetermines what to expect (Udupa et al, 2018). Based on this perspective, hate speech can be perceived as a metadiscursive term used in specific contexts by particular categories of speakers for the purpose of labelling and evaluating speech, indicating speakers' identification with (and against) other sociopolitical categories of speakers. That is, from this view, the hate speech concept can be used for promoting moral and political agendas and undermining those of others as (Boromisza-Hamashi, 2021).

The extreme speech framework has been mainly applied in ethnographic studies of extreme speech in minority-majority contexts, such as that by Max Kramer (2021) in relation to Indian Muslims; to show how different regimes label oppositional discourse as hate speech, such as the study on the pre-election context of social media in Kenya by David Katiambo (2021) or the analysis by David Boromisza-Habashi (2021) on the various uses of the hate speech concept in public discourse in Hungary and the competing systems of meanings that inform those uses and understandings. However, it has also been applied to analyse online vitriol expressed in more subcultural contexts, such as the use of memes on 4chan (Tuters and Hagen, 2021) or fun as a metapractice among online users in ideological battles in India (Udupa, 2021).

The contributions by the extreme speech framework are crucial for bringing in the historic and situated context of violent speech. It also helps breaking the binary division between "hate speech" and "acceptable speech", showing the dynamic nature of online vitriol. Nevertheless, the strong focus of extreme speech studies on contexts of violent speech against minorities in specific situated contexts, as well as the lack of extreme speech studies from a gender or intersectional perspective, does not make this concept ideal for the topic under study in this doctoral thesis. Instead, as reflected throughout, I have chosen to evaluate whether the online gendered violence that targets feminists in Spain can be categorised and analysed as a kind of hate speech, in terms of its functions,

impacts and underlying structures, but also in terms of legal tools and support. In the next sections, I shall review the literature in relation to gender-based online hate speech and other types of gendered online violence.

3.3. Gender-based online hate speech

Having addressed the concept of online hate speech in general, as well as other related concepts, this section aims to zoom in on the gendered aspects of online hostility, violence and hate speech. Neither gendered hostility online, nor the study of this, are new phenomena. Early research in the field of computer-mediated communication was based on the notion that Internet communication was anonymous (Herring, 2001) and that the structure of the internet would make gender irrelevant to online interactions (Balka, 1993; Graddol and Swann, 1989). In the early 1990s, however, as the presence of women increased online, studies of gender in relation to online communication started appearing (Herring, 2001). The findings of these studies started problematising claims of gender-neutral equality on the web. For example, Cynthia Selfe and Paul Meyer (1991) found that men and in general participants who enjoyed high status offline also dominated the interactions online, both under normal conditions and under conditions of anonymity. Further, in the early 1990s scholars also began reporting the use of more aggressive behaviour by men in online discussions, with some of this behaviour explicitly targeted at female participants (Herring 1993; Kramarae & Taylor, 1993).

Whilst Susan Herring, from the field of Information Science, has examined gendered dimensions in online interactions (2004, 2010; 2011); feminist responses to 'trolling' (2002); and gender differences in online posting behaviour (Herring 1996a; 1996b), she was also one of the first scholars to recognise the gendered nature of much online abuse. For example, Herring (1993) found that women were more likely than men to react aversively to aggression in online interaction, including falling silent and dropping out of groups. Other early studies found that women introduced fewer topics of discussion and received fewer public responses than men online (Herring 1993; Herring, Johnson and DiBenedetto, 1995).

Contemporary studies of gendered power relations on social media have often argued that the absence of traditional gatekeepers makes social media a space where historically marginalised groups, such as women, could be presumed to exert greater social and political influence (Megarry, 2014). However, several authors have, over the last two decades analysed the androcentrism of digital communities, and how these informally exclude women by creating an unwelcoming climate (e.g., Kendall, 2002; Martin, Vaccaro, Heckert, & Heasley, 2015; Reagle, 2012;). So while women currently outnumber men on many social media platforms, this is not necessarily indicative of their ability to influence online public spheres (Megarry 2014). In this regard, Julian Aussenhofer and Axel Maireder (2013) found that women were much underrepresented in the Austrian political Twittersphere and that Twitter users, regardless of gender, addressed men more often than they addressed women.

In the next few sections, I will focus on the conceptualisation of online gender-based violence and hate speech, as well as a review of findings from contemporary studies divided by different themes, including different typologies of digital gender-based violence; prevalence and targets; perpetrators and social underpinnings; content and discourses; experiences and life impacts; responses and contestations; and intersectional aspects.

3.3.1. Nomenclatures related to online misogyny

In examining contemporary studies on gender-based online hate speech, one of the first obstacles that appears in gaining an adequate overview of this research field is the lack of consensus on how to name and conceptualise this phenomenon. Although misogynistic violence represents an important social problem, both institutional and scholarly research have often dismissed the recognition of online misogyny as a form of gender-based hate speech (KhosraviNik and Esposito, 2018). As scholars such as Jessica Megarry (2014) argue, social media remains grounded in the material realities of women's everyday experiences of sexism. That is, it is important to explore online abuse against women as an extension of offline gender relations, as "it seems that gendered epithets are dismissed as rowdy protests or robust criticism, or simply

par for the course, rather than being recognised as sex discrimination or gender hatred” (Weston-Scheuber, 2012).

The absence of gender from most institutional definitions of hate speech highlights the institutional unwillingness to acknowledge gender as a social factor that, per se, can trigger hate (Titley, 2012; KhosraviNik and Esposito, 2018). There are several factors that potentially influence this: First, the widespread assumption that gender equality has been largely achieved may account for the reluctance to discuss sexist speech in terms of hate speech (Lillian, 2007); second, while hate speech targets “members of vulnerable minorities” (Waldron, 2012: 5) with racial and xenophobic aggressions aimed at their elimination from a given community, the same cannot be argued for women (Lillian, 2007). However, these assumptions fail to account for the numerous ways in which women are targeted by sexist and sexually violent speech in what can be regarded as one of the “most complex and pervasive system[s] of oppression” (Lazar, 2007: 143). A conceptual framework of online abuse towards feminists as violence against women and at the same time as a form of hate crime has previously been proposed by Ruth Lewis et al (2019), as well as by Noelia Igareda (2019), more specifically in terms of legal debates.

Further concepts that have been proposed to cover verbal gender-based abuse online are “online misogyny”, “gendered cyberhate” or “e-bile” (Jane 2016; 2017), “gender hate” as proposed by Debbie Ging and Eugenia Siapera (2019) in the context of anti-feminism, or “gender-trolling” (Mantilla 2015). E-bile (Jane, 2015: 66) has been defined as a concept to “gather under one heading a variety of ostensibly variegated denunciatory speech acts that share characteristics and signal features, and so demand a broad field of inquiry”. At the same time, it points to the need for a specific terminology to acknowledge the intricacy of the actions being analysed, as well as the fluid and evolving nature of these practices and the technologies upon which they depend. (Jane, 2017b: 46). Whilst Emma Jane has predominantly studied gendered e-bile, this concept is aimed to be wider, going beyond solely gendered forms of denunciatory speech acts. Gender-trolling, on the other hand, has been proposed by Karla Mantilla (2015) as a sub-

category of trolling. It describes a different pattern of harassing, abusive, and threatening behaviours than generic trolling - a pattern of trolling specifically targeted at women. This is a separate and distinct category from other types of trolling, which arises out of different motivations. The so called “gendertrolls” aim to inspire abject fear in their targets, to drive the target out of public discourse online, and attacks are often sustained over long periods of time (Ibid.). That is, this concept lies very close to what this thesis understands as gender-based online hate speech. Nevertheless, neither “e-bile” nor “gendertrolling” has gained much traction in scholarly debate.

3.3.2. Review of studies of gender-based online violence and hate speech

Over the recent years, a few legal studies have focused explicitly on the topic of “gender hate speech” (Souto Galván, 2015), “sexist hate speech” (Álvarez, 2019) or “anti-gender hate speech” (Igareda, 2022) in the context of Spain. Further, Gómez (2016) has analysed the incorporation of the grounds of gender following the 2015 reform of the penal code in relation to hate crimes and hate speech. With relevance in the field of media and communication studies, a few studies have been based on big data analyses to explore hate speech related to gender and sexual identities on Twitter (Arcila-Calderón et al, 2021; Arce-García and Menéndez-Menéndez, 2022). Further, Teresa Piñeiro-Otero and Xabier Martínez-Rolán (2021) have conducted a quantitative study on hate speech and misogyny in the Twitter conversations surrounding Spanish women with a high profile, whilst Cilia Willem, Lucas Platero and Iolanda Tortajada (2023) have analysed the discourse of anti-trans or transexclusionary feminists on social media. Further, in the last few years, several scholars have focused on hate speech against LGBTQI+ people on social media platforms in Spain (e.g., Rivera-Martín, 2022; Martínez-Valerio, 2021).

For research extending beyond an exclusively legal domain or beyond quantitative or big data analyses, especially in the context of Spain, it is essential to also delve into the broader subject of gender-based online violence. It is crucial to note that the literature on gender-based online violence extends beyond academic studies; it also includes grey literature generated by various

institutions and organizations. When analysing this literature, it is important to emphasize that grey literature often adopts a perspective that addresses all forms of online gender-based violence, potentially overlooking the distinctive nature of hate speech. This makes it difficult to limit the literature-review exclusively to the phenomenon of online gender-based hate speech, as aspects of this phenomenon tend to be studied in an integrated way with other types of gendered online violence. Therefore, some of the sections below, such as prevalence, speaks more about the broader phenomenon of online violence against women. It should also be kept in mind that studies published by international organizations tend to focus on quantitative information, with the aim to elucidate the pervasiveness of the phenomenon and ultimately awaken political action, whilst lacking theoretical frameworks set to examine the issue from a more in-depth angle.

To highlight what types or categories of violence that tend to be included under the broader umbrella of gender-based online violence or “digital violence”, the next section briefly discusses different typologies of violence included under the umbrella terms “gender-based online violence”, “cyber violence against women” or “technology-facilitated violence” as proposed by grey literature and previous scholarly studies.

Categories of gender-based online violence

As previously mentioned, most studies englobe a great variety of gender-based online violence and not only that which can be defined as “hate speech”. Nevertheless, one type of violence is often accompanied by other types, also in the case of hate speech. That is, the limitations must be left somewhat dynamic and flexible for the purpose of fully understanding and analysing the types of violence that tend to target feminist or activist women online. For this reason, this section presents an overview of the different typologies or categories of violence included in the literature reviewed.

First, different organizations and institutions have suggested typologies mainly from a perspective of covering gendered digital violence as a broad spectrum for policy-making purposes and to raise awareness. In this regard, one of the most

exhaustive international typologies is included in the 2015 definition of online gender-based violence established by the Internet Governance Forum, which includes the following categories of violence (with the numerous subcategories reflected in the footnotes): Infringement of privacy¹⁷; Surveillance and monitoring¹⁸; Damaging reputation and/or credibility¹⁹; Harassment (which may be accompanied by offline harassment)²⁰; Direct threats and/or violence²¹; and Targeted attacks to communities²². Further in this regard, the UN Broadband

¹⁷ Including accessing, using, manipulating and/or disseminating private data without consent (by cracking personal accounts, stealing passwords, using/ stealing identities, using another person's computer to access a user's accounts while it is logged in, etc. The term 'cracking' is used rather than 'hacking' to indicate a forced entry or takeover of content with malicious intent, while 'hacking' could include similar actions that are bona fide and/or done in the public interest. See e.g. Internet Governance Forum, 2015: p. 22.); taking, accessing, using, manipulating, and/or disseminating photographs and/or videos without consent (including 'revenge pornography'); sharing and/or disseminating private information and/or content, including (sexualised) images, audio clips and/or video clips, without knowledge or consent; doxing (researching and broadcasting personally identifiable information about an individual without consent, sometimes with the intention of providing access to the woman in the 'real' world for harassment and/or other purposes); contacting and/or harassing a user's children, extended family, colleagues (etc) to gain access to her.

¹⁸ Including monitoring, tracking and/or surveillance of online and offline activities; using spyware or keyboard loggers without a user's consent; using GPS or other geolocator software to track a woman's movements without consent; stalking.

¹⁹ Including deleting, sending and/or manipulating emails and/or content without consent; creating and sharing false personal data with the intention of damaging a user's reputation; manipulating and/or creating fake photographs and/or videos; identity theft; disseminating private information for the purpose of damaging someone's reputation; making offensive, disparaging and/or false online comments and/or postings that are intended to tarnish a person's reputation, including libel/ defamation.

²⁰ Including "cyber bullying" and/or repeated harassment through unwanted messages, attention and/or contact; direct threats of violence, including threats of sexual and/or physical violence; abusive comments; unsolicited sending and/or receiving of sexually explicit materials; incitement to physical violence; hate speech, social media posts and/or mail; often targeted at gender and/or sexuality; online content that portray women as sexual objects; use of sexist and/or gendered comments or name-calling; use of indecent or violent images to demean women; abusing and/or shaming a woman for expressing views that are not normative, for disagreeing with people (often men) and also for refusing sexual advances; counselling suicide or advocating femicide; mobbing, including the selection of a target for bullying or harassment mobbing by a group of people rather than an individual and as a practice specifically facilitated by technology.

²¹ Including trafficking of women through the use of technology, including use of technology for victim selection and preparation; sexualised blackmail and/or extortion; theft of identity, money and/or property; impersonation resulting in physical attack.

²² Including cracking websites, social media and/or email accounts of organizations and communities with malicious intent; surveillance and monitoring of activities by members in the community; direct threats of violence to community members; mobbing, specifically when selecting a target for bullying or harassment by a group of people, rather than an individual, and

Commission for Digital Development Working Group on Broadband and Gender 2015 report defines “cyber violence against women” as including: “hate speech (publishing a blasphemous libel), hacking (intercepting private communications), identity theft, online stalking (criminal harassment) and uttering threats. It can entail convincing a target to end their lives (counselling suicide or advocating genocide). Looking at the Latin American context, the Mexican association Luchadoras (2018) has put forward the following typology of aggressions against women through ICT: Unauthorised access (tapping) and monitoring access; Control and manipulation of information; Spoofing and identity theft; Monitoring and cyberstalking; Discriminatory statements; Harassment; Threats; Dissemination of personal or intimate information without consent; Blackmail; Discrediting; Sexual abuse and exploitation associated with technology; Attacks on women’s channels of expression; and Oppression by those with regulatory power.

In the Spanish context, a policy brief by the Spanish National Observatory of Technology and Society (Ontsi) (2023) puts forward a typology and categorization based on the Council of Europe’s (2018) definition of digital violence, which includes: Cyber harassment, including threats of violence, coercion, insults or threats, non-consensual dissemination of sexually explicit images; Direct threats or physical violence related to digital technologies; Gender-based Hate crimes related to digital technologies; Privacy violations related to digitization and the Internet, including doxing and identity theft; and Online sexual exploitation.

Reviewing scholarly work, Anastasia Powell and Nicola Henry (2017: 6), analyse what they refer to as *technology-facilitated sexual violence*. Within this type of violence, they propose the following broad categories: enabling rape and/or sexual assault or another unwanted sexual experience; image-based sexual abuse (including “revenge pornography” and “sextortion”); and online sexual harassment (including sexual solicitation, image-based harassment, gender-

as a practice specifically facilitated by technology; disclosure of anonymised information like address of shelters, etc.

based hate speech and rape threats). Further, based on a review of 33 international articles on gender-based online violence, Nuria Vergés Bosch and Ariadna Gil-Suarez (2021), categorise this violence in three broad categories: first, violence with highly sexist discourse and that based on sex or gender try to depreciate women, LGBTIQ people and any advances that they have managed; second, highly technological gender-based violence that aim to magnify its effects through the use of technology; and third, violence with a large sexual component that aims to cause harm to and control women's bodies and sexuality.

Focusing specifically on Spain, a report by Noelia Igareda et al (2019) commissioned by the Institut Català de la Dona to analyse "gender-based cyber violence" problematises the definition itself of gender-based violence in Spain, which in the legal definition exclusively refers to violence by a (heteronormative) partner or ex-partner.²³ Due to these legal limitations, Igareda et al distinguish between, on the one hand, sexist violence ("violencia machista") that is legally categorised as gender violence in Spain. That is, where there exists or has existed a relationship between victim and offender. And, on the other hand, forms of criminal violence that tend to be gender-based but that are not categorised as gender violence in Spanish law (due to the absence of a relationship between victim and offender). In this regard, Igareda et al include the following types of online violence as "gender-based cyber violence": Discriminatory or derogatory expressions or insults for being a woman; Access to accounts and devices without consent and / or manipulation of private data such as passwords, accounts, profiles; Threat through digital channels; Contact and harassment targeting a person and/or their family or friends; Tracking and monitoring movements; Gaslighting; Discredit or defamation on the internet; Blackmail through digital channels; Internet publication of personal information with the intention that other people will find, attack or harass you; Identity theft through e-mail accounts, social media profiles; Censorship or silencing by deleting profiles or

²³ As I shall expose in Chapter 5 of this thesis, this limitation of the Spanish framework has important implications both for the conceptualisation of this type of violence among professionals, and for the formal support for those women who are targeted.

content, server attacks (DDOS), blocking web pages or domains; Sexual blackmail (sextortion) through digital channels; Induction to femicide/suicide²⁴; Dissemination of images, videos or information of a sexual character without consent; Deception for the purpose of sexual exploitation; Sexual harassment; Mansplaining, contempt, condescension.

As can be observed in these typologies, most, but not all, include 'hate speech' as a specific category. However, even if they do not include this category, they still include elements that can be defined as aspects of online hate speech, such as threats and discriminatory statements, for example, in the case of the typology of Luchadoras. Further, attacks on women online tend not to include *only* verbal elements, but also other elements such as doxing, cracking websites, direct threats of violence, and so on. That is, the dimensions of violence overlap and do not necessarily signal discrete categories.

An examination of these typologies illuminates the complexity of limiting the scope of violence. Even if this study were to establish a highly defined boundary for what it addresses; given its foundation in ethnographic fieldwork, the narratives of those targeted will inevitably shape the analysis of the violence they have experienced, whether categorised as "hate speech" or within the broader framework of "online gender-based violence". As previously stated, I have chosen the label "hate speech", to construct a framework that is able to shed light on how the structures that underpin this violence are similar to those underlying other types of hate crimes, namely structures that aim to uphold (hetero)normative power structures, thus punishing and "putting in their place" those persons who do not stay neatly within these norms or who try to counter the norms.

Further, the extensive list of online violence included in some of these typologies clearly goes far beyond what is considered violence by the law. This is the case, for example, with gaslighting or mansplaining, in Igareda's typology. It reminds us that what is experienced as violence is highly subjective and that, as ethnographers, we therefore cannot adhere exclusively to the legal definition of

²⁴ That is, invites to commit suicide or repeating that a person should be dead.

violence, but that we need to base the definition and understanding of gendered online violence on the perceptions and experiences of those persons who are targeted by this violence. This is where ethnography holds a crucial role, to offer a situated and grounded perspective on what is *lived* as violence online; how this violence impacts on women's lives; and how it is contested. This means that this thesis applies a perspective on violence that is grounded in the experiences and observations collected in the ethnographic fieldwork. Even though I label the violence under study "hate speech", I also consider other types of violence that accompany this verbal abuse and that possibly have the same aim of silencing subaltern voices and identities, in digital spaces and beyond, to the extent that these types of violence are reflected in the fieldwork.

In the next sections, I will summarise some of the findings extracted from the literature reviewed, with focus on the elements that have been of relevance to limiting the scope and formulating research objectives and questions of this doctoral research; however, with incorporation also of relevant studies published throughout this research. It should be noted that the great majority of these studies are of a quantitative nature.

Prevalence and targets of online gender-based violence

Reviewing international studies on online gender-based hate speech or online misogyny, there is agreement that women with a public presence are especially targeted, including journalists, politicians and human rights defenders (Igareda, 2022; Jane, 2017; Internet Governance Forum, 2015); but targeted groups also include academics (Beard, 2013); women who are active in technology industries (Internet Governance Forum, 2015) and gamers (Burgess et al 2017; Jane 2016). Another category highlighted by scholars are women who engage in feminist debate online (Lewis et al, 2016; 2019; Crosas Remon and Medina-Bravo, 2019), which is the focus of this thesis. In light of this, Piñeiro-Otero and Martínez-Roldán (2021) have shown how women in communication and especially politics experience a particularly toxic environment on Twitter. More than one in ten tweets directed to or about high-profile women in their sample included insults or derogatory terms.

In terms of intersectional perspectives on prevalence and targets, legal scholar Danielle Keats Citron (2014), highlights that in the US non-white women face cyber harassment more than any other group, with 53% reporting having been harassed online, compared to 45% among white women, concluding that “what is beyond dispute is that being a woman raises one’s risk of cyber harassment, and for lesbian, transgender, or bisexual women and women of colour, the risk may be higher” (2014: 14). Outside of academia, Amnesty International (2018) have shown, through a quantitative crowd-sourced study, how racialised women are disproportionately targeted by online abuse on Twitter.²⁵

In the case of Spain, a policy brief by the Spanish National Observatory of Technology and Society (Ontsi) includes results from a 2019 survey on issues of violence against women. Based on approximately 10,000 women, it includes questions about online sexual harassment, allowing for an approximate idea of the incidence of the broader problem of online gender-based violence. Findings indicate that 7.4% of women aged 16 or older have received inappropriate, humiliating, intimidating, or offensive insinuations through social media. Over 15% had experienced repeated harassment from the same person, with almost a quarter having to endure offensive or embarrassing comments or improper propositions online. The policy brief underscores that younger women are more prone to digital harassment, with over 25% of women aged 16 to 25 experiencing inappropriate insinuations on social media.

Further, the association Calala has produced a report, together with Hybridas and Komons, on online gender-based violence against activists in Spain (2020). The scope of the predominantly quantitative research underpinning this report is fairly similar to the scope of the present thesis, as it concerns violence against activists. This report argues that the violence affects to a higher degree activists with visible online information, such as name, photo or work, and those who represent non-normative identities; and the further from a heteronormative

²⁵ Based on an analysis of tweets sent to women journalists and politicians in the US and the UK in 2017, 7.1% of the tweets sent to these women were labelled "problematic" or "abusive". Among these women, Black women were the most targeted; they were 84% more likely to be mentioned in abusive or problematic tweets than white women.

identity, the more violence is received. In terms of platforms, the majority (72.7%) mention Twitter as the channel where the violence had taken place. Igareda et al (2019), on the other hand, showcase findings from a survey implemented by the association DonesTech. Again, considering the relatively small sample of 262 respondents, conclusions should be drawn carefully; however, considering the general lack of research on the topic of online gender-based violence in Spain, it is still of interest to briefly review some findings from this study. Of the respondents who reported that they had suffered online violence, 70.0% had been targeted by more than one type of online violence, with the most prevalent type being insults or discriminatory or derogatory expressions for being a woman (54.6%). That is, behaviour that the present thesis would categorise as gender-based hate speech. Further, this study found that participants with public relevance to a higher extent received insults or discriminatory or denigrating expressions for being a woman (64.3%), followed by digital threats (46.4%), harassment towards themselves or people around them (29.8%), and discredit or defamation (29.8%). Regarding the frequency of being targeted by online violence, for 55.2% the violence had taken place on a one-off basis, whilst on the other end of the spectrum, 6.2% had been exposed daily. In terms of durations, the violence had lasted at least a week in 72.1% of the cases; for 27% it had lasted one day, whilst 10.4% of the respondents had suffered violence for more than 3 years.

Perpetrators and the manosphere

Calala (2020) highlights the existence of misogynist communities on social media in Spain, including pages and groups on Facebook, often with a paramilitary culture. As noted by Igareda (2022: 101), anti-gender activists are extremely active online, taking advantage of the opportunities offered by new technologies. However, not all perpetrators stem from organised movements. In Calala's study on online gender-based violence against activists in Spain (2020), in most of the cases (76%), victims perceived the attackers as acting individually and not collectively, and that the attackers were anonymous (63,6%). Nevertheless, in approximately one fifth of the cases, the attackers were identified as belonging to a certain community including political parties; sectors of abolitionist or

transexcluding feminists; groups advocating for shared custody; groups that deny gender-based violence; and organised far-right groups and parties. Similar to this, Diana Morena-Balaguer et al (2021) identify five different types of perpetrator accounts related online violence against feminist activists in Spain, including anonymous radical far right profiles, bots, harassers, transphobic feminists and transexcluding anonymous women.

“Triggers”, content, discourses

Sarah Sobieraj (2018) suggests categorising the strategies of aggressors in three overlapping strategies: intimidating, shaming, and discrediting, arguing that the common function of these is to silence women or to limit their impact in digital spaces. Exploring these strategies, “intimidation appears in threats of physical violence, such as death and rape threats, as well as in intimations that the attacker knows where the person lives or works, and vague but frightening missives indicating that the target should shut up or their families could be affected” (2018: 5). In the context of Spain, different discursive strategies have been signalled in cyberviolence against activists and feminists, including insults, sarcasm, imposition, desire to harm, sexual objectification, criminalisation and defamation, general threat, and sexual threat (Crosas and Medina-Bravo, 2018), as well as assessments and insults based on binarism, assignment of gender roles, sexualisation and objectification of women; use of sexual terror, threat of sexual assault and rape; discredit of opinions; evaluations of physical appearance (Calala, 2020).

In terms of “triggers” or trigger topics for violence expressed against activists and feminists, research highlights several topics discussed by women as unleashing attacks, including feminist messages in general; opinions about geopolitical tensions or politics in general; commenting on or calling out sexist violence; comments about racism in Spain; criticisms against persons, entities or parties that are considered to be against women’s rights (Calala, 2020); as well as tweets related to transfeminism (Morena-Balaguer et al, 2021). The latter is signalled as a key trigger for massive and coordinated attacks, both from the far right and by transphobic profiles (Ibid.)

Experiences and life-impacts

International research points towards both online and offline repercussions of gender-based online hate speech. For example, the impact on an online-offline continuum can be showcased looking at the 2014 GamerGate case, which has been subject to several instances of scholarly analysis (e.g., Burgess et al 2017; Jane, 2016). In this case, several feminist gamers received significant harassment and abuse, including death-, rape- and bomb threats, after speaking up about misogyny in video games. As the most extreme element, one gamer received a threat to disrupt one of her speaking engagements with a mass shooting, which ultimately lead her to cancelling this event. That is, what had first been limited to the online environment quickly evolved into rape and death threats; and at least two of the victims fled their homes (Moloney and Love, 2017). These studies, thus, highlight the impact that this violence had, not only on online interactions but on a continuum of the online and the offline. They also explore the backlash that feminists are subject to in online environments.

Previous studies further point towards the long-term detrimental impact of online abuse on the well-being of targeted women, often leading to feelings of anxiety, sadness, vulnerability, and terror, with targets suffering on a social, psychological, professional, financial, and political level (Jane, 2017). Looking at evidence from Spain, Calala (2020) detects a considerable impact on physical and mental health,²⁶ with the most common feeling after being targeted by online violence being frustration and helplessness. More than half of the respondents had self-censored or had to lower their profile online. In the Igareda et al (2019) survey, 76.1% of the respondents who had been targeted by online gender-based violence reported that they had suffered psychological harm, whilst in more than half of the cases the violence had harmed their participation in public life in terms of self-censorship, withdrawing from social media, and withdrawing from participating in public acts in general. Further, respondents had suffered physical

²⁶ In their survey, anxiety appears in 40.8% of the cases, sadness and depression in 35.9% and fear in 30.4%. A third of respondents had suffered from sleep alterations, and 16.9% had suffered from physical symptoms associated with stress and depression, such as hair loss, skin problems, or headaches.

health consequences, such as headaches or nausea (26.3%); or economic harm, including income loss, difficulty of getting work due to loss in prestige, and economic loss due to judicial costs such as fees to lawyers (10.8%). Also international studies have signalled online harassment in terms of having a continuous economic impact, whilst also limiting the potential for further employment opportunities (Citron, 2009; Megarry, 2014).

Further, in the case of attacks against women politicians, this violence has been shown to have consequences for the democratic process, as women may hesitate to participate in elections, or may disconnect from political debate online (Igareda, 2022). In this regard, Igareda (Ibid.) highlights how online harassment thus works as a correction mechanism for female politicians who dare to occupy a space traditionally reserved for men (Ibid.).

Responses and contestations

Some studies have also focused on forms of response and contestation to gendered online violence, as even though digital spaces reproduce masculine cultures and power hierarchies, these spaces can also enable women to engage in new forms of contestation (Sassen, 2002). From an international perspective, most notably the many contributions to the debate by Jane (e.g., 2017) highlight how women push back in different ways, including a more widespread feminist rhetoric and increased individual and collective feminist activism. The latter may take the shape of “feminist vigilante responses” to gendered cyberhate, often expressed as “naming and shaming” of perpetrators (Ibid.). Further, Jasmine Linabary and Bianca Batti (2019) have analysed women’s public narratives of resistance to gendered online violence.

The Spanish survey studies referenced in previous sections highlight some immediate responses by targets. In this regard, Diana Morena-Balaguer et al (2021) argue that targets react in very distinct manners, with responses stretching from ignoring the attack to publicly denouncing the attack, and also, in lines with Jane’s evidence, applying strategies of “naming and shaming”.

Intersectional perspectives

Hate crime scholarship has only just recently begun to analyse how various identities interact and are read by victims and perpetrators (Burnap and Williams, 2016). Intersectionality frameworks have also to some extent been applied to examine hate crime policies and perpetrators (e.g., McPhail, 2002; Perry, 2001; Strolovitch, 2007). Further, Doug Meyer (2010) has analysed intersectional differences between LGBT hate crime victims, suggesting that the social position of LGBT people plays an instrumental role in structuring how they evaluate the severity of hate-motivated violence.

Studies on islamophobic hate crimes stand out as the field where an intersectional perspective is most notably present, with increased attention paid to the gendered aspect of this phenomenon over the last few years. As Islamophobia is often triggered by visible Muslim identities, studies have underlined its strongly gendered dimension, often involving physical aggression against headscarf-wearing Muslim women (Allen, 2015). However, it is not clear whether Muslim women are more targeted than Muslim men by online hate speech. Giulia Evolvi (2018), based on a rather limited study on Twitter, concludes that online islamophobic attacks do not specifically target Muslim women, but that rather, anti-Islam tweets tend to advocate for the protection of Western women against Muslim men. Other studies, however, indicate that the majority of reported victims of online Islamophobia are women (Feldman and Littler, 2014); one of the reasons suggested for this is that women may be more likely to report online abuse (Gerard and Whitfield, 2016).

To the extent that intersectional perspectives are included in research on gender-based online violence and hate speech, or online misogyny, existing scholarship indicates that women are attacked on multiple elements of their actual or perceived traits. For example, Bridget Harris and Laura Vitis (2020) have discussed how bisexual, lesbian and transgender women report threats of sexualised violence. These threats can be interpreted as “corrective” to a perceived transgression from traditional gender roles and heteronormative sexuality (Jane, 2014). Other studies have analysed how Black women,

Indigenous women, and, in general, racialised women receive abuse and harassment that is at the same time racialised and sexualised (Gray, 2012; Madden et al, 2018; Glitch, 2023). Moyla Bailey and Trudy (2018) have referred to the intersection of misogyny and racism that black women suffer as “misogynoir”, with a report by the British NGO Glitch (2023) highlighting “digital misogynoir” as the violent dehumanisation of Black women on social media, as well as through other digital forms, such as algorithmic discrimination. Further, as related above in relation to hate crime studies, those who practice a religion, particularly visibly Muslim women, may find their faith is attacked, alongside their sex or gender (Awan and Zempi, 2015). Finally, women with cognitive, intellectual or physical disabilities or mental illnesses have spoken of being sent aggressive messages where their bodies, perceived capabilities and femininity are critiqued (Hackworth, 2018; Harris and Vitis, 2020).

Spanish quantitative studies related to digital violence have in the last few years increasingly included intersectional considerations, showing how racialised and LGBT+ users are exposed not only to misogynist abuse, but that this abuse tends to be coupled with racist and/or LGBT-phobic insults (e.g., Calala, 2020; Morena-Balaguer et al, 2021). In terms of online feminist activism in Spain, a couple of studies have included intersectional considerations in their analysis of feminist practices on social media (Caballero-Gálvez et al, 2022; Willem and Tortajada, 2021); whilst others have analysed Roma feminism on social media from an intersectional perspective (Marques Gonçalves and Willem, 2021).

In most of the existing studies referenced in previous sections of this chapter, however, intersectionality is only addressed in passing. For example, Lewis et al (2016), briefly refer to intersectionality, stating that “for some of the respondents, the sexism or misogyny in the online abuse intersected with other forms of oppression, such as racism or homophobia”. However, this is left without further analysis. Jane (2017) explicitly states that her examination of gendered cyberhate does not negate or diminish the significance of online hate speech targeting racism, homophobia, or transphobia. While acknowledging the undeniable political intersectionality of gender with other social identities, she

clarifies that her current research does not delve into cyberhate concerning issues of race, class, sexual orientation, and gender in any greater detail.

As argued by Lewis et al (2019), precisely the intersectionality of prejudice often expressed in hate speech against women makes the debate on whether to interpret it within a hate crime framework even more complex. According to these authors, a victim-focused perspective needs to acknowledge that these crimes or incidents are experienced in a wider social, cultural and personal context and may be related to multiple forms of oppression. However, historically, the law has struggled to respond to this intersectional dimension of hate and abuse (Ibid.). Racist or LGBT-phobic hate is exacerbated by a combination with misogyny, but in ways that are unpredictable and mediated by context (Ibid.). The misogyny expressed in online abuse in their study – and in the present research – may as a matter of fact not be extended to all women, but may encompass primarily feminist women, women who speak out or women who challenge patriarchal and heteronormative norms, that is, women who step out of the expected norms of femininity and gender norms (Ibid.).

To conclude, returning to the importance of labels and categories, Nasreen Rajani (2022) makes a particularly interesting contribution to the field of study of technology-facilitated violence (TFV) from an intersectional perspective. As argued by this scholar, the very concepts applied to talk about this kind of violence, including “online misogyny,” “gendered cyberhate”, “networked misogyny”, “technologically facilitated sexual violence”, or “online sexual harassment, makes the focus lie explicitly on the sexual or misogynist nature of the enacted violence. This, in turn, “risks framing a generalized and homogenous picture of TFV that is ill-equipped to account for women who are also racialized, queer, disabled, or transgender” (Rajani 2022: 222).

3.4. Towards a working definition of gender-based online hate speech

The review of previous research has revealed numerous studies covering the broader issue of gender-based online violence by both known and unknown perpetrators. Many of these studies propose taxonomies of types of violence,

among which gender-based hate speech tends to be included. As I have previously discussed, it is difficult to separate “hate speech” from other types of online violence, as these often appear interlaced, and many women are exposed to different forms of online violence in parallel. What the label “hate speech” contributes with to the discussion is signalling the functions, the structural roots and the collective impacts of this violence.

Igareda (2022: 100) discusses “anti-gender hate speech”, which she defines as hate speech against the collective of women in general in a specific society, or against specific groups of women; speech that denies, ridicules or stigmatises discourses that show how in our society there are gender roles and stereotypes that assign different characteristics, behaviours and values to men and women; speech that denies or undervalues gender-based violence; and speech that is directed against certain population groups in relation to their gender identity, gender expression or sexual orientation. Of importance is also the ODIHR definition of “gender-based hate crimes”, as

criminal offences motivated by bias against a person’s gender. Such crimes target people, property or associations connected with people or groups due to their actual or perceived gender (...) One of the motivating factors behind this type of crime is the perpetrator’s perceptions of gender norms. The victims of such crimes are often targeted due to their perceived deviation from gender norms, including on the basis of their sexual orientation and gender identity. Gender-based hate crimes may also target people or property due to their association, professional affiliation with or activism on gender issues, such as women’s rights groups and civil society organizations working with victims of violence.

Based on an extensive literature review, following a grounded and intersectional approach and for the purpose of this thesis, I propose the following working definition of gender-based online hate speech:

Gender-based hate speech encompasses speech against the collective of women; as well as speech that denies, ridicules, or stigmatises discussions that highlight the existence of gender roles and stereotypes, or persons who deviate from these norms, including due to their sexual orientation or gender identity. Speech may also target individuals associated with gender issues, such as feminist associations or individuals who express

feminist opinions. In addition to text, hate speech in digital spaces may be expressed through visual or audiovisual elements, and may be reinforced by other types of violence such as doxing²⁷. Discourses and attacks against women and gender non-conforming individuals are often based not only on misogyny and sexual violence but may also include other factors recognised by hate speech frameworks, such as LGBT-phobia, racism or islamophobia.

Based on this working definition and on the literature review, in the next section I shall justify further the specific research focus and scope of this doctoral thesis.

3.5. Rationale of the thesis

As exposed throughout, although most studies have a broader focus on online gender-based violence, there is an emerging field of study focusing on online hate speech expressed against women as well as online abuse against activists and feminists. However, despite the increasing importance of digital anthropology, social anthropologists are rather absent in academic debates on gendered hate speech. Whilst scholars such as Jane and Lewis have, to some extent, used qualitative methods, surprisingly little ethnographic attention has been given to this topic. Further, most scholarly work - regardless of disciplines - on this topic stems from Anglophone countries such as Australia, UK and the United States. Thus, to fill this gap and based on the literature review conducted, the scope of this thesis research focuses on online gender-based hate speech as defined above including its intersectional dimensions; experiences, responses, contestations, and impacts both in terms of individual consequences and societal repercussions.

In geographical terms, the study is limited to Spain, where, as shown by the state-of-the-art, the topic of gender-based online hate speech has received little academic attention. As noted, women are disproportionately targeted by gender-based online hate speech, and certain women are the most targeted, namely those who are public figures, such as journalists, or those who engage in political and especially feminist debates online. This factor makes the topic crucial to the

²⁷ Doxing is the action or process of searching for and publishing private or identifying information about a particular individual on the internet, typically with malicious intent.

wider debate on whose voices are represented in the public sphere. Drawing on these insights, the thesis focuses on women who engage in political debates online, and especially those who take feminist stances. Evidence from previous studies suggests that women espousing feminist views are particularly targeted for abuse online, and thus may provide an important starting point. Further, racialised women seem to be disproportionately targeted.

Considering the scarcity of ethnographically focused studies in the context of Spain, there is a need to unravel whether experiences described in Anglo-Saxon countries are also lived in Spain and what consequences this may have for women's participation in digital spaces. Is the ongoing backlash from far-right sources against women, and especially against women expressing political views and/or racialised women, also palpable in the lives of women who express feminist and/or anti-racist views online in Spain? I do not claim that feminists' experiences can necessarily be extrapolated more widely; however, given their interpretations and perspectives, those who engage in feminism online offer useful insight into their experiences, given the centrality of violence against women in feminist politics (Lewis, 2017).

In relation to an intersectional perspective, this is still largely missing in scholarly debates on hate speech, regardless of discipline and geographical scope. There is an emerging field of studies taking this perspective into consideration in Spain, however, predominantly stemming from the research work of NGOs. There is still little academic debate on this topic and there is a lack of in-depth understanding of the specific experiences of being targeted by gender-based hate speech in digital spaces, both on its own terms and intersecting with other types of abuse or abuse on other grounds. As argued above, applying the concept of "hate speech" facilitates recognition of the different grounds and the structures beneath the attacks instead of perceiving the attacks in an individualised manner. At the same time, we need to go down to the micro, individual level to understand the experiences, impacts, responses and contestations based on significant and situated instances of lived experience.

Further, not only is evidence lacking on the experiences of the targets, but also on the perspectives and experiences of the professionals who in different manners provide support to victims and targets of gender-based violence and hate crimes, on how to tackle online violence and how to provide advice and support to targets and victims.

4. A continuum of online – offline ethnography: Methodological and ethical considerations

4.1 Introduction

This research is methodologically situated within the ethnographic tradition of social anthropology. A “traditional” ethnography often opens with the ethnographer's arrival to the field – an opening narrative that, as Mary Louise Pratt (1986) highlights, is not trivial, but that plays the crucial role of anchoring that description in the intense and authority-giving personal experience of fieldwork. During the years of research for this thesis, I have often reflected on how to depict the research process from the viewpoint of an ethnography, given that I have not travelled to a faraway place to study "the other" and cannot delimit a physical field. Nor can I describe my arrival to the field, as I am in a sense constantly in it. I have rather adventured upon the studying of "us" (e.g., Nader, 2002a).

Due to this complexity in defining the field in concrete geographical or cultural terms, the first part of this research, then, precisely consisted of delimiting the field. As mentioned at the end of Chapter 3, in relation to the scope of this study, the first step limited the case study to Spain. The second step consisted in delimiting the field further, in terms of a more detailed scope. In this regard, I perceive the field as situated within the intersection of four pillars, and at the

same time as an interaction of these pillars: First, the targets of gender-based online hate speech, with a focus on women who express feminist and/or anti-racist opinions; second, the current legal framework and the socio-political landscape with its undercurrents of neo-liberalism, post-feminism, and anti-feminism; third, the institutions upholding the legal framework and the services provided by these institutions and by other actors such as civil society organization; and fourth, the technological affordances of and the interactions in online spaces.

The methodologies applied, further described in the sections below, have been those traditionally used in anthropological research, with a focus on ethnographic or in-depth interviews, whilst also including digital ethnography or participant observation in digital spaces. This combination of online and offline tools has been particularly valuable for picking up on different experiential and interactional aspects that cannot be captured solely by interviews nor by digital observations, but that have emerged from the combination of the two.

Further, the research has coincided with the Covid-19 pandemic, which has led to a digitalisation of the fieldwork itself, moving from face-to-face connections to connecting through a computer screen. That is, the research itself has been conducted on the continuum of the online and the offline, giving less importance to the physical context and placing more emphasis on the connection and content. Nevertheless, while the “digital” in digital ethnography is often used synonymously with “remote” or “virtual” online-only ethnography (Hine, 2015; Boellstorff et. al. 2012), I take a “postdigital” approach to digital ethnography (e.g., Coleman and Jandric, 2019; Pfeifer and Jovovic, 2023). This extends beyond the online/offline or analogue/digital divide, enabling an examination of the contextual nature of digital media and associated practices in everyday life from a critical perspective. That is, it encompasses various methods and approaches across the online/offline spectrum, as described below.

In the next sections I will discuss the reflexivity of the research process and the application of a feminist intersectional approach, after which I shall address the

specific methods and sample, and finally, I will explore the ethical considerations of this type of research.

4.2 Reflexive ethnography and the intersectional feminist lens

Whilst considerations of reflexivity are important for all forms of research, for social anthropological research this aspect becomes even more crucial. Reflexivity can be defined as “turning back on oneself, a process of self-reference” (Aull Davies, 1999: 4). At its most immediate level, this refers to the ways in which the research is influenced by the persons who conduct the research and the processes of conducting it. These effects can be found throughout the research process; from the selection of the topic to the final reporting of results (Ibid.). Issues of reflexivity are particularly crucial to ethnographic research, as, traditionally, the social interactions and relationships between the researcher and the informants form the basis of the subsequent theorising. That is, the impact of the presence and personal configuration of the researcher on the field must be considered.

In the introduction to this chapter, I mention that this is a study of “us”; however, this “us” must be nuanced. Like most of my informants I am a woman, but I am a cis heterosexual woman, a white, close to middle-aged woman. At the same time, I am a foreigner in this country, but a privileged foreigner from the North. As a woman, I have experienced harassment and violent nuances on the continuum of gender-based violence. As a foreigner in Spain, I have been screamed at on the street, “go back to your country”, but this is far from the constant abuse and discrimination suffered by many racialised persons. I am a feminist, but I’m not a particularly active feminist. If you analyse my Twitter feed and the profiles I follow, you can most likely deduce that I am left-wing and pro-trans rights. That is, I am not neutral, but rather positioned within a set of political stances with more or less relevance for the topic under study. I am not a racialised, lesbian, bi or trans feminist activist. And I have not – at least up until now – experienced online hate speech or other types of online attacks. This means that whilst I’m in a sense studying “us”, I’m also an outsider to the field in many ways. Further, my openness on social media regarding my political stances may have an impact on

people's willingness to open up to me. For example, as previously mentioned, the current climate on social media in Spain shows an increasing gap between those feminists who are pro trans rights, and those who are embracing transexclusive or transphobic stances. In this regard, some individuals may have felt hesitant to commit to an interview or to talk to me even informally. Being open with who I am – that is, not a neutral research tool, but rather a social being with all their intersections and contradictions - impacts both the sample and the discussions during the interviews. As Margery Wolf puts it, “we as anthropologists can only try to be sensitive to the implications of our perceived status, implications that may be even more troubling for the fieldworker who works in her own society” (1992:13).

For David Graeber (2004: 12), when carrying out an ethnography, the anthropologist observes what people do and then tries to disentangle the hidden symbolic, moral, or pragmatic logics that underlie their actions; trying to understand how people's habits and actions make sense in ways that they are not themselves completely aware of. Then, if informed by reflexivity and assessed by a critical scholarly community, the results of ethnographic research “are expressive of a reality that is neither accessible directly through native text nor simply a reflection of the individual anthropologist's psyche” (Aull Davies, 1999: 6). Nevertheless, whilst the anthropologist listens to as many voices as possible, she eventually takes the responsibility for putting down the words, “for converting their possibly fleeting opinions into a text” (Wolf, 1992: 11). This action of the anthropologist of putting a variety of voices into words is, to a greater or lesser extent, always still an exercise of power (Ibid). In this sense, what the ethnography reveals is an interpretation of reality created between the anthropologist and the informants that would not be counted in the same manner by a complete “insider”, as the ethnographer has boiled down the variety of voices and conducted an analysis based on their understanding and background. In James Clifford's (1986: 2) words,

ethnography is actively situated *between* powerful systems of meaning. It poses its questions at the boundaries of civilizations, cultures, classes, races, and genders. Ethnography decodes and recodes, telling the grounds

of collective order and diversity, inclusion and exclusion. It describes processes of innovation and structuration, and is itself part of these processes.

Further, as I have already pointed towards above, from a feminist methodological point of view, we must be transparent with the fact that all research is always determined from a specific ideological position and imbued with certain values that influence and determine the research (Igareda et al, 2019). Feminist anthropology has questioned dualities such as subject/object, personal/political, thinking/feeling, building instead a dialogical and intersubjective path between the “researcher” and the “researched” (e.g., Gregorio, 2014). To break these dualities and balance out power relations in the research process, different feminist ethnographers have proposed various approaches. For instance, Ruth Behar affirms that sharing our vulnerabilities may be a way of balancing power relations: “We ask for disclosures from others, but we share very little or nothing about ourselves; we make others vulnerable, but we remain invulnerable” (1997: 273). On the other hand, feminist research that promotes personal positioning and involvement can also be more ethically precarious – as there is a danger of instrumentalising personal relationships or imposing points of view during the interview (e.g., Stacey, 1988). This means that it is also necessary to check that the researcher’s story or voice does not overshadow or silence the voices of the informants (Ellis-Sloan, 2014). With this perspective appropriately applied, the narrative in a sense becomes a meeting place between two subjectivities, that of the researcher and that of the informant, who both become collaborators and participants in the production of knowledge. In this way, the interviews become a form of collaborative theorising, as Patti Lather (1986) puts forward. The participant is not just a “source of data”, to be analysed, but rather an expert on their own experience, treated as capable of being analytical and reflective on that experience and what it means.

Finally, in the sense that method concerns the way one thinks (MacKinnon, 2013: 1019), adding an intersectional lens to the methodology not only provides prompts for the sample in terms of significance of different experiences. It also provides a different way of thinking about inequalities – as dynamic interactions

between structural oppressions – and how to capture these in research. Intersectionality, as proposed by Avtar Brah (2004: 76), signifies “the complex, irreducible, varied, and variable effects which ensue when multiple axis of differentiation – economic, political, cultural, psychic, subjective and experiential – intersect in historically specific contexts”. That is, the concept emphasises that different dimensions of social life cannot be separated out into discrete and pure strands. Or, as put forward by Kimberlé Crenshaw, it addresses “the combined effects of practices which discriminate on the basis of race, and on the basis of sex” (1989: 149). That is, intersectionality as a method does not solely add variables as separate axes of analysis, but rather looks at the imbrications of the different dimensions. Then, with each new intersection, new connections emerge, and previously hidden exclusions come to light (Davies, 2008). By asking another question, the research may take on a new and sometimes surprising turn, exploring the consequences of intersecting categories for relations of power (Ibid). As Kathy Davies (Ibid.) pinpoints, intersectionality, in this manner, presents numerous possibilities for scrutinising personal blind spots and converting them into analytical tools for subsequent critical examination. Due to its inherent openness, an intersectional perspective can set in motion a discovery process that produces fresh, more comprehensive, and reflexively critical insights (Ibid).

4.3 Methods and sample

4.3.1. Digital ethnography or Participant observation in online spaces

The digital provides many new opportunities for anthropology to help us understand what it means to be human (Miller and Horst, 2012). In this regard,

the internet provides access to intimate spaces, [...] that let us see people making sense of the world on their own terms, and at a profoundly human scale. Everyday actors make available their thoughts and experiences online in a variety of ways and at a great diversity of sites (Recuber, 2017).

In this sense, the affordances of digital technology can enhance traditional forms of ethnography, as methodologies that are attuned to questions of meaning and that work with small sample sizes (Smith, 2014).

According to Dhiraj Murthy's (2008) definition, "digital ethnography" involves the incorporation of digital technologies such as online questionnaires, digital videos, social media sites, and blogs into the ethnographer's traditional toolkit. Other scholars go further and advocate a kind of "netnography" in which online interactions constitute the entirety of the research site (e.g., Kozinets, 2010; 2015). Yet this approach has been criticised for drawing an untenably sharp distinction between the on- and offline world and ignoring the latter, or the interactions between the different spaces (Dumetricia, 2013). As I aim to analyse lived experience with a perspective of the online and the offline as an interlaced continuum, my digital ethnography has been closer to Murthy's notion than to a "netnography", and thus, as mentioned above, situated within what can be defined as a post-digital approach. Digital ethnography from this approach refers to ethnographic research with and through the "digital" and not limited to the remote or online. In this sense, observations in digital spaces are not essentially different from traditional participant observation, as a way of participating, observing, and understanding the world of our research participants, complementing in-depth interviews. As such, observations contribute to highlighting the possible gaps between what people say they do and what they actually do (Hylland Eriksen, 2010).

As I have argued throughout Chapter 2, digital spaces must be studied in relation to the structures underlying society and that therefore also affect these spaces. In the absence of a contextual analysis implied by a relational view of social networks, "social media is simply a bubble of communication floating in the ether, detached from the structural conditions created by states, polities and economies" (Wilson, 2019). This means that ethnography, with its comprehensive approach examining the interconnections among various social domains like technology, gender, racialisation, and sexual orientation, emerges as a highly relevant method for investigating social media.

Methodological approach and sample for the digital ethnography

During the research for this thesis, I embarked on the digital ethnography before I started the process of interviewing. In this regard, the online ethnography helped

me identify research participants and relevant associations, as well as the scope for the interviews. However, it also constituted a collection of ethnographic material in itself, through the observation and analysis of online interactions.

In terms of the use of digital spaces throughout this research, the main platform where research has been conducted is Twitter. This decision was based on the fact that Twitter, now known as X, was still the main platform for political debate and activism, whilst at the same time it has been described as the most hostile of social media (e.g., Lewis, 2016; Jubany and Roiha, 2018). Nevertheless, to a certain extent I also followed the same persons and organizations on other platforms such as Instagram, YouTube and Twitch, and, to a lesser extent TikTok. This movement between platforms was informed by the case study profiles' and interview participants' own use of different platforms for different interactions, from a perspective of "polymedia" (Miller and Madianou, 2013).

To systematise this online participant observation, my main focus was on case studies on Twitter. These cases included persons who auto-define as women and feminists, and who actively and openly debate feminist and and/or other political issues such as anti-racism or trans rights on social media, and who in the past few years have been targeted by gendered online violence. That is, they correspond to women, including trans women, and non-binary persons who have been targeted by hate speech or other types of attacks, either based on their gender and their online activism, or these factors in intersection with other factors, such as racialisation or LGTB+ identities. The cases were collected through a snowball sample, which was constructed by following feminist, LGTB+ and anti-racist organizations as well as individual feminists, LGTB+ and anti-racist activists on Twitter, aiming to cover a sample as heterogeneous as possible within the context of Spain.

I started building the list of profiles to follow in October 2019 and have expanded it throughout the ethnography. Through following these profiles, I was alerted to cases of gender-based hate speech or attacks against feminists, either through re-tweets or original tweets by these profiles, where they publicly denounced an attack. This manner of sampling means that the majority of cases involve a

public denouncement by the target, or by a follower of the target. Nevertheless, some cases where the target has not denounced the abuse have also been included. In these cases, I have myself spotted misogynist, racist or LGBT-phobic abuse in the comments to their tweets. In total, I have collected 26 cases of gender-based online hate speech on Twitter, corresponding to 26 different individuals. The first case was collected in May 2020 and the last cases were collected in May 2023. The entire anonymised list of cases can be found in Table 1, below.

Table 1: Cases collected in the digital ethnography

Code	Profile	Date of the attack or public denouncement²⁸
DE-1	Feminist journalist	June 2021
DE-2	Feminist journalist	July 2021
DE-3	Antiracist activist	July 2021
DE-4	Politician, councilor in a local government.	April 2022
DE-5	Non-binary person, local politician.	May 2022
DE-6	Trans woman, activist, academic	June 2022 July 2022
DE-7	Journalist	August 2022
DE-8	Antiracist activist, feminist and communicator	August 2022 January 2023
DE-9	Trans woman and activist	September 2022
DE-10	Trans woman and activist	September 2022
DE-11	Trans feminist, activist	September 2022
DE-12	Trans woman and comedian	June 2022
DE-	Feminist gamer	December 2022

²⁸ In the cases where I have analysed several attacks or public denouncements for the same individual there are several dates. The code numbers do not follow a strictly chronological order. This is because I collected the first cases in my field diary in 2020 without systematising the information collected. A few cases were also collected months or even years after the attack had taken place.

13		January 2023
DE-14	Social media feminist	December 2022 January 2023 February 2023
DE-15	Racialised journalist	October 2022
DE-16	Trans woman, local politician	April 2022
DE-17	Roma feminist	March 2020 May 2020 April 2022
DE-18	Feminist journalist and activist	May 2020 June 2022 February 2023
DE-19	Feminist journalist	February 2023
DE-20	Social media feminist	June 2022 September February 2023
DE-21	Local politician, feminist	September 2020
DE-22	Anti-racist activist and feminist	July 2020
DE-23	Journalist	June 2022
DE-24	Online communicator and gamer	April 2023 May 2023
DE-25	Science communicator	May 2023
DE-26	Sociologist specialised on gender-based violence	May 2023

To keep track of the information gathered in the online spaces, I kept a field diary in the same way as I would have done with in-situ fieldwork. This means that I took dated field notes, thus recording my observations, reflections and questions, in relation to cases of attacks that emerged, as well as general observations of

gendered interactions and violence mainly on Twitter. This diary also includes screenshots as well as links, which helped to assist my memory, and to collect comments and tweets to feed into Atlas TI at the stage of analysis. I also kept a table of the case studies, with links to the tweets or other sites where the abuse had either been denounced or directly perpetrated. To be able to analyse the interactions and comments to the posts in a more systematic manner, I introduced the text from the relevant tweets and the comments to the tweets in Atlas TI and used labels in the same way as for the interview material. As previously mentioned, the cases have been anonymised in the analysis. To further protect the anonymity of these profiles, I have chosen not to cite literally, which I will discuss further in section 4.5. Ethical Considerations.

4.3.2. Ethnographic interviews

The scope of my study and the lack of a clearly delimited territorial field lead to the decision to focus on ethnographic interviews as the main research tool, supported by digital ethnography, as described above. These types of interviews are often referred to as “semi-structured”; however, as Bill Gillham (2000) notes, the structured-unstructured dimension of interviews is in a sense false, as the interviewer will always have some sort of structure in mind, which is used flexibly according to what emerges during the interview process. Nevertheless, whilst the interviewer considers the broad aim of the interview, it is the person interviewed who has something to tell the interviewer and only they can provide this information from their perspective. That is, the interviewer needs to know how to listen actively, with focus on the person interviewed, whilst being aware of their own aim. This calls for technique and at the same time “naturalness” – responding naturally to the person interviewed, who will sense the interviewer’s interest and concern (Ibid.). This genuine interest was critical to the interviews conducted during this research, and particularly for the interviews with targets of online violence, who had been through often traumatising experiences, whilst at the same time they had often been met with a lack of understanding of the impact of online harm. Empathic active listening, then, was key to making the persons interviewed feel safe and listened to. I also believe – or hope – that the

interview benefited them to some degree, as a chance to talk about what happened to them, with a person who was genuinely interested and took the topic seriously.

Further, in relation to the reflexivity discussed above, I find the following remark by Gillham highly relevant to the discipline of anthropology: “in a research interview you are the research instrument and you are not a standard product” (2000: 4), which links to the discussion of reflexivity above. In this regard, not only does the researcher bring their own background into the research, but the style of interviewing is also very personal, and each ethnographer must find their own style of relating to the field and of interacting with people. In my case, my status as a foreigner in Spain to some degree facilitated my status as an interviewer. On the one hand, for some interviewed professionals this worked as a prompt to go to further lengths of explaining issues with their own words, which they perhaps would have taken for granted that a native interviewer knew. On the other hand, some persons may feel less trust and limit themselves when speaking to an “outsider”. My “liminal” status as a “not quite outsider” also helped in that I could make comparisons with the situation in my country of origin, Sweden, and the context surrounding gendered online hate speech there. During some of the interviews, this helped unblock some topics and may also have made the persons interviewed gain some new information on the topic.

Interviews with key professionals and experts

The initial research plan for this thesis included only a few interviews with legal experts. However, after the first few interviews, I decided to conduct more interviews with experts and professionals working with the topic of gender-based online hate speech from different perspectives. This decision was based on several factors: considering the lack of previous studies on the topic, the research called for a better contextualisation of the problematic of gender-based online hate speech, that is, more situated knowledge on the issue in Spain. I also lacked information on what services are available in terms of reporting and support, and related to this, there was a need to better understand the perceptions of the actors involved in these services. This was because a first overview of the legal

framework and the first few interviews with legal experts revealed a fragmented framework with significant gaps in terms of an important lack of gender perspective. As this gap would have an impact also on reporting to the police and access to support services, I decided to interview representatives from different institutions, predominantly in Catalonia. In addition to the institutional representatives, I also interviewed persons who are experts on the issue of digital gender-based violence from different perspectives, including academics and activists, as well as NGOs that offer support to victims of hate crimes and gender-based violence, or that are generally active in the fields of feminism and LGBT+ rights. These interviews focused more on triangulating and complementing the literature review with further perspectives as there is still very little academic research on this phenomenon in Spain. Finally, these interviews also served as further guidance in elaborating the interview guide for the interviews with targets of online gender-based hate speech.

I started conducting in-depth interviews for this thesis research in November 2020, during the Covid-19 pandemic. This had an important impact on the sampling and interview process. Because of my role as a researcher within action-research projects for the last few years, I had previously established contacts with many institutions and organizations across Spain, which I drew upon in the early stages of fieldwork. Through these previous contacts, I received new pointers on relevant people to talk to. This way, the access was accomplished through snowball sampling, where a previous interview person introduced me to another one, often by e-mail.

Through the interviews with professionals, I tried to capture the importance these institutions give to the overall issue of online gender-based hate speech; what services they may offer to victims; the role of the legal framework; and how prepared they perceive that they are to offer support to victims of digital violence. At the moment in time when I conducted the main part of the interviews with professionals (late 2020 and early 2021), these questions had become increasingly relevant in the context of Catalonia, with the approbation of the *Llei 17/2020, del 22 de desembre, de modificació de la Llei 5/2008, del dret de les*

donec a erradicar la violència masclista, which is an administrative law that for the first time includes the concept of "digital violence".

In total, I conducted interviews with 20 professionals and experts. Most interviews were conducted between November 2020 and April 2021; however, two complementary interviews were conducted in April 2023. The entire sample of professionals can be found in Table 2, below.

Table 2: Sample of experts and professionals interviewed

Code	Profile/expertise	Date
EX-1	Representative of an LGBT+ organization	November 2020
EX-2	Representative of a feminist association	November 2020
EX-3	Representative of a feminist association	November 2020
EX-4	Representative of an LGBT+ organization	November 2020
EX-5	Representative of an LGBT+ organization	November 2020
EX-6	Lawyer specialised on hate crimes and gender-based violence	November 2020
EX-7	Representative of a public institution (regional) in the field of gender equality and LGBT+ rights	December 2020
EX-8	Representative of a public institution (regional) in the field of gender equality and LGBT+ rights	December 2020
EX-9	Prosecutor	December 2020
EX-10	Police officer	December 2020
EX-11	Police officer specialised on victim support	December 2020
EX-12	Representative of a public institution (local) in the field of gender equality and LGTB+ rights	December 2020
EX-13	Representative of a public institution (local) in the field of anti-discrimination	December 2020
EX-14	Representative of a public service (local) in the field of gender-based violence	January 2021
EX-15	Law scholar specialised on gender-based and anti-gender violence	January 2021
EX-16	Representative of a public service (local) in the field of gender-based violence	January 2021

EX-17	Sociologist and representative of a feminist association	January 2021
EX-18	Representative for an information service on gender-based hate crimes (NGO)	April 2021
EX-19	Psychologist specialised on gender-based violence	April 2023
EX-20	Researcher in the field of trans rights	April 2023

Due to the Covid-19 pandemic, I let the persons interviewed decide whether they wished to meet online or face-to-face. Based on these preferences, only three of the interviews with professionals were conducted face-to-face: one with a prosecutor, and two interviews with police representatives. It is quite telling that those participants who preferred to meet face-to-face were the institutional actors for whom it was not possible to work from home during the pandemic and who continued to be engaged in face-to-face contact with citizens throughout the pandemic. The remaining part of the interviews were conducted with the help of digital tools, including Google Meet, Teams, Zoom and Webex.

Interviewing professionals and experts through digital media proved to have some advantages: roughly a year into the pandemic, professionals from different fields were highly accustomed to using digital tools, meaning that they were quite easily accessible; perhaps even more so than in a face-to-face setting. In most cases they managed to find time in their busy schedules with short notice. Further, the digital setting meant that I was able to interview persons located also outside of Catalonia, predominantly in Madrid. Finally, there are also other advantages of these tools, including high-quality audio recordings as a built-in feature, whilst they also grant the interviewed individuals full control over the management of their privacy and intimacy throughout the process (Roca and Folguera, 2021).

Interviews with targets of gender-based online hate speech

As a second phase of fieldwork, I conducted interviews with targets²⁹ of online gender-based hate speech. Supported by the literature review and interviews with professionals and experts, one of the primary groups of targets of this type of online violence are women with an activist or feminist profile and with an online presence. Previous international research also shows that women with this profile, who are intersected by further structural oppressions (e.g. lesbian, trans, racialised or Muslim women), suffer from online aggressions to a higher degree than cis-white-hetero women. Thus, my aim was to build a sample as wide as possible in terms of different backgrounds, of women with an activist or feminist online presence, who had been targeted by verbal violence online, or what I conceptualise as gender-based online hate speech. Some of these women had also suffered from other types of online violence, as the attacks often take different shapes and include other elements that might not fall directly under a definition of hate speech, but that nonetheless are highly relevant to this research, as they carry the same function as processes of dominance and exclusion.

Whilst the pandemic context did not pose any major challenges to the interviews with professionals, unfortunately, Covid-19 was somewhat detrimental to the access to interviews with targets. My initial strategy was highly based on participant observation in different in situ assemblies and events, where I had planned to establish a broader network through participant observation. Due to the Covid-19 restrictions, during the main part of my fieldwork time there were hardly any face-to-face events and assemblies of this type, meaning that I was only been able to do this towards the end of my fieldwork. Whilst, as I shall discuss further below, the online events that I have participated in have provided me with valuable information on the topic at large, during this type of events it is

²⁹ As pointed out in Chapter 1, I have chosen to use the word “target” as opposed to other words such as “victim”. These persons are targeted by violence and hate speech, but most of them would not define themselves as “victims”. Denominating them “targets” also speaks of the continuity of attacks and the crimes or incidents as processes rather than one off events.

not possible to informally interact with other participants and network in the same sense as during face-to-face events.

Thus, due to the circumstances of Covid-19, I used the following paths of access:

- Throughout the interviews with professionals who work directly with victimised persons, I asked them whether they had any persons in mind who would fit the profile of the sample and who they believe could be interested in participating. Nevertheless, in practice, most of the professionals and services had not provided support to persons with this profile.
- I elaborated a list of more than 40 organizations and platforms, including feminist associations, cyberfeminist associations, associations of Muslim women, association of Roma feminists, afrofeminist organizations, associations of lesbian women, a digital feminist magazine, organizations of trans women, associations of women- and feminist lawyers, an association of women journalists, a platform against gender-based violence, as well as associations of women gamers and developers of video games. I contacted these organizations with an e-mail that briefly exposed the research and explained that I was interested in interviewing women who had been targeted by gendered online violence. Whilst several of the associations replied to my e-mail, saying that they had forwarded the message to their members or to other organizations, through this strategy I only identified two persons interested in participating.
- Through the digital ethnography and the case studies on Twitter, I elaborated a list of around 30 women who had been directly targeted by online gender-based hate speech. Contacting these women directly was my last strategy. I hesitated to do this, as I did not have a personal connection with most of these persons. This lack of personal connection posed two problems: first, for self-protection, most of these profiles have established their online privacy settings so that people whom they do not follow on Twitter cannot send them a direct message through the platform. This means that if I follow them on Twitter and they have not

followed me back, I cannot contact them through a direct message on the platform. Second, considering ethical considerations in relation to the risk of secondary or re-victimisation, I hesitated regarding what their reaction might be to a person contacting them to invite them to participate in research where they would be asked to talk about their experiences of online violence. Nevertheless, I managed to contact a number of these persons. All of these were positive to the research and several agreed to participate; however, due to their busy schedules, on numerous occasions interviews were cancelled and rescheduled, and in the end, I only managed to interview two of these persons. I subsequently decided to delete these individuals from the list of case studies, in order for these samples not to overlap.

- Finally, based on the cases collected on Twitter and through a few personal connections, I created a snowball sample.

The final sample of targets interviewed consists of 10 women, detailed in the table below. As previously discussed, these include women who identify as feminists and/or anti-racist and/or LGBT+ activists, who are active online, and who have experienced online violence due to these factors. These women live in the following autonomous communities: Andalucía, Cataluña, Madrid, and País Vasco. To protect their identity, in the sample table below, I have not linked the profiles to their respective locations.

Finally, whilst the sample is fairly heterogeneous in terms of age, sexual orientation, racialisation, and location, I did not manage to include any trans persons or Roma activists among those interviewed, despite contacting several relevant profiles. Their omission to participate or lack of response to my intents of contact, may be related to a research saturation of certain minority groups, who constantly find themselves in the position of “objects of study” (Guyan, 2022). Nevertheless, these profiles are reflected in the digital ethnography, as detailed above.

Table 3: Sample of targets interviewed

Code	Pseudonym	Profile	Age	Date
T-1	Julia	LGBT+ activist	26	April 2021
T-2	Elena	Feminist and LGBT+ activist; Gamer	30	May 2021
T-3	Daniela	Feminist involved in youth politics	22	May 2022
T-4	Amara	Feminist and antiracist activist	39	June 2022
T-5	Sara	Journalist focused on issues of gender and feminism	31	January 2023
T-6	Lucía	Feminist journalist	33	February 2023
T-7	Alba	Feminist activist	42	April 2023
T-8	Aurora	Feminist mainly active on social media	32	April 2023
T-9	Naomi	Antiracist and feminist activist	28	July 2023
T-10	Susana	Feminist mainly active on social media	25	September 2023

When conducting the interviews, I followed an interview guide, which I had elaborated based on the literature review, on the interviews with professionals and experts, and on the digital ethnography. The guide covered the main themes identified when defining the research objectives, including the experience(s) of online violence; the impacts and consequences of being targeted; as well as the reactions, responses, and contestations, including the experiences of formally reporting to the police or another authority in the cases where this had been done. As reflected above in relation to ethnographic interviews, the guide was used as a dynamic and flexible tool, serving as a reminder to myself as the interviewer of what topics I aimed to cover, through open questions. Nevertheless, in most of the interviews I started the interview with a general open question about their trajectory on social media – that is, what channels they use and for what purposes. This allowed for a “neutral” start of the interview and often also served as an introduction to the topic, allowing space for the person interviewed to express their own narrative.

All the interviews with targets were conducted online. This was for several reasons: First, the pandemic context was still highly present when I started these interviews in the spring of 2021. Second, as this thesis project concerns online gender-based hate speech in Spain, I aimed to cover more territories than those that are geographically the closest to me. Conducting interviews online thus provided me with the opportunity to talk to persons in different parts of Spain, , reflecting a wider range of experiences. Third, the persons I interviewed in this phase of the fieldwork can be characterised by their “living in media” (Deuze et al. 2012), that is, they are highly used to communicating through digital tools. Finally, I have found that it is often an advantage to talk about digital experiences through a screen, as screen-sharing, screenshots and direct messaging of links by the participants are valuable tools that have given me further hints and information and creates a “thicker” description of their experiences. These tools turn the interview into a more dynamic encounter, that could be compared to “walking interviews”, in that in some cases I let the participants take me to one or several of their social media profiles through screen-sharing, and moving through that digital space, new topics and conversation turns were prompted, either by me or by them.

4.3.3. Participation in events, courses and workshops

The fieldwork conducted has also gone beyond formal and informal interviews and observations, and into other events that provide opportunities to delve deeper into the topic or address it from other perspectives. An important aspect of approaching the studied topic has been achieved through attendance, observation, and, in some cases, participation and presentation at specialised workshops, seminars, courses, and congresses.

During the first part of the research, I participated in a wide range of informal online events, most of which, in a non-pandemic context, would have been held face-to-face. Most of these events were organised by feminist organizations or by informal groups of feminists, to discuss different topics related to gender inequalities, gender-based violence and LGBT+ issues. These events were organised as interactive live streams, such as the series of Instagram live sessions

called "De víctima a superviviente"³⁰ or livestreams on Twitch such as "Discurso de odio y la educación"³¹. My participation in these events allowed for a better comprehension of how issues surrounding online gender-based violence in general, and hate speech in particular, are perceived and discussed collectively. The digitalisation of interaction spaces allowed me to participate in a much wider range of events than had they all been arranged face-to-face. However, online events do not allow for the same type of informal interactions, or "mingling" as face-to-face events, which is why these events to a lesser extent than planned provided access to the field. Still, these events provided me with highly relevant and valuable information. In this regard, I found fieldnotes to be particularly useful to reflect upon interactions during these online events. To put forward an example, during an Instagram live stream where two feminists discussed the journey from "victim" to "survivor" of gender-based violence, a few men insisted on calling the attention of the participants of the discussion by asking non-relevant questions in the live comment field and demanding a response time upon time. When they did not receive the attention that they wished for, and did not manage to disrupt the event, they started an anti-feminist discussion in the comment field. I have several further examples of similar situations noted down, which relate to the conquering of online spaces and both explicit and symbolic violence that women are subjected to. This happens both when they enter more traditionally male-dominated spheres online, and when they try to create their own public discussion spaces, such as in the example above. Further, as expressed in the example above, this participant observation on a few occasions made me a direct bystander to gendered power relations in online spaces.

I also participated in more formal workshops, conferences and courses related to the topic of online hate speech and gendered violence, organised both online and offline by public institutions and civil society organizations. These included the *Curs de Comunicació i Periodisme Intercultural #ComuniCanvi*³²; *Jornada El*

³⁰ Organised by psychologist Aina Troncoso and lawyer Carla Vall in November 2020.

³¹ Organised by Maldita Educa in May 2021.

³² Organised by Llatins/es per Catalunya, ITACAT and Xarxa BCN Antirumors, at the Espai Societat Oberta, in November 2019.

*ciberassetjament i altres delictes virtuals envers les dones*³³; *Jornada ¿Estamos reparando el daño? Justicia y prácticas restaurativas en el abordaje de casos de odio y discriminación*³⁴; *Virtual symposium Restorative Justice over Distance*³⁵, *Jornada de violències masclistes en entorns digitals i videojocs*³⁶; *Cicle de Debat Libertinas y cancelados: Laura Bates & Proyecto UNA: De la crueldad sádica a la «machoesfera»*³⁷. Participating in these events has not only made it possible to deepen my knowledge on the topic, but it has also brought me closer to current debates and to the different perspectives that exist in relation to the topics investigated.

4.4. Data management and analysis

All the interviews were recorded with a voice recorder and were subsequently transcribed. The interviews were conducted and transcribed in either Spanish or Catalan. Interview quotes used in the thesis have been translated into English, always aiming to keep as close to the original meaning as possible.

To guarantee the anonymity and confidentiality, each participant was assigned a code instead of their name or a pseudonym, as further described in section 4.5.1. The codes follow the pattern of “EX-n^o” for professionals, “T-n^o” for targets, and “DE-n^o” for cases from the digital ethnography. Nevertheless, when writing the empirical chapters of this thesis, I decided to use pseudonyms instead of codes for the women interviewed as targets, to highlight the person behind the experience and the narrative. Further, for the interviews with these women, any details that would allow for an identification of the participants were omitted from the analysis.

Regarding the interviews with experts and professionals, to protect the anonymity of the representatives of different organisations and institutions, these

³³ Organised by Observatori Català de la Justícia en Violència Masclista and Centre d’Estudis Jurídics i Formació Especialitzada (CEFJE), in November 2020.

³⁴ Organised online by the Institut de Drets Humans de Catalunya and Sos Racisme in March 2021.

³⁵ Organised online by the European Forum for Restorative Justice in June 2021.

³⁶ Organised by the Institut de Seguretat Pública de Barcelona, at the l’auditori del Cibernàrium, Barcelona, in December 2022.

³⁷ Organised by CCCB, Barcelona, May 2023.

entities are not named in the analysis, but rather described as e.g. “Local support service for victims of gender violence”. The persons interviewed have been assigned with a code instead of their name, also omitting their specific position in the organisation or institution.

All interview transcripts, fieldnotes and material from the digital ethnography were introduced in Atlas TI as a hermeneutic unit. With the help of this software, I attached a set of thematic labels and sub-labels to the content. These labels were dynamic themes that were first established through the literature review, and later revised and changed during fieldwork and analysis, with more subthemes emerging during the process.

4.5 Ethical considerations

From a broad perspective, the basic ethical principles to be maintained in any research process include doing good, not doing harm, and protecting the autonomy, wellbeing, safety and dignity of all research participants (Iphofen, 2021). I have already addressed some of these principles above in the discussion of a feminist methodological perspective, in terms of managing possible power imbalances, and making sure that participants feel safe, and that they feel free to talk, or not talk. Researchers, in this regard, must be aware of how their activities may cause distress to others. For example, it is difficult to anticipate the emotional effects that an apparently innocent question may have.

Further, the researcher should make sure that participants understand the purpose of the research and that their participation can be withdrawn at any stage. This means “ensuring participants have, and perceive themselves to have, adequate power to determine their role in the research is seen as ethically necessary. The ultimate test of the enhanced power of research subjects lies in their knowing that they have the ability to withdraw from the study at any point” (Ibid: 4). To ensure this, I have provided participants with a document summarising the research and its purposes, as part of the informed consent process, as described below.

4.5.1 Anonymity and confidentiality

Considerations of anonymity and confidentiality have been integrated in the research process to make sure that participants are safe and that they can freely share their stories without fear of consequences, respecting the privacy and confidentiality of participants. To this aim, any information that might identify an individual participating in the project has been excluded from the analysis process. To protect the anonymity and confidentiality of the interview participants, and to facilitate data management, each participant has been assigned a code. This code – in combination with a pseudonym for the interviews with targets – has been used for all references to that specific participant, including transcriptions, and any quotes used in the analysis included in this thesis, as well as in any publications. However, it is necessary to be able to re-trace a participant's identity to fully respect the right to withdraw from the study. Instead of anonymisation, it is therefore more relevant to talk about “pseudonymisation”. Pseudonymisation entails substituting personally identifiable information (such as an individual's name) with a unique identifier. This difference is important as in cases where it is possible to re-identify the individual data subjects by reversing the pseudonymisation process, data protection obligations still apply. They cease to apply only when the data are fully and irreversibly anonymised (DG Research and Innovation of the EC, 2021). Therefore, I have informed participants that I will keep their personal data (name, last name and e-mail address) only as long as it is needed for the purpose of fulfilling the objectives of the research. These issues are further discussed below in relation to informed consent and research in digital spaces.

4.5.2 Informed consent

Gaining consent cannot be easily separated from the giving of information (Iphofen, 2021). Subjects should be able to choose “freely” to participate in research and should have been given enough information about the research for them to know what their participation involves. Due to the open and long-term nature of fieldwork in anthropological studies, participants' giving of consent may have to be treated as ongoing throughout the research engagement, and may have to be negotiated repeatedly (ASA, 2011). In the present study, all the

research participants were explicitly asked for their written informed consent, including permission to record data. That is, during fieldwork, I aimed to ensure that the nature, purpose, results, duration, conclusions and sharing of the research were clearly explained, so that participants understood fully what they were agreeing to. The participants were also informed that they could withdraw their consent and participation at any time during the research process. To give their consent, participants were asked to sign a consent form, written in accessible language. A Project Information Sheet (PIS) accompanied the consent form, with the aim to inform participants about the research and provide them with my contact details and the contact details of my doctoral supervisor.

For the online interviews, I sent the document by e-mail to participants prior to the interview, together with a summary of the research. They returned it to me signed also by e-mail, either before the interview, or, in a few cases, immediately after the interview. For the in-situ interviews, the participants received and signed the document directly on paper. I also explained the purpose of the research and the conditions of their participation orally and provided them with space to ask questions before starting the interview. The informed consent forms were stored securely offline (that is, not in a cloud location).

4.5.3 Ethical considerations for research in digital spaces

In terms of specific ethical considerations for research in digital spaces, I have relied on the guidance provided by the Association of Internet Researchers (AoIR), particularly their updated ethical guidelines (Franzke et al, 2019) and the companion *Feminist Research Ethics*, as well as the AoIR's previous recommendations (2012).

As highlighted by Daniel Miller³⁸, the ethical considerations of participant observation online would in many ways be similar to those in an offline setting, in the sense that when conducting offline participant observation, the ethnographer does not always walk around with a consent form, but rather pick up on things happening around them, establishing informal conversations. Nevertheless,

³⁸ Video by Daniel Miller (May 3, 2020), "How to conduct an ethnography during social isolation", available at <https://www.youtube.com/watch?v=NSiTrYB-0so>, last accessed on April 19, 2024.

researchers in digital spaces still need to, on the one hand, be as open as possible about their role as researchers and, on the other hand, they need to make sure to maintain the anonymity of any data collected and used.

Further, the increased reliance on internet platforms and applications to gather research data means that researchers will increasingly confront terms of use that dictate the conditions under which such research activities can take place, if at all (Franzke et al, 2019). This means that digital research must also consider the specific terms and conditions of the different platforms where research is conducted. Further, different platforms have different user cultures that lead to different ethical implications. Depending on users' personal perception of privacy, the role of social norms, and the governing of the platforms, people may perceive research as intrusive in differing degrees³⁹ (e.g., Beninger et al., 2014; Patterson, 2018).

Following these guidelines and considerations, I defined the following ethical protocols for the digital fieldwork of my investigation:

- In my Twitter bio, I am open with being a social anthropologist who studies online gender-based hate speech.
- I have only collected and analysed material written on public (open) profiles, which anyone⁴⁰ can read without following the user. In terms of comments on the posts by the persons in the online cases, which are also an important part of the digital data to understand interactions and responses, I have only used material from *public* profiles. That is, again those accessible without following the user who made the comment. This is based on AoIR's guidelines as well as Twitter's own terms and conditions.⁴¹

³⁹ Compare e.g. an open thread on Twitter with a closed group on Facebook.

⁴⁰ When I started this research, open profiles could be accessed by anyone on the Internet. However, currently on X (formerly Twitter) a user profile on X must be registered in order to access also most of the public profiles.

⁴¹ Twitter's terms of service for third-parties requires that they "respect users' control and privacy" by deleting any "content that Twitter reports as deleted or expired", as well as any content that has been changed from public to private.

- I have anonymised the cases collected through the digital ethnography, by providing each case with a code (e.g., DE-1), whilst also deleting any names and other identifiable information when discussing the cases.
- To further ensure anonymity, I do not use entirely literal quotes from the digital ethnography in this thesis or any publications based upon my PhD research, as these might jeopardise anonymity in an online context through a simple google search. Quotes have also been translated to English.

5. The legal framework: Overview, interpretations, and gaps

What is not quantified doesn't exist and you can never justify it as an issue of attention from a legal or political perspective (EX-15)

5.1. Introduction

In March 2021, the OSCE Office for Democratic Institutions and Human Rights (ODIHR) launched a factsheet on Gender-Based Hate Crimes, clearly showing the aim to encourage OSCE states to prosecute gender-based crimes as hate crimes. According to ODIHR, "gender-based hate crimes are criminal offences motivated by bias against a person's gender" and,

One of the motivating factors behind this type of crime is the perpetrator's perceptions of gender norms. The victims of such crimes are often targeted due to their perceived deviation from gender norms, including on the basis of their sexual orientation and gender identity. Gender-based hate crimes may also target people or property due to their association, professional affiliation with or activism on gender issues, such as women's rights groups and civil society organizations working with victims of violence (ODIHR, 2021).

Any individual or group, regardless of their gender, can be a victim of gender-based hate crime, however, gender-based hate crime disproportionately affects women and girls (Ibid.). Victims may be selected solely due to their gender, or

based on multiple identity traits, such as gender and religion. As I have argued in Chapter 2, I include online hate speech on the grounds of gender within the conceptual framework of hate crimes. In this chapter, I will look at how this can be addressed from a legal perspective, with a focus on the legal framework available to address hate speech.

From an anthropological point of view, legal praxis is inevitably mediated by the set of values, norms and symbols that constitute the way of life of the society to which the legal tools belong (Canyelles, 2018). In this sense, both the legal framework itself and the implementation of this framework are deeply imbued in culture. Following Laura Nader (2002b), a user theory of law would reject the illusion that the law is a neutral and autonomous reality, thus considering that it is a vehicle for both different interests and processes of legitimation. Nevertheless, in our legal culture, it is generally taken for granted that a legal decision is not arbitrary, and the objectivity of the people who interpret and implement the laws is usually not problematised. Similarly, the subjectivity of legal operators as social agents is often ignored (Canyelles, 2018). As a social anthropologist lacking formal legal training, my understanding of the multitude of differing interpretations of the legal framework has been greatly enhanced through interviews and informal discussions with legal experts (including legal scholars and lawyers), police officers, and prosecutors. These interactions have facilitated the interpretation of the legal framework and enhanced comprehension of its application concerning gendered online violence in contemporary Spain. Furthermore, these interviews have served as essential groundwork for contemplating the criminal boundaries of gender-based online hate speech in collaboration with other professionals.

This chapter is located halfway between a theoretical and empirical chapter. It does not aim to offer an exhaustive analysis of the legal framework but rather serves as an exploration of legal responses available to victims of gender-based online violence. The intention is to provide both an overview of available legal mechanisms for those targeted by such violence and to establish a contextual backdrop for the subsequent chapter, which delves into professionals'

perceptions of gender-based online violence and hate speech, along with their readiness to provide support.

The subsequent sections, thus, aim to provide a foundational understanding of the application of hate speech laws. Initially, I will briefly address the fundamental rights implicated when employing legal tools to combat hate speech. Following this, I will present an overview of the principal legal framework in Spain. This includes an examination of hate crimes and hate speech, encompassing relevant jurisprudence. Additionally, consideration will be given to the framework pertaining to gender-based violence, regional laws, and other legal instruments. Finally, I will explore diverse interpretations of these frameworks and identify gaps based on insights gained from the interviews conducted with legal experts and other professionals.

5.2. Freedom of expression versus the right to dignity

From a legal perspective, hate speech and hate crime are two separate concepts and are as such constructed differently. Generally, hate speech legislation will criminalise "incitement to hatred" or "stirring up hatred", offences that have a much longer statutory pedigree than hate crime (Schweppe and Perry, 2021). To categorise an offence as a "hate crime" requires that a criminal offence known to the law was committed, with, for example, a racist motivation. "Incitement to hatred", on the other hand, is an offence in itself, with no "non-hate" version available (Ibid.).

Discussions regarding hate speech from a legal standpoint commonly revolve around the boundaries imposed on freedom of speech—an explicit fundamental right outlined in both the Universal Declaration of Human Rights (UDHR) and the International Covenant on Civil and Political Rights (ICCPR), closely intertwined with freedom of opinion or ideological freedom. Despite being unequivocally acknowledged as a paramount element of human rights and fundamental freedoms, freedom of expression is not absolute and is subject to diverse limitations. Examples of such restrictions are respect for others' rights or reputation, national security, public order, or public health (Paulos and Çelik,

2021). In the European tradition, post-World War II, the pre-eminence of freedom of speech underwent a shift, with human dignity assuming a more prominent position (Krotoszynski, 2004). The matter at stake has been to find a balance between these two basic pillars of human rights, namely the right to freedom of expression and the right to non-discrimination and human dignity, or expressed in another way, as effects on the victim versus the right to free speech (Matsuda, 1993: 2321).

The harm of hate speech includes not just an increased threat of violence and discrimination, but also “a jolting failure or undermining of the assurance that people need to rely on: the assurance that they can go about their daily life and their ordinary business without fear of being denigrated and excluded as subhuman or second-class citizens” (Waldron 2012: 171). The human element under attack by hate speech is often defined as “‘dignity’, in the sense of their basic social standing, the basis of their recognition as social equals and as bearers of human rights and constitutional entitlements” (Ibid: 59). From a philosophical standpoint, one might argue that dignity is innate to humans. However, as a social and legal concept, it necessitates establishment, preservation and validation by both society and the law (Ibid.). This upholding of human dignity is something that all members of society are required to partake in. It should be highlighted that although this debate concerns group dignity, the point of reference is the individual members of the group, and not the dignity of the group as such or the dignity of the culture or social structure that holds the group together (Waldron, 2012). Attacking this dignity,

defamation is not just an idea contributed to a debate. In its published, posted, or pasted-up form, hate speech can become a world defining activity, and those who promulgate it know very well—this is part of their intention—that the visible world they create is a much harder world for the targets of their hatred to live in (Ibid.: 74).

Placing dignity at the center of this discussion, then, helps us keep the focus on discourse that is harmful and not just offensive. That is, the ultimate concern, according to this perspective, is what happens to individuals when defamatory imputation is associated with shared characteristics such as “race”, ethnicity,

religion, gender, sexuality, and national origin. Or, as Catharine MacKinnon (1993: 99) puts it:

We argued that group libel ... promotes the disadvantage of unequal groups; ... that stereotyping and stigmatization of historically disadvantaged groups through group hate propaganda shape their social image and reputation, which controls their access to opportunities more powerfully than their individual abilities ever do.

The issue of whether legal measures should address hate speech, and to what extent, has been a matter of intense debate. In Europe, the prevailing perspective is that restrictions on hate speech not only align with human rights standards but are actively mandated by these standards. In the US, on the other hand, the stronger stance is that all hate speech bans violate the First Amendment (Phillipson, 2015). Nevertheless, scholarly positions do not neatly align with geographical distinctions, as some scholars in the US express, at least in principle, support for hate speech bans, while certain European scholars, influenced to some degree by North American arguments, are critical of the scope, if not the existence, of many hate speech bans (Ibid.).

Hate speech laws are commonly defended by reference to the harm that hate speech causes. As we have seen above, this can be defined as the harm inflicted by hate speech upon human dignity. Hate speech regulation, thus, can be understood as the protection of “a certain sort of precious public good”: a visible assurance offered by society to all its members that they will not be subject to abuse, defamation, humiliation, discrimination, and violence on grounds of “race”, religion, and so on (Waldron, 2012). That is, laws regarding group defamation are

set up to vindicate public order, not just by preempting violence, but by upholding against attack a shared sense of the basic elements of each person’s status, dignity, and reputation as a citizen or member of society in good standing—particularly against attacks predicated upon the characteristics of some particular social group (Ibid.: 47).

This means that legal tools against hate speech are not just a matter of protecting people from sporadic insult, offense, and wounding words, but rather

of securing, in a systematic fashion, a particular aspect of social peace and civic order under justice: the dignity of inclusion and the public good of mutual assurance concerning the fundamentals of justice (Ibid.: 103).

The call for legal sanctions against hate speech can also be understood as a clear message from the authorities that the dignity of vulnerable groups is worthy of legal protection, as “the choice of public sanction, enforced by the state, is a significant one” (Matsuda, 1993). As Mari Matsuda has argued, the places where the law does not go to redress harm tend to be the places where those most vulnerable live (Ibid.). This absence of law is itself a story with a message, perhaps unintended, about the relative value of different human lives. Thus, a legal response to hate speech can be understood as a statement that victims of hate speech, whether racist, anti-LGBT, gender-based or based on intersections of several grounds, are valued members of our community, as “tolerance of hate speech is not tolerance borne by the community at large. Rather, it is a psychic tax imposed on those least able to pay” (Matsuda, 1993: 18).

Returning for a moment to the hate crime framework, this message that society does not tolerate attacks motivated by bias tends to be a central justification of hate crime laws (Stanford Law School, 2021). However, the challenges contained in investigating, charging, and proving hate crimes may undermine this expressive function. Failures within the legal system to investigate hate crimes might in the end communicate quite a different message: that protecting people from hate crimes and hate speech is not a priority (Ibid.). In this regard, Mark Walters (2019) refers to the gap between the estimated total of hate crimes committed and those convicted with a sentence uplift as a hate crime, as the “justice gap”. Based on hate crime statistics collected in the UK, there are two fundamental arguments against the approach to combat hate crimes solely through enhanced punishments (Ibid.). First, the enhancement of punishment and additional criminalisation of hate-motivated perpetrators does little to repair the actual harms caused by hate incidents and crimes; and second, enhancing the penalties of offenders is unlikely to challenge in an effective way the underlying causes of prejudice. That is, there is little evidence to show that hate crime laws

generate any meaningful reparative benefits directly to victims, or to society by reducing overall levels of hate crime offending (Ibid.).

Similarly, there is an absence of clear empirical evidence demonstrating the tangible effectiveness of hate speech bans in preventing harm to vulnerable groups or mitigating the damage inflicted (Heinze, 2016). Neither is there hardly any empirical evidence that hate speech bans are effective in minimising harmful messages and decreasing discrimination or hate motivated violence (Phillipson, 2015). In instances where the speech is offensive or provocative, the labelling of an expression as “hate speech” could overstate the connection between the speech and the alleged harm—either by misjudging the influence of a speaker or the likelihood of harm occurring—or overlooking the propensity of individuals to engage in effective counter speech (Gagliardone, 2019). Hate speech laws may also lead people to express the same basic ideas, but in less crude language (Phillipson, 2015). This, in turn, may be counterproductive: “Given that the use of crude racial epithets is likely to put many readers off, a more apparently reasonable - because more moderate-sounding argument - might actually have wider appeal, thus possibly causing the harmful ideas themselves to gain wider currency” (Ibid.:7).

Hate speech laws can also have other counter-productive effects: prosecutions may turn a thus far obscure and despised group into “free speech martyrs”, shifting public focus from the prejudiced nature of their views to the alleged injustice of their prosecution by criminal law (Ibid.). As a further negative effect, prosecution of hate speech may result in greater publicity of the original message, as it is amplified and repeated in media and on social media. Importantly, hate speech laws may also miss their intended targets and instead be used against historically oppressed minorities. For example, there were almost as many prosecutions of black power activists as of white racists under the UK’s race hate laws (Ibid.). Similarly, before the Ethiopian elections of 2015, a group of bloggers seeking to pressure the government to respect fundamental freedoms included in Ethiopia’s own constitution were arrested with the accusation of producing speech that could destabilise the country (Gagliardone and Pohjonen,

2016). Thus, there is a tension surrounding the concept of hate speech and its regulation:

Hate speech regulation imagines itself as simply enforcing the given and natural norms of a decent society [...]; but from a sociological or anthropological point of view we know that law is always actually enforcing the mores of the dominant group that controls the content of law (Post, 2009: 130).

Further, hate speech laws do little to address the real problems and underlying structures of hate speech (Gelber, 2002). First, because they counterpose the goals of securing free speech and ameliorating the harms of hate speech. Secondly, these laws tend to invoke a private and individualised resolution process for what is essentially a public problem, which means that there is little scope for directly addressing the broader harm inflicted by hate speech. Thirdly, these laws tend to rely on a statutory definition of hate speech, which is either too vague, and thus makes laws difficult to enforce, or too restrictive.

However, there are middle grounds in the debate surrounding the legal regulation of hate speech. Laia Serra (2018b), in this regard, refers to a European consensus on the categorisation of hate speech according to three levels: First, non-punishable “hate speech”, which is harmful or disturbing speech in terms of coexistence and tolerance, but that does not deserve any type of sanction; Second, “hate speech” of medium intensity, which deserves a civil or administrative sanction, and third, serious “hate speech”, which deserves a criminal sanction. In a similar vein, Gavin Phillipson (2015), while raising concerns about hate speech laws and acknowledging that certain existing bans on hate speech may have had unintended negative consequences, contends that this does not automatically imply that all bans are ineffective or counterproductive. A starting point when arguing for a narrower approach to hate speech bans could be “the widely-shared insight that hate speech that urges people to regard the target group as non-human, or less than human, is in principle the most dangerous and destructive idea that can be uttered” (Phillipson, 2015: 9). Consequently “bans on such speech seek to enforce a fundamental premise of the social contract, or of deliberative democracy, namely mutual recognition

extended by each to all of a shared citizenship and humanity” (Ibid.: 11). From this standpoint, hate speech laws should encompass three primary categories of speech acts. Firstly, speech explicitly or implicitly denying the equal humanity of the targeted group, exemplified by the use of dehumanising terms like "germs", "cockroaches" or "vermin." This extends beyond overtly insulting language or racial epithets to include more subtle arguments portraying a specific group as less human. Secondly, such laws should address speech that advocates for violence against an identifiable group. Thirdly, they should cover speech that fails to acknowledge members of a given group as citizens, effectively calling for the complete deprivation of their ordinary constitutional rights—manifested in expressions advocating mass deportation, enslavement, property confiscation, and similar actions (Ibid.). In essence, this middle-ground approach is distinctly centred on addressing the potential harm inflicted by hate speech.

The digital context of online hate speech adds a further layer of complexity to the application of hate speech laws. For example, there are challenges connected to the place of crime, which concern the location of the crime and the legal status of the social media company in the country in which the service is delivered (Paulos and Çelik, 2021). While a derogatory social media post might be deemed an online hate crime in the nation of its origin, the same may not hold true in the jurisdiction where the online service was rendered. Further, in contrast to traditional media, identifying the authors of internet content is often difficult. Content produced in a particular country has the potential to be replicated, altered, and disseminated across geographic borders, and it may find hosting in various countries, each governed by its own distinct set of regulations (Ibid.) Notably, the legal prosecution of hate speech cases differs between mainstream websites and social media platforms, especially within the European Union, where servers hosting websites are legally mandated to furnish information (Jubany and Roiha, 2018). Nonetheless, despite the possibility of international collaboration concerning web pages, the existing system is perceived as considerably inefficient (Ibid.). Consequently, the challenge of tracking individuals on the internet is more of a bureaucratic matter than a technical one, contingent

on the willingness of service providers to collaborate.

5.3. Hate crimes and hate speech in the Spanish criminal code

As I have reflected upon previously in this chapter, different interpretations exist among scholars regarding the kind of speech that hate speech laws should address. From a legal point of view, there is often a fine balance between what can be defined as hate speech and what cannot be defined as such and thus addressed under this framework. There is agreement among legal scholars that not all hate speech deserves criminal sanctions and that, for example, a civil or administrative sanction could come into play (e.g., Phillipson, 2015; Heinze, 2016; Serra, 2018).

In Spain, the Organic Law 1/2015, amending the Penal Code, was enacted on March 30, 2015.⁴² This law provides for several amendments to the penal code. With relevance to this thesis, it amends Article 510, which criminalises discriminatory speech by adding a criminal offence of negating, minimising or praising publicly the commission of or perpetrators of the crimes of genocide, crimes against humanity or crimes against protected persons and property in times of armed conflicts committed against a specific group or against members of such group. In its current version, since 2015, Article 510.1a of the Spanish criminal code refers to publicly encouraging, promoting or inciting, directly or indirectly, hatred, hostility, discrimination or violence against a group, part of it or against a specific person because of their membership of that group, for racist, anti-Semitic or other ideological reasons, religion or beliefs, family situation, the belonging of its members to an ethnic group, "race" or nation, national origin, sex, gender, sexual orientation or gender identity, or for reasons of illness or disability. Article 510.2a further refers to injuring the dignity of people through actions that involve humiliation, contempt or discredit of any of the groups referred to in the previous section, or of a part of these groups, or of any person determined by reason of belonging to these groups for racist, anti-Semitic or

⁴² Organic Law modifying the Penal Code, 2015, *International Humanitarian Law Databases*, available at <https://ihl-databases.icrc.org/en/national-practice/organic-law-modifying-penal-code-2015>, last accessed on April 19, 2024.

other reasons referring to ideology, religion or beliefs, family situation, the membership of its members to an ethnic group, "race" or nation, national origin, sex, sexual orientation or gender identity, for reasons of gender, illness or disability. 510.2b refers to producing, elaborating, possessing in order to distribute, provide third parties with access, distribute, disseminate or sell writings or any other kind of material or supports that, due to their content, are suitable to damage the dignity of persons because it represents a serious humiliation, contempt or discredit of any of the mentioned groups, of a part of them, or of any person determined by reason of their belonging to these groups. In addition to art. 510, article 22.4 regulates the aggravating circumstances of crime, in which gender, sexual orientation and sexual identity are considered as particular circumstances, among others, which negate the principle of equality. That is, for a crime to be prosecuted as a hate crime it must be categorised as a crime in the criminal code, and with an aggravating circumstance. Alternatively, it must correspond to a crime categorised within the previously mentioned article 510 CP on incitement to hatred (or, as it is more commonly denominated, hate speech).

The particularity of the crime addressed by article 510 is that the injury – as I have discussed in section 5.2 above – is caused by conduct carried out by the author in the exercise of a fundamental right, namely freedom of expression (Lizardo-González, 2021). The legal good is located in human dignity, but with a social or collective dimension, and in the right to equality and non-discrimination (Landa Gorostiza, 2000: 217, cited in Lizardo-González, 2021). This fine and often complex balance between the right to freedom of speech and the right to dignity exists in relation to all hate speech laws. However, the Spanish art. 510 has some additional complexities due to its specific composition and wording. In this regard, the current wording of art. 510 moves away from the provisions of the EU Framework Decision 2008/013/AJ, as it extensively expands prohibited conduct. Further, the wording of art. 510 is ambiguous, with indeterminate legal concepts that cause legal uncertainty in its application. This way of writing, together with the lack of authorised and consolidated jurisprudential interpretations, opens the

door to an incorrect, arbitrary, and sometimes contrary application of the criminal concepts (Lizardo, 2023). Critique in this regard was also brought up in the interviews with legal experts, which I shall reflect upon in section 5.7, below. There, I will take a closer look at the interpretation and implementation of the legal framework from the perspective of the legal experts interviewed, with an emphasis on article 510. Nevertheless, to begin clarifying how article 510 is currently being applied in Spain, in the subsection below I will conduct a brief review of jurisprudence of this article linked to its application in cases of hate speech on the grounds of gender in online spaces.

5.3.1. Jurisprudence linked to article 510

The unclear wording of art. 510 and the generalised lack of knowledge of what kinds of expressions can be considered hate speech from a legal point of view may contribute to reinforcing the underreporting of online hate speech. The lack of dissemination of the convictions under the legal framework further contributes to this lack of awareness and lack of response (Lizardo, 2023). In this sense, analysing jurisprudence contributes to facilitating a better understanding of the balance between the limits between freedom of speech and hate speech from a legal perspective.

An examination of jurisprudence, academic literature, and various reports has uncovered only two instances of online hate speech convictions under Article 510 specifically based on gender between the introduction of gender as an aggravating circumstance in Spanish criminal law in 2015 and the completion of the current research in 2023.⁴³ The first sentenced case exclusively concerns gender, whilst the second case includes publications that incite hatred on several grounds, including religion, nationality, sexual orientation and gender. In the next sections, I will provide an overview of these cases.

Cases prosecuted and sentenced under art.510 with a gender motivation

Case 1: STS 396/2018 Supreme Court

⁴³ The review of jurisprudence was conducted on the official online platform Centro de Información Judicial (www.poderjudicial.es), using the search phrase “510 Y discurso de odio Y género Y internet”, filtering the results for the years 2015, 2016, 2017, 2018, 2019, 2020, 2021, 2022, 2023.

The case STS 396/2018⁴⁴ was brought to the Supreme Court after an appeal by the representative of the prosecuted party. The case can be summarised as follows: On January 26, 2017, the Fourth Section of the National Court handed down a conviction against the accused, Miguel, for a crime of incitement to hatred (art. 510 CP), derived from the publication by the accused of a series of messages with illicit content through Twitter. Among the proven facts that motivated this ruling, Miguel had been the owner of two accounts on this social media platform for several years, with around 2.000 followers in total, through which he disclosed a series of posts containing denigrating comments towards the group of women, in which he also celebrated and incited the commission of acts of violence against them. More specifically, he published the following messages through the first of his accounts:

- December 17, 2015: "53 murdered by sexist gender violence so far this year, it seems few to me with all the whores that are loose"⁴⁵
- December 30, 2015: "And 2015 will end with 56 murdered women, it is not a good mark, but we did what we could, let's see if we double this figure in 2016, thank you"

As a result of these publications, the Police Unit *Grupo de Redes II* received various e-mails from citizens reporting the comments, and two people even filed formal complaints at two different police stations. In parallel, on January 7, 2016, Twitter suspended the defendant's first account, which made him use his second account to publish further posts, with similar content:

- January 14, 2016: "Beatriz was a feminist and she jumped into the river because women get wet for equality."
- January 14, 2016: "I like to fuck against the counter and the stove, because I put the woman in her place twice"
- January 16, 2016: Sharing the image of a woman, regarding whom it is unknown if she has been a victim of abuse or gender violence, the accused published the comment "I have already battered her, you're up next"

The Supreme Court sentenced Miguel as the responsible author of a crime of incitement to hatred of article 510, sections one and three, to the penalty of 2

⁴⁴ STS 396/2018 - ECLI:ES:TS:2018:396, available at <https://www.poderjudicial.es/search/AN/openDocument/6c9b9b325578e688/20180219>, last accessed on April 19, 2024.

⁴⁵ Own translation from Spanish.

years and 6 months in prison and a fine of nine months with a daily fee of €40. The Court clarified that Article 510 CP penalizes individuals who promote discrimination, hatred, or violence against specific groups or associations based on the criteria outlined in the provision. It emphasized that the fundamental element of such criminal acts involves the articulation of epithets, qualifiers, or expressions conveying a generic message of hatred. The Court further asserted that the offense of incitement to hatred does not necessitate a specific intent but relies on the presence of a basic malice discernible from the content of the expressions employed. The intent of these crimes is established by confirming the voluntary nature of the act and ensuring that it is not an uncontrolled situation or a momentary, even emotional, reaction to a circumstance beyond the individual's control. In this sense, the factual account refers to the publication of the messages on different dates, which makes the conduct voluntary and not a reaction to an external stimulus. On the other hand, the content of the sentence reveals the aggressive character of the expressions and the verification of hatred when referring to situations in which the accused wants to place women, who are referred to in aggressive terms in a gender-based context (art. 510 1a).

Case 2: STS 2085/2022 Supreme Court

The second case, STS 2085/2022⁴⁶ entails publications by Antonio, also known as Argimiro, who used three different profiles on Facebook to publish a several posts against different groups. The Provincial Court of Madrid, initially handling the case, categorised the various posts based on the groups they targeted. In other words, they organised them according to the specific grounds or protected characteristics that were the focus of the discourse's attack. The groups that were targeted by these messages included: other branches of Islam than the one that the defendant had converted to (religion); Spanish people and in particular people from Andalucía (nationality/origin); LGBT+ persons (sexual orientation); people from the political “right-wing” (ideology); and finally, women as a group (gender). The defendant published several posts in relation to each of these

⁴⁶ STS 2085/2022 - ECLI:ES:TS:2022:2085, available at <https://www.poderjudicial.es/search/AN/openDocument/955b1a273d339e12/20220608>, last accessed on April 19, 2024.

groups. In relation to the collective of most relevance this research, namely women as a group, the sentence describes nine different posts that were published on seven distinct dates. These are exemplified as follows with extracts from four of the posts⁴⁷:

- July 20, 2017: The rebel woman / You hit her / The rebel woman/ you slap her a couple of times
- July 20, 2017: Whores and shitty drunks / They are the Spanish whores / They eat multiple dicks / They're not even valid as slaves
- August 13, 2017: (...) If a man obeys to feminist indications / It will be the man who suffers / Physical violence / Economic violence / Family violence / I'm tired of that feminine egocentrism that makes the harm that women suffer / more important than the harm that men suffer / I think I haven't murdered an ex-partner who kidnapped my daughter /Because of the enormous quantities of joints that I smoke / But if instead of smoking joints / I drink whiskey / I would have run over her or beat her up or shot her or anything to /end her drunk dirty cocaine addict existence.
- August 30, 2017: Whether it's a man or a woman / If they attack me, I defend myself / Violence against women is good
- January 1, 2019: You feminists are whores who eat pig / and you are what you eat / You try to destroy the biology of man forcing him to be cold in situations where you don't wear proper clothes that cover your erogenous zones / In conversations you don't put any limits / But the man has to stay cold / Is that what you want whores? To be as slutty as you can without suffering the consequences? Is that what you want to be, whores?

Antonio was sentenced as the criminally responsible author of a crime against the fundamental rights of art. 510.1a and 3 of the Penal Code, to the penalty of two years and six months in prison and a fine of nine months and one day with a daily fee of two euros. The sentence was appealed and later tried and confirmed in the Supreme Court. In regard to the expressions of gender-based hate speech, together with the hate speech published against the other groups, the Supreme Court stated that this behaviour does not concern the use of social media as a simple vehicle to express a merely hurtful or offensive opinion, but that some of the passages - alone or interrelated with the rest - reflect that the message invites to action, violence, and armed struggle. The Court further stated that Antonio not only incites others, but he offers himself as the first aggressor of those whom he despises because of their ideology, their gender, their sexual

⁴⁷ Own translation from Spanish.

orientation or their national origin. The Court concluded that it cannot protect this hate speech, which invites online users to join in the violence and which proposes hitting women as a model of coexistence.

Cases prosecuted under art.510 and not sentenced

Article 510 has also been tried for incitement to hatred based on gender in a few other cases, but without sentence. One case⁴⁸ involved threats against a woman on Facebook, including, amongst others, the following expressions: "If you do that to my daughters, I won't break your phone. I'll break your head, you disgusting bitch," "Next time her head is broken... if she plays the victim... let it be for good reason... ". Here, the court interpreted that this would not amount to incitement to hatred under art. 510, as

these are expressions addressed to a specific person and as a result of a previous incident carried out by the same, without it being possible to estimate that it integrated a discourse whose purpose was only to propagate violence or promoting hatred against a certain group, even though the complainant may belong to that group and this without prejudice to the fact that, as we have already indicated, the expressions discharged through the social network FACEBOOK could constitute a minor crime of threats and/or harassment.⁴⁹

Both the cases tried and prosecuted and those that were tried without sentence highlight, first, the need for the hate speech to include *generic* hatred expressed towards the group of women to be defined as hate speech in a legal sense. Nevertheless, as reflected in the court's statement in the case above, this hatred towards the collective of women could be integrated in an attack that targets a specific individual. Second, the cases underline that there is no need for a specific intention to cause physical harm to the collective in order for the speech to reach the threshold of a criminal act. That is, the aggressive character of the expressions against the group is sufficient to limit the freedom of expression and sentence this type of expressions as incitement to hatred. Third, the cases show that the intent of these crimes is also proved by the voluntary nature of the act

⁴⁸ AAP L 882/2019 - ECLI: ES:APL:2019:882A, available at <https://www.poderjudicial.es/search/AN/openDocument/6c9b9b325578e688/20180219>, last accessed on April 19, 2024.

⁴⁹ Own translation from Spanish.

and the verification that it is not a one-time uncontrolled situation or a momentary or emotional reaction before a circumstance that the defendant has not been able to control. Instead, the cases that have been sentenced consist of several publications of generic hatred against the same collective on different occasions, and in the sentenced cases exposed here, even across different social media accounts and platforms.

Whilst, as previously mentioned, the application of art. 510 will be further discussed in relation to the interviews with legal experts in section 5.7, in the next section, I will provide a brief overview also of the legal framework related to gender violence.

5.4. The legal framework on gender violence in Spain

The Istanbul Convention establishes that violence against women is a

violation of human rights and a form of discrimination against women and shall mean all acts of gender-based violation that result in, or are likely to result in physical, sexual, psychological, or economic harm or suffering to women including threats of such acts, coercion or arbitrary deprivation of liberty, whether occurring in public or private life.⁵⁰

In this convention, "gender-based violence against women" is understood as "violence that is directed against a woman because she is a woman or that affects women disproportionately". As the Convention further states, the acts against a woman with the purpose to discriminate, dominate or subjugate her, can be carried out in any context, that is, including digital spaces.

The Istanbul Convention entered into force in 2014 and has been signed and ratified by the Spanish state. However, the concept of gender violence has been erroneously incorporated in the Spanish legal framework, as the *Organic Law 1/2004, of December 28, on Comprehensive Protection Measures against Gender Violence*, only contemplates gender violence in those cases where the man who commits the violence maintains or has maintained a romantic relationship with the woman (Marin de Espinosa, 2018). That is, the ratification of the convention

⁵⁰ Council of Europe Convention on preventing and combating violence against women and domestic violence (CETS No. 210), available at <https://www.coe.int/en/web/conventions/full-list?module=treaty-detail&treatyid=210>, last accessed on April 19, 2024.

has not included an adaptation of Spanish laws to the concept of gender-based violence. Thus, it can be understood as an incomplete law, which understands gender violence as a problem reduced to the private sphere, and which, consequently, does not recognise gender-based violence that is exercised in the public sphere (Ibid.: 13). Spanish law, thus, does not define as gender violence other forms of violence such as forced marriages, genital mutilation, sexual violence and abuse, or digital violence (outside of the couple).

In terms of the existing legal framework, gendered digital violence can be reported under several articles, which are not specific to violence in the *digital* sphere, but which can nevertheless be applied (Igareda et al, 2019). Here, it is important to consider that scholars who have analysed what articles can be applied tend to include a broad concept of all possible types of online violence against women, and not only verbal violence (e.g., Igareda et al, 2019). Different forms of gender-based online violence fit into different legal typifications of violence, but these are not always identified as sexist- or gender-based violence (Ibid.). The regulations applicable to gender-based online violence can be divided into two categories: those falling under the legal framework of gender violence (pertaining to cases involving an existing or past relationship between the victim and the offender) and those outside this framework, which are not classified as “gender violence” in Spain (Ibid.). The first category, that is, within the context of intimate partner violence, includes Articles 147 and 1489 - Injuries; Article 153 - Physical or psychological abuse; and Article 171 - Threats. In the second category, that is, outside of a relationship and thus outside of the Spanish criminal code’s definition of gender violence, we can include the following articles: Art. 171.7 - The minor crime of minor threats; Art. 172.3 - The minor crime of coercion; Art. 173 - The crime against moral integrity; Art. 178 - The crime of sexual assault; Art. 184 CP - The crime of sexual harassment; Art. 172 ter - Stalking; and Art. 197.7 CP - The dissemination without consent of intimate content.

Further, in the Organic Law 10/2022, of September 6, on the comprehensive

guarantee of sexual freedom⁵¹, popularly known as the “solo sí es sí” (“only yes is yes”) law, the crime of digital identity theft on the Internet is introduced (new art. 172 ter 5 CP) as follows: "Whoever, without the consent of its owner, uses the image of a person to make advertisements or open false profiles on social networks, contact pages or any means of public dissemination, causing the same situation of harassment, harassment or humiliation". This could, at least in theory, be applied to some of the violent situations analysed in the present research. Nevertheless, due to the relatively recent character of this new law, a review of jurisprudence has not been conducted.

5.5. Laws from autonomous communities

There are also examples of regional legislation in Spain, with relevance to gendered violence and with some mentions of “digital violence”; however, these do not imply changes in the criminal code, but rather act at a civil level.

During the first stages of research for this thesis, the regional law against gender violence in Catalonia was modified in *Law 17/2020, of December 22, modifying Law 5/2008, of the right to eradicate male violence*. In the Catalan context, the definition of gender violence was already before this reform broader than in the Spanish-wide criminal law. Nevertheless, this reform has led to some important changes by including the concept of digital violence within the different forms of gender-based violence that the law considers. This law includes an explicit reference to digital violence, which is defined as

acts of sexist violence and online misogyny committed, instigated, amplified or aggravated, in whole or in part, with the use of information and communication technologies, social networking platforms, websites or forums, email and instant messaging systems and other similar means that affect the dignity and rights of women. These acts cause psychological and even physical damage; they reinforce stereotypes; damage dignity and reputation; violate the privacy and freedom of action of women; cause economic losses, hindering their political participation

⁵¹ *Ley Orgánica 10/2022, de 6 de septiembre, de garantía integral de la libertad sexual*, available at <https://www.boe.es/eli/es/lo/2022/09/06/10/con>, last accessed on April 19, 2024.

and freedom of expression.⁵²

Likewise, in Galicia, *Law 15/2021, of December 3, which modifies the Galician Law 11/2007 of July 27, for the prevention and comprehensive treatment of gender violence*, has incorporated digital violence as an additional typology in gender violence, with the following definition:

Digital gender violence or online violence against women, which includes any act or conduct of gender violence committed, instigated or aggravated, in part or in full, by the use of new information and communication technologies (ICT), such as the Internet, social network platforms, messaging and email systems or geolocation services, with the purpose of discriminating, humiliating, blackmailing, harassing or exercising dominance, control or interference without consent in the privacy of the victim; regardless of whether or not the aggressor has a marital, partner or similar affective relationship in the present or in the past, or kinship with the victim. Likewise, acts of digital violence against women will be considered those carried out by men from their family, social, professional or academic environment.⁵³

In the autonomous community of Catalonia there is also, since 2014, a regional law to uphold the rights of LGBT people, *Llei 11/2014, del 10 d'octubre, per a garantir els drets de lesbianes, gais, bisexuals, transgènere i intersexuals i per a erradicar l'homofòbia, la bifòbia i la transfòbia*. Through the *Àrea per a la Igualtat de tracte i no-discriminació de persones LGBTI* of the *Direcció General d'Igualtat*, the Generalitat de Catalunya collects civil infractions, including those taking place online. In section 5.7 of this chapter, I shall discuss the legal professionals' perceptions also in relation to regional laws.

5.6. Other tools and mechanisms

There are also other tools and protection mechanisms, with relevance for gender-based online hate speech or gender-based online violence in a broader sense, which I will briefly highlight in this section.

⁵² Own translation from Catalan. The full text in Catalan is available at <https://portaljuridic.gencat.cat/ca/document-del-pjur/?documentId=889760>, last accessed on April 19, 2024.

⁵³ Own translation from Spanish. The full text in Spanish is available at <https://www.boe.es/eli/es-ga/l/2021/12/03/15>, last accessed on April 19, 2024.

5.6.1. Detection and intervention protocol for victims of digital gender violence in Andalusia

Since 2015, the Andalusian regional government has a detection and intervention protocol for victims of digital gender violence, which is defined as "gender violence that is carried out taking advantage of ICTs" (Instituto Andaluz de la Mujer, 2015). The protocol seeks to facilitate the detection and registration of technology-facilitated gender violence as well as to incorporate this view in the actions of professionals who provide direct attention to victims (Vergès, 2017). Positively, the protocol does not only refer to gender violence as defined by Spanish law, i.e. violence within the (ex)-couple, but lists a range of behaviours, including discovery and disclosure of secrets (data theft of data, photos/videos, accounts, profiles); injuries, slander through ICT; denigrating treatment through ICT; dissemination of degrading or harmful images through ICT; identity theft; fraudulent use of cards; damages (equipment, companies, professional activity); grooming and crimes related to corruption and prostitution of minors; induction to the abandonment of the home to minors through the Internet; induction to suicide through the Internet; threats and computer coercion; sextortion; child pornography with use of ICT. Importantly, the law text also mentions that "At the collective level, there is also a crime that is increasing through ICT: The apology of discrimination and gender violence" (Ibid.). This implies the use of a broader concept of gender violence, which would include what the present project refers to as online gender-based hate speech.

5.6.2. Right to be forgotten

Among mechanisms of a civil nature, where there is not a criminal case, the protection of personal data can be highlighted. This is relevant predominantly in cases of dissemination of images or distribution of other personal data, for example doxing of private addresses. In these cases, the right to be forgotten can be exercised, in accordance with the *European Regulation 2016/679 of the Parliament and of the Council of 27 April 2016*, known as the General Regulation of Data Protection. The exercise of the right to be forgotten can be used as an instrument to empower women who suffer from cyber-violence, without the need

to report to police or judicial authorities. That is, it is a tool that allows for the direct request for the deletion and/or non-indexing of specific content without previously placing a police report (Igarada et al, 2019). Nevertheless, there is very scarce information available regarding to what extent the right to be forgotten is used, and particularly in cases of gendered online violence, or to what extent the petitions for deletion of content are successful.

5.6.3. Protocol to combat illegal hate speech online

The Spanish *Protocol to combat illegal hate speech online*⁵⁴ is coordinated by the Observatorio Español de Racismo y Xenofobia of the Ministry of Inclusion, Social Security and Migrations (OBERAXE). It stems, on the one hand, from the *Code of Conduct*, signed in 2018 by the European Commission and several social media platforms, and, on the other hand, from the *Recommendation of the EU 2018/334 of the European Commission, of March 1, 2018, on measures to effectively combat illegal content online*.

The Protocol seeks to combat illegal hate speech online; establish a national focal point - the Computer Crime Unit of the State Attorney General's Office - as the interlocutor of the public administration with internet companies; accredit and train trusted flaggers (or trusted users); preferably process communications from duly accredited trusted flaggers; establish homogeneous notification circuits for hate speech; and finally, implement the protocol and its monitoring, within the framework of the Interinstitutional Agreement to combat racism, xenophobia, LGBTI-phobia and other forms of intolerance. Whilst gender-based online hate speech is not explicitly included in the protocol, this type of hate speech and its intersections with other factors, can be interpreted as included in "other forms of intolerance".

In the next sections, I shall discuss the interpretations and use of the legal framework, based on interviews with legal experts and, to some extent, other professionals. That is, in the following sections I will enter the space between the

⁵⁴ *Protocolo para combatir el discurso de odio ilegal en línea*, available at https://www.inclusion.gob.es/oberaxe/ficheros/ejes/discursoodio/PROTOCOLO_DISCURSO_ODIO.pdf, last accessed on April 10, 2024.

written law and its implementation.

5.7. Perceptions and interpretations of legal protection mechanisms for victims of gender-based online hate speech

The legal framework that is applicable to behaviour categorised as gender-based online hate speech, or more broadly speaking, online gender-based violence, can be defined as fragmented and piecemeal. Legal definitions of many of the aspects included in this type of violence are still lacking, and particularly in terms of applying a gender perspective. A first obstacle consists in categorising and naming gendered violence in the digital sphere:

There is a diversity of nomenclature to name the same phenomenon, so to speak. And of course, no denomination has an official or legal definition. Well, we are interested in the legal definition and of course you cannot find any legal instrument that defines misogynist cyber violence. Sometimes we find that they talk about cyberbullying, other times they talk about digital violence, but what worry us is that the gender dimension is often lost [EX-15].

That is, depending on the conceptualisation and categorisation of this violence, the recognition and visibility of the gender aspect varies. The obstacle to recognising the gender-based motivation in online violence is twofold. On the one hand, as we have seen above, the legal framework related to gender violence in Spain is limited to what is sometimes defined as “intimate partner violence”. On the other hand, gender-based hate crimes or hate speech tend not to be recognised as such, although gender is included as a protected characteristic in Spain's Penal Code. As the review of the currently existing jurisprudence has exposed, in the cases where hate speech is recognised, the cases that have reached a sentence concern generic attacks against the whole collective of women, as opposed to attacks against an individual on the grounds of gender (or on the grounds of a feminist ideology).

As previously mentioned, sex and gender are included as protected characteristics both under article 22 and article 510, however, whether gender-based hate speech can and should be prosecuted under article 510 is subject to

different interpretations. This is further discussed in the next section.

5.7.1. The applicability of Spanish hate speech laws on gender-based hate speech

When reflecting upon the different legal options to address gendered online hate speech, and searching for jurisprudence on cases of hate speech on grounds of gender, it becomes clear that article 510 has been scarcely applied in general, and even less in cases of incitement or promotion of gender-based hate. Taking a closer look at article 510, there are differing and sometimes conflicting views on how to interpret this article:

There has already been a tremendous legal discussion about whether what counts is the motivation. In fact, as 510 is articulated, it's a kind of strange mix, because on the one hand there is an international trend that defines hate crimes by discriminatory motivation, but there is another sector that leaves in second place the motivation, as it is quite difficult to prove many times, and instead puts on the table the effects on socially recognised groups as vulnerable [EX-06].

That is, to begin with, the general interpretation of article 510 is blurry or unclear. In drafting the article, the legislator does not include any limiting clause, nor do they link the prohibition of behaviours with the values and fundamental rights that it tries to protect (Lizardo González, 2021: 384). Likewise, the article does not incorporate the publicity requirement, essential to be able to punish only those conducts that effectively endanger any legal right worthy of protection. This situation results in a lack of application by the courts (Ibid.). The confusion around the correct application of this article was further reinforced by a circular disseminated by the General Attorney of the State in 2019.¹ Although the intent of this circular was to clarify the application of article 510, some of the legal experts interviewed argue that this circular only served to create more confusion, and to undermine the original purpose of hate crime laws, namely to protect vulnerable groups:

The 2019 prosecution circular no longer speaks of motivation, it speaks of intolerance (...) I consider that it's the formula that the prosecution finds to finally exterminate the political foundation of hate crimes. Then it's no

longer about categories or historical oppressions or sociodemographic justifications. It's about censoring intolerance and all radical critical discourses can be considered as such. And also, if we depoliticise it for historical or geographical reasons, hate speech of women against men could be tried because it's an intolerance towards a category or a lifestyle or identity based on masculinity. And this worries me because it's directly at odds with the international warnings of the ECRI and company, keeping an eye on how the fight against hate speech is being applied. Here in Spain the aggravating factor of ideological discrimination against antifascists who have attacked Nazis is already being applied... [EX-06]

The very purpose of hate speech laws, that is, to protect historically or potentially socially vulnerable groups, becomes increasingly blurred by the wording of article 510 and the additional institutional clarifications surrounding it. This also brings up questions around the gaps between the written law and its interpretations and implementation. As I have exposed above, there is very little jurisprudence to lead the way in interpreting the legal framework in this regard.

There are also diverging opinions on whether this article should or even could be applied in cases of gender-based online hate speech. In this regard, scholars such as Víctor Gómez (2016) have argued that article 510 could not even in theory be applied on sexist discourse, arguing that "it seems reasonable to demand, nonetheless, at least, that it is a minimally adequate behaviour to alter in some way the collective security or the honour of women as a collective. This, as is easy to see, will not happen easily"⁵⁵ (Gómez, 2016: 25). That is, Gómez argues that as the security or dignity of *all* women is not likely to be threatened, this type of discourse should not be prosecuted as hate speech. Some of the interviewed experts argue along the same lines:

Normally the classical or dominant interpretation of hate speech is when hate speech is apologetic of the holocaust, anti-Semitic, racist speech, but it would be very difficult for them to come to accept that anti-gender hate speech is really hate speech (...) If we try to apply hate speech against women or against certain women, it would be impossible to prove, because you cannot prove that this hate speech is inciting violence against all women in Spain and that it implies an imminent threat of persecution

⁵⁵ Own translation from Spanish.

and annihilation of women, right? [EX-15]

Nevertheless, other experts interviewed argue that article 510 can be applied to gender-based hate speech: "Can you apply it? It can be applied. It's been contemplated like that. Another thing is that statistically I don't know of any case in which it has been applied" [EX-06].

The scarce jurisprudence that exists and that I have previously exposed, shows that article 510 both can be and is being applied to hatred on the grounds of gender. A remark by Noelia Igareda (2022) is relevant in this regard, as she makes an important distinction regarding the application of hate speech laws in the case of gender-based hate speech. From this perspective, hate speech laws would apply to anti-gender hate speech that targets women in general, groups of women in particular, or groups of people because of their gender, sexual orientation, or gender identity. But in the case of anti-gender hate speech directed at a specific woman, who due to her public projection is especially exposed on the internet and social media and the attack is perpetrated by an unknown number of people, it would be difficult to apply these legal frameworks and jurisprudential interpretations (Igareda 2022: 116). As I have previously concluded, according to this interpretation, then, article 510 would apply only in cases of hatred directed against women as a group. As we shall see in the next chapter, following interviews with police officers, this also seems to be the route followed at least by some police forces in Spain.

Reflecting further on the interpretations of article 510 by scholars and other experts, in addition to the converging opinions of whether hate speech laws should at all be applied to hate speech against women, the scarce use of article 510 can also be understood as linked to the difficulty in drawing the line between generic speech against the collective of women that can be perceived as sexist, but within the legal limits, and that which amounts to a crime. This is particularly the case for discussions or disputes in the political sphere, where the topics of gender roles and norms are to a certain extent treated as part of a democratic discussion among political parties. As a legal scholar points out:

It's practically impossible that speech that is clearly sexist or denying of

gender violence as we have in Spain with Vox, could be considered hate speech. Because you could never say that Vox is instigating people to murder women, starting tomorrow, nor could you really ever prove that their speech poses a threat to 51% of the Spanish population that are women. That's why it would never be considered hate speech no matter how much the content disgusts us. That is, being sexist is not a crime. This is also sometimes interesting, to differentiate between hate speech and discriminatory or sexist or homophobic content, or xenophobic. Being xenophobic or being sexist is not a crime in itself and that sometimes is difficult to understand [EX-15].

Here, again, it is important to consider that hate speech does not need to amount to a criminal limit to fulfil its function of dehumanising and ridiculing, and creating harm; however, to apply criminal law measures, it needs to tick off certain parameters.⁵⁶

The lack of application of art 510 in cases of gender-based online hate speech is also perceived as related to underreporting. Ample evidence shows that underreporting continues to be a major problem in relation to both hate crimes and hate speech (e.g., European Union Agency for Fundamental Rights, 2021; Ministerio del Interior, 2021), as well as gender-based violence (e.g., Government Delegation against Gender Violence, 2019). Few victims report gendered online violence and hate speech to the police, and if reported, it tends not to be categorised under art 510, nor with the aggravating factor of gender.

Another limitation to the application of article 510 for gender-based hate speech is, as argued by a lawyer, that "although it could be applied to legal persons, it is focusing on individual offenders when in the Spanish state those who exercise great misogynistic hatred are religious confessions, right-wing parties and various groups. That should also be taken into account" [EX-06]. The difficulty of applying criminal law on collective crimes (i.e. with several perpetrators) online has also been recognised by Igareda et al (2019) in their study on misogynist cyberviolence, further highlighting how the specific affordances of digital violence puts into question the current legal instruments.

⁵⁶ In Chapter 6, I shall look further into the difficulty of establishing these limits even for the police and representatives of civil society organisations that work with hate crimes and hate speech on a day-to-day basis.

Further, the scarce application in cases where the bias is gender or sex is also perceived as linked to a lack of effort or decision in the feminist movement in including sexist or gender-based discourses in the broader category of hate speech:

So there is a debate about whether it is more strategic for women to embrace this broader, more diffuse category. For example, when the Code of Conduct of the large internet intermediary platforms was approved in 2016, gender discourse has remained there half volatilized. Or, the other option to claim it as political violence that affects half the planet, and turn to tools for women's rights and the fight against sexist violence [EX-06].

That is, there has not been a strong civil society movement aiming towards a higher level of inclusion of gender specifically in instruments against online hate speech. Instead, as we have seen in the previously mentioned regional laws from the autonomous communities of Galicia and Catalonia, efforts have been made to include digital violence as an aspect in legislation surrounding gender-based violence in general. Hate speech laws, thus, seem to still be considered to cover mainly racism and LGBT-phobia, rather than gender-based speech. Part of the explanation of this lack of application of hate speech laws in cases of gender-based hate speech can be linked to the view of violence against women as belonging exclusively to the private sphere, predominantly of (heteronormative) romantic relationships:

The system has a tendency for everything that is violence against women to pass through a kind of private filter or conflict, let's say personal. Therefore, it is always easier for a certain discourse against a feminist journalist to be read as an attack on honour instead of political violence. So between the under-reporting of women through 510, the internal dynamics of the judicial system and feminisms that have not made it a battle to include misogynistic hate speech in this category, there has been a reality of general underuse of this article [EX-06].

That is, the structural components of gender-based violence in contexts outside of intimate partner violence are not visibilised, and this violence is not read as based on gender, or as political violence. Over the last few years, a gender perspective on online attacks against certain women, such as feminists or women journalists might have become more common in media and social debates;

however, in terms of legal protection, the gendered aspects are not recognised.

Finally, in terms of including an intersectional perspective in prosecution through article 510, there is agreement that judicial operators are still far from recognising intersectional grounds. However, theoretically, it would be possible from a legal point of view:

The system in general is very reluctant to acknowledging the political foundation of any violence. Asking the system to determine this interaction of axes of discrimination, that is, that it recognises that speech is at the same time racist and misogynistic, is quite complicated. Legally it is possible, but it is quite complicated because I insist, the courts have much difficulty in capturing the motivations [EX-06].

The reluctance to recognising intersecting oppressions and applying the concept of intersectionality in practice is interesting, given the concept's very origin in discrimination law. The original contribution made by legal scholar Kimberlé Crenshaw responded directly to the issue of justice with respect to antidiscrimination legislation in the US for black women, arguing that "black women were located at the intersection of racism and sexism. Their experiences were thus the product of both and equivalence of neither" (Crenshaw, 1989: 22). Despite the decades that have passed since the birth of this concept, lawmakers have been slow in incorporating a perspective where more than one axis of oppression is considered at the same time. This in turn invisibilises the specific vulnerabilities of racialised and LGBT+ women, who are targeted by online attacks to a higher extent than white, heterosexual women.

5.7.2. Applying the framework of gender violence on online gender-based hate speech

Returning to the Spanish legal framework on gender-based violence, as I have previously exposed, although the Istanbul Convention has been signed and ratified by the Spanish state, the concept of gender violence has been erroneously incorporated in the Spanish legal framework. That is, this incompleteness of the Spanish definition of gender violence highly restricts the acknowledgement of a series of crimes as gender-based, amongst others, online hate speech, in cases where the offender is not the woman's partner or ex-partner. As a legal scholar highlights:

The instruments of the EU and the Council of Europe speak of 'gender-based violence' but Spain has translated this into 'gender violence', which is not the same as 'gender-based violence', so it seems like a translation trick so to speak, but of course it has very important legal consequences. So we find laws that don't really talk about gender-based violence, because they only talk about violence within the couple [EX-6].

Consequently, this imperfection of the Spanish legal definition of gender violence restricts the recognition of a series of crimes as “gender-based”. Although it must be recognised that violence carried out within heteronormative couples is a serious problem in Spain, this should not be to the detriment of other types of gender-motivated crimes: "I think it leaves out many types of violence, it ends up being explosive, it ends up reinforcing binary divisions, heteronormativity, some of the things that seem very interesting to me to be questioned but not to invalidate everything else” (EX-17).

Some of the new penal figures that have been introduced into the Spanish penal code as a result of the impulse of the Istanbul Convention, and which were originally aimed at protecting women from sexist violence, have in the Spanish context been drafted in a gender-neutral way (Serra, 2018a). An example of this is the case of stalking. Due to the limitation of the state legislation on gender violence to the framework of the (ex-) partner, digital violence against unknown women is not conceived as sexist or gender violence. Thus, in cases of stalking, the severity of the sentence is heightened only when it occurs within the context of a current or former intimate relationship. Otherwise, regardless of whether the perpetrator or victim is male or female, the punishment remains consistent. Consequently, within the Spanish criminal code, the gender aspect becomes invisible in instances of digital violence occurring outside the realm of intimate partnerships (Ibid.).

Finally, considering the recent focus by the OSCE-ODIHR on gender-based crimes as hate crimes, it also becomes interesting to analyse briefly the application of the aggravating factor of gender of 22.4 in the Spanish criminal code. Also this article is subject to widely different interpretations of its scope, much related to the legal definitions of gender violence within the Spanish criminal framework.

Scholars such as Marin de Espinosa (2018) have analysed this aggravating circumstance and argue that, following the erroneous concept of gender violence in the *Organic Law 1/2004, of December 28, of Comprehensive Protection Measures against Gender Violence*, the aggravating circumstance of gender is restricted to cases where the victim maintains or has maintained a sentimental relationship with the aggressor. That is, according to this view, an online threat or online harassment could only be prosecuted with this aggravating factor if the threat comes from the victim's partner or ex-partner, placing this type of harassment within a heteronormative framework.

Nevertheless, other experts argue that it is applicable and that it is in fact being applied also outside of cases of intimate partner violence, although still very scarcely:

The aggravating factor of gender of 22.4 of the penal code is hardly being applied, but since 2017-2018 the supreme court has begun to make sentences that are quite good, incorporating and validating the gender perspective. There are a few sentences, like one that I got last week, that are beginning to apply the aggravating factor of gender in sexual violence, outside the framework of the couple [EX-06].

Even if this article is not directly applicable to hate speech, the onset of its use to signal that a crime has been aggravated by gender may speak of the start of a slow change in the justice system, towards a better detection of gendered structures behind crimes beyond those perpetrated by a partner or ex-partner of the victim.

The lack of recognition of the gender-based motivation for these online crimes means that there is a lack of much needed specialisation, both for addressing victims' needs and for more technical aspects, such as skills related to the collection of evidence. This lack of specialisation and awareness on the specificity of these crimes has consequences for the whole criminal justice process, including the reporting, prosecution, court procedures and outcomes, and particularly for the experience of the victim.

5.7.3. Perspectives on regional (civil or administrative) laws

As previously mentioned, the Catalan regional law against gender violence was modified in *Law 17/2020, of December 22, modifying Law 5/2008, of the right to eradicate male violence*. As this change took place during my fieldwork with professionals, this law was brought up for discussion on numerous occasions during interviews in the Catalan territory. Whilst this law does not apply to criminal behaviours, the amended law could –at least theoretically– open the way for administrative penalties for digital violence or online gender-based hate speech. Whilst some of the experts interviewed perceive this law as a great impulse to recognising digital violence against women, others –mainly those whose work is limited to addressing behaviours within the criminal code– perceive it as a merely symbolic change as

the competences in criminal matters only belong to the state. Then there are autonomic laws like the Catalan that go beyond the definitions of the violence stipulated by the state law. That's very well from a political point of view, in a symbolic sense, but in legal effects, it has no repercussions because it does not modify the penal code or modify the main laws [EX-15].

Nevertheless, looking beyond merely criminal matters, this and other similar laws could provide victims of gendered online violence with access to support mechanisms for victims of gender-based violence. Unfortunately, these support mechanisms are already much overwhelmed by the number of victims of intimate partner violence and would need more resources to be able to cope with victims of other types of violence.

Experts also point to how this law could contribute to compiling statistics:

Perhaps this would allow for all these cases of digital violence that reach the Mossos d'Esquadra⁵⁷ to be identified as gender-based violence and registered in a specific way, because if not, now when you ask what's the prevalence of digital violence in Spain? We have no idea (...) And what is not quantified doesn't exist and you can never justify it as an issue of attention from a legal or political perspective [EX-15].

⁵⁷ The Mossos d'Esquadra is the Catalan police force.

Returning to Serra's (2018b) categorisation of hate speech, where hate speech of “medium intensity” would be subject to civil sanctions, the reformed law could in theory open up for a path for civil sanctions for gender-based online hate speech. However, some experts interviewed view this path as unlikely to have any real consequences, as the institution responsible for upholding this administrative law has no real powers in investigating on the internet and on social media.

Drawing comparisons with the Catalan regional LGBT law, which was established in 2014, LGBT organizations interviewed express that very few infractions that are reported administratively actually lead to sanctions. According to a prosecutor interviewed, this is due to a blurring between administrative infractions and criminal offenses in the Catalan LGBT law: "It's an ineffective law, because all the behaviours it describes are criminal, they fall within the criminal jurisdiction. Of course, they complain that the law is not enforced, well they don't have administrative infractions to apply, because everything described in it is actually a crime" [EX-09]. However, the same prosecutor perceives that a well-defined civil law could facilitate addressing some infractions that take place online:

On the internet there are behaviours or there is dissemination of expressions that are not criminal and that don't have to be criminalised because not everything is criminal law, but it would be good if that type of insults were classified or this type of comments that are sometimes done when a piece of news is published in a digital newspaper or on social media, that this could be pursued at an administrative level [EX-09].

That is, the incorporation of digital violence into the reformed Catalan law on gender violence could serve as a means to gather statistics on instances of digital violence, even those not constituting criminal offenses. Additionally, it might pave the way for victims to access support services and resources aligned with this law.

5.8. Conclusion: A fragmented legal framework

As I have argued in this chapter, the legal framework applicable to gender-based online hate speech, or in a broader sense, to online gender-based violence, can be defined as fragmented. The obstacle to recognising gender-based motivation

in online violence is twofold. On the one hand, the legal framework on gender violence in Spain is limited to intimate partner violence. On the other hand, gender-based hate crimes tend not to be recognised as such, although, as we have seen, gender has been included as a protected characteristic or aggravating circumstance in Spain's Penal Code since 2015.

This means that there is no straightforward or coherent legal framework contemplated for application on hate speech or violence targeting women online. If digital violence is exercised within the couple or ex-couple, the legal path is clearer, as the legal framework of gender violence is applicable. However, in cases where the violence takes place outside of the couple or ex-couple, for example by unknown perpetrators against feminists and other activists, as in the scope of this research, bits and pieces of law can be adapted to the various types of online violence that women suffer. Nevertheless, the most common case is that this violence, as we shall see in Chapter 9, is neither reported nor prosecuted.

When examining instances of verbal violence manifested online and the potential application of existing hate speech laws, diverse interpretations emerge regarding the motivations encompassed within the scope of hate crimes. Additionally, there is a noticeable underutilisation of these laws in instances of gender-based online hate speech. Based on the limited available jurisprudence and scholarly analyses, the current application of hate speech laws in cases of gender seems to be restricted to instances where the hate speech is directed towards women collectively, rather than towards an individual woman based on her gender. Even when the speech against an individual aims to denigrate women or feminists as a collective, these laws tend not to be invoked for individual cases, necessitating the application of alternative legal instruments that may not adequately address the structural – gendered – underpinnings of such speech acts. Moreover, hate speech often involves intersecting grounds, such as sexual orientation and gender, or racialisation and gender, which remain inadequately addressed by existing legal frameworks.

In addition to this, there is a clash between different definitions of gender-based violence from a legal perspective. Following the Spanish penal code, and in

contrast with the Istanbul Convention, incidents and crimes tend to be labelled as “gender violence” only when they take place within a heteronormative couple or ex-couple. As we have seen, some scholars argue that this is currently changing and that crimes outside of the couple or ex-couple are now to a higher extent labelled with the aggravating circumstance of gender of Article 22.4. However, deeper research into this jurisprudence is beyond the scope of this thesis.

Violence against women is still to a high extent interpreted as belonging to the private sphere, and there is a lack of articulation of this violence as gendered in other spheres, and outside of the heteronormative couple or ex-couple. Although the reform of the Spanish criminal code in 2015 included the aggravating circumstances of a crime the grounds of sex and gender, this has been very scarcely applied, particularly in relation to hate speech. Further, some of the new penal figures that have been introduced into the Spanish penal code due to the Istanbul Convention, and which were originally aimed at protecting women from sexist violence, have in the Spanish context been drafted in a gender-neutral way. As I shall expose further in Chapter 6, women who are victims of other types of gender-based violence, such as online hate speech or digital violence at large, fall into a protection gap. These crimes or incidents are generally neither labelled as “gender-based violence” nor as “hate crimes”, meaning that a structural analysis of and approach to these crimes is lacking, that statistics are not collected, and that specialised support services are not offered. As we shall see in later chapters, this legal gap also has a direct impact on the institutional services available to the targets of this type of hate speech, and to the experiences of the victims when reporting.

Nevertheless, more than new laws or a tightening of the criminal code, what might be needed to close some legal protection gaps is a reformulation of what already exists in the criminal code, linked to an updated definition of gender-based violence in lines with the Istanbul Convention. It is also crucial to keep in mind that not all discriminatory behaviours are necessarily criminal behaviours, and that this is a complex interpretation, which is perceived as even more difficult in online spaces. Examining the functions and effects of online gender-

based hate speech, which I shall discuss in Chapter 8, these are noticeable also at a collective level, with women being scared away from participating in discussions in public spaces. Consequently, it becomes evident that additional protective measures are essential to safeguard women who express their views in public discourse from online violence. These protective mechanisms should encompass a combination of criminal law, civil law, and an elevated level of accountability assumed by social media platforms to ensure the protection of women and other groups especially susceptible to online violence.

Finally, it is important to signal the importance of laws in making clear the lines between socially and culturally acceptable and unacceptable behaviours (e.g., Dixon and Gadd, 2006; Mason-Bish, 2011; Smith, 2019). That is, whilst improvements to laws are far from the only way of improving societal responses, the symbolic value of legislating against gendered online violence is crucial.

6. Professionals' perspectives on online gender-based violence

We have that idea that physical violence is the one that generates the most impact, and obviously it does, but there are other, slightly more invisible forms of violence that can play a very powerful role (EX-16)

6.1. Introduction

It is essential for institutions and organizations committed to combating gender-based violence online to grasp both the dynamics and consequences of this violence. Additionally, recognising its structural roots is equally important. Nevertheless, there is a tendency among entities that women turn to for assistance, such as platforms, tech companies, and justice agencies, to perceive online violence as less significant or separate from offline violence (e.g., Harris and Vitis, 2020; Harris, 2020; Citron, 2009). Despite this, there is a lack of research on the perspectives of practitioners and professionals who engage with victims of gender-based online violence or hate speech. One notable exception is the study conducted by Anastasia Powell and Nicola Henry (2016), which delves into the viewpoints of police and service sector professionals regarding the

policing of technology-facilitated sexual violence. In the Spanish context, there is a dearth of research examining practitioners' attitudes to gendered online violence. This chapter aims to fill that gap by analysing the perceptions and categorisations of verbal gender-based violence in digital spaces in contemporary Spain.

As exposed in Chapter 4, the analysis is based on ethnographic interviews with professionals who provide support to victims of gender-based violence and anti-LGBT violence from different perspectives, as well as practitioners from the justice system and law enforcement agencies. Against the backdrop of the legal debates introduced in Chapter 5, I will concentrate on these professionals' perceptions of online violence and their readiness to address it within various institutions and organizations in Spain.

The chapter thus centres on the perspectives of these professionals regarding the online context, their classifications of gendered online violence, and their preparedness to address such instances. It presents its findings organized into four distinct subtopics: understanding how institutions and organizations perceive and categorise gendered online violence, assessing professionals' readiness to respond to online violence, examining perceptions of responses from law enforcement and the justice system, and exploring training requirements and potential strategies for progress.

6.2. Perceptions and categorisations of gendered online violence

In tackling matters related to gender, institutions and organizations typically encompass legal and social support within the broader framework of "gender violence" or "gender-based violence", as opposed to a hate crime framework, as explored in Chapter 5. This holds true, as well, when addressing gendered violence in online environments. In the context of Spain, this can be observed, for example, in the work of the organization Calala, which, in its report "Diagnosis on gender violence against feminist activists in the digital sphere", analyses this violence within the framework of gender violence: "We start from a broad conception of gender violence, which includes the set of violence that derives

from the cisheteropatriarchal system and that is directed at people because of their gender position, their gender identity, their gender expression or their sexual orientation” (Morena-Balaguér et al, 2021). Also Noelia Igareda et al. (2019), in their examination of "sexist cyberviolence," adopt the framework of gender violence. However, as discussed in the preceding chapter, the categorisation becomes more intricate when considering the perspective of available legal tools, which often underpin protection mechanisms. It is apparent that legal definitions for many aspects encompassed in online gender-based violence or hate speech are lacking.

During the fieldwork that underlies this chapter, when making contacts with different public institutions across Spain, these generally redirected me to services catering to victims of gender violence, because those are the ones that address, at least to some extent, digital violence. Although many of these increasingly take into consideration the aspect of technology-facilitated violence; it should be noted that whilst they are aware of this problematic, they cannot place much focus on this type of violence due to a scarcity of resources. Further, these services tend to predominantly address intimate partner (or ex-partner) violence, following the definition of the Spanish legal framework. Even so, the interviews with these services provided valuable information on perceptions by professionals about the different types of online violence and their impacts, as well as the availability of support.

The above indicates that, initially, when addressing the issue of online sexist violence, a considerable number of the interviewed professionals confine this type of violence solely to the realm of heteronormative relationships. In other words, they perceive it as a form of gender violence perpetrated by a man against a woman, who is either his current or former partner. Within this framework, numerous services are available to assist victims of gender violence, as defined by the established laws, and these services are offered by both public institutions and non-governmental organizations (NGOs).

Broadly speaking, the public services catering to victims of gender violence acknowledge its prevalence in digital realms. Both governmental bodies and non-

governmental organizations demonstrate a certain level of comprehension regarding the significance and ubiquity of these online spaces, especially in the lives of younger individuals.: “This is a kind of thing where you are constantly connected to your mobile (...) It’s a violence that doesn’t stop, 24 hours a day, every day of the year” (EX-14). Even though most services primarily concentrate on intimate partner violence, several professionals also acknowledge that digital violence is not confined solely to within couples.:

Here we also differentiate between online violence received from someone known, and violence received by someone unknown or anonymous. We find the most cases of someone known, but we know that there is violence through persecution on social media, through platforms such as Wallapop⁵⁸, Instagram, Facebook, that someone starts to follow you and coerce you. We don’t have it quantified, unless a woman comes and identifies very clearly that she is suffering from online violence. But I think that the aspect that we could find the most is the digital element as another instrument used by a known aggressor or the abuser (EX-14).

The potential impact of online violence on the individual victim is also acknowledged by several professionals:

The level of violence is very extreme, it’s very high and it has a big impact. We have that idea that physical violence is the one that generates the most impact, and obviously it does, but there are other, slightly more invisible forms of violence that can play a very powerful role (EX-16).

Nevertheless, as is reflected in the quote above, even when recognising the possible impact of online violence, professionals tend to implicitly refer to a hierarchy of violence. Within this hierarchy, physical violence is categorised as “real violence” and perceived as the type of violence with the highest impact.

In the Spanish framework, services are equipped to address gender violence, encompassing both physical and psychological forms, using their many times limited resources. However, these services generally lack specific training or formal protocols to handle online violence. In fact, certain services do not even have a designated category in their internal computer systems for incidents occurring in digital spaces:

⁵⁸ Wallapop is a mobile application for selling second-hand items.

We would mark the item as sexist violence and we would look for psychological, physical, and economic support, but we don't have an item where we know the quantification of violence that occurs through mobile phones or online violence. We know that it exists, and we know that in cases where psychologists or social workers are working with the women, WhatsApp messages are mentioned in the interviews, the pressure through 'where are you?' (...). And of course it exists, but we haven't quantified it. We don't know how many cases we address each year (EX-14).

The absence of designated categories not only results in a shortage of statistics but also leads to a deficiency in the preparedness of professionals to handle instances of online violence. When acts of violence in digital spaces are not categorised as such within public services, they remain unseen. Consequently, tools, protocols, and specialised training are not developed, leaving practitioners without the specific knowledge and skills required to address the needs of victimised women in these situations.

There are further obstacles to providing support to victims of online gender-based violence. First, the vast majority of services are aimed at victims of gender violence as defined by the categorisations of the Spanish legal framework:

I started working in the department 3 years ago and what I found was an obviously very heteronormative approach to violence services, because it has a historical logic and a legal logic (...) I remember strong discussions with teammates, and saying 'but let's see, not all women have the same sexual orientation', and that has forced us to make an analysis of what gender violence we are leaving out or not attending to (EX-12).

There are examples of local public services, for example in Catalonia, that have begun to extend their support services beyond the strictly heteronormative definitions of gender violence. They have expanded their categorisation of this violence, incorporating a somewhat intersectional approach, particularly in terms of the intersection between gender and sexual orientation. The more expansive approach aligns with the broader definition of gender violence outlined in Catalonia's autonomous law on the subject. This legal framework is more closely aligned with the Istanbul Convention compared to the Spanish framework:

So we already have that X-ray identified, we have already done training with all the teams. What's more, this caused, for example, that [a service

for victims of gender violence] is already providing services not only to heterosexual women, but it also caters to other situations of gender-based violence (...) We know that we have to generate care for different cases of violence (EX-12).

Secondly, the online dimension itself is perceived as a major obstacle to providing support to victims. On the one hand, online incidents are perceived as aggravated by the Covid-19 pandemic, as much social and work life -and thus also hate based violence- have been transferred online: "So it's true that in 2020, with the lockdown, obviously the Internet and social media is an item that skyrocketed. It's obvious" [EX-05]. On the other hand, analysing the interviews with professionals, it is evident that violence expressed in digital spaces is still generally perceived as a great challenge. Most of the professionals interviewed express a series of doubts related to victimisation in digital spaces. Part of this perception among professionals is linked to the perceived immensity of the internet itself. Especially among the representatives of NGOs interviewed, who typically take a more proactive stance in addressing or monitoring various forms of violence, the internet is perceived as too expansive a space to effectively monitor or tackle violent expressions. Organizations tend to lack the resources to go beyond addressing incidents that are reported directly to them: "What happens with the Internet? It is true that we have a more, so to speak, passive attitude. That is, it has to come to us, we don't have the resources to go looking for incidents online" (EX-04). This perspective might be partially associated with the perceived hierarchy of violence, wherein physically expressed violence is labelled as "real" violence. This means that efforts by institutions and NGOs focus on protection from physical violence.

However, professionals and representatives from civil society also voice scepticism regarding the boundaries of freedom of speech in digital spaces. The threshold for online expression not being classified as hate speech is much higher compared to in-person incidents, as face-to-face hate speech is perceived to entail a much higher immediate risk of physical violence. This is understandably so; however, recent years have also provided examples of acts of physical violence that have been linked to misogynist discourses in online forums

connected to the manosphere and particularly to incel forums (e.g., Barcellona, 2022). That is, to comprehensively understand the consequences of anti-feminism and anti-LGBT movements, and the interplay between the online and offline realms in this context, one must consider not only physical violence but also the impact of violent expressions online.

The professionals' perceptions of the online context as less significant and too intricate to manage or address may not only restrict their understanding of what constitutes violence but may also create challenges for victims of gendered online violence in coming forward to report. This perspective may also result in impractical or unrealistic advice from these professionals, who are intended to offer support, as I will elaborate on in the following section.

6.3. Preparedness to respond to online violence

Among the interviewed professionals and practitioners, there persists a perception of the online and offline as distinct, fixed binary categories, rather than as a continuum or highly interconnected spaces. As highlighted by Mark Deuze et al. (2012), contemporary life is increasingly lived in media rather than merely with media, and a more profound understanding of the implications of this phenomenon is lacking not only among professionals responding to violence incidents but also within society as a whole. This contributes to a view of the online as a confined space that those experiencing online violence can simply exit. Hence, it is imperative to give more consideration to acknowledging the inherent and ongoing connections between interactions and communication across both online and offline domains. This involves recognising that an individual's social environment is substantially experienced in both realms (Powell and Henry, 2016). When the digital and physical spheres are perceived as separate and importance is being given mainly to what happens in the physical context, this not only downplays what happens online, but it also leads to often unrealistic advice from professionals on how to avoid attacks in digital spaces. As explained by a representative of a public service for victims of gender violence, who details the advice they tend to give to women regarding digital spaces:

So the guidelines that we give them is to cut with all the digital footprint they may have, neither Instagram nor Facebook, we change the phone card, we remove the location of the phone, that is, we try to suppress absolutely any possibility to locate them through any device (EX-16).

In instances of digital gender-based violence, professionals often find themselves offering the advice to disconnect from online platforms, a challenging task for young individuals whose lives are intertwined with social media. Economic factors may also play a role, especially for women reliant on social media or online spaces for income generation. The interviewed professionals do acknowledge, to some extent, the challenges of disconnecting. However, due to a lack of in-depth knowledge and specific training on potential online protection mechanisms, they frequently feel limited in offering alternative advice. Their primary objective remains the protection of victims, primarily from the risk of being located by the perpetrator. Here too, then, we can see how concern is formulated around possible physical threats, with physical violence defined as “real” violence, while the online violence is not perceived with as much concern. Therefore, regardless of who is exposed to online violence, the onus to protect themselves tends to fall on the victims, preferably through total disengagement from online spaces:

We tell them that if they open another profile on social media again, because many are young women, who use Instagram or TikTok and things like that and now it is very difficult not to be present on social media... The guidelines are always not to take any photograph with elements that can locate them, that can identify them, that they need to be very prudent in this sense to avoid new re-victimizations in the future (EX-16).

As Powell and Henry (2016) have noted, there may be adverse impacts associated with advice that victims should disconnect. Such advice may result in extending the social abuse and control of the perpetrator over the woman’s life, as well as the denial of digital citizenship (see also Henry and Powell, 2014). In essence, the concept of self-protection (including setting social media profiles to private or exiting social media entirely) as the only available protection strategy, can result in the silencing of certain voices. This is especially relevant in the context of activists engaging in discussions within digital spaces. The individuals

most susceptible to being silenced are those facing intersecting forms of oppression, such as racialised or LGBT+ women. Another issue that specifically LGBT organizations struggle with is that a high percentage of reports are presented to them by white gay men, whilst other groups within the LGBT+ collective, such as lesbian and bisexual women, as well as trans persons, report incidents to a much lesser extent:

I think it has to do with the logic, it has to do with the logic that the white gay man can report. And in fact he reports, or has the chance to report it. Maybe women, and I'm talking hypothetically because I haven't come across so many cases, maybe lesbian women have other itineraries that don't go through reporting many times when it comes to repairing their dignity (EX-05).

Finally, related to reporting, the balance is often complex - and even more so online - between behaviour that could constitute a criminal offense and incidents that should be treated as administrative infractions. Representatives of feminist and LGBT organizations interviewed admit the difficulty in categorising online incidents and they recognise that "at the level of legal regulation it's much more complicated to manage this, when things happen on social media, because there is so much, so much. And we set the criteria: if you report to us, we record it" (EX-05). Linked to this, in the next section, I shall discuss the perceptions and responses to online violence of the police and other agents within the justice system.

6.4. Perceptions on reporting and responses by the police

As mentioned earlier, civil society organizations that address feminism and LGBT+ rights harbour reservations concerning the online dimension, particularly regarding the boundaries of freedom of expression in digital spaces. The balance is frequently intricate, especially online, when navigating between behaviours that might amount to criminal offenses and incidents that could be considered administrative infractions: "We think we have to polish and see what we are talking about when we talk about online hate incidents, because this also leads

to a much deeper debate that has to do with freedom of expression..." (EX-04). Coupled with the perceived difficulty in monitoring digital spaces, this contributes to a reticence in reporting incidents to the police. Additionally, there exists a widespread belief among organizations and institutions that prosecuting cases of digital violence, especially online hate speech, is intricate, primarily due to territorial aspects that may impede actions in online cases:

There is a clearly territorial theme (...) Even if they live in [a Spanish city], the police may tell us that he is a man from that city and perhaps they can even prove it to you, but then the account is I don't know where. And so, this is very complicated, if not almost impossible to address (EX-07).

For some professionals, the uncertainty that surrounds the online context has even been reinforced through consultations with the police, as pointed out by a representative of a public anti-discrimination service, who states that "the police themselves have recognised many times that tracking and being able to reach the origin of many of these discourses or many of these situations of virtual discrimination is difficult, it's not easy" (EX-13). The classification of online crime as inherently challenging to prosecute could, in turn, create hesitancy among professionals to encourage victims to report, as they may view the process as futile. Nevertheless, some of the civil society associations interviewed argue that they encourage those that seek their advice to report to the police or other institutions:

And we always say, go and report it, for legal reasons, if you don't report, at the legal level [the crimes are not recorded]. In fact, the vast majority of incidents don't have a legal itinerary. They don't go through a legal or criminal route, although it's the most common of our cases, this year I think we have received 30 criminal cases (EX-05).

Another aspect of the online dimension that creates doubts and insecurities among professionals is the collection of evidence in case the victim wants to report digital violence to the police:

If you are going to report, from what perspective do the police advise on what to do? Because, of course, it's a crime, but between having evidence, that your profile on social media is being questioned, if you have published a photo like that and so on (...) Then we don't know very well how

someone should have the mobile phone programmed so that it really becomes forensic evidence, how is that done, do women know how to export those conversations, are screenshots useful or not? (EX-14).

As this quote highlights, in addition to lacking knowledge about the collection of evidence in digital spaces, professionals from different services also question the treatment that victims may receive by the police. This further raises the question of prevailing biases of sexism, racism and LGBT-phobia within law enforcement agencies themselves. In addition to fears of being questioned, the digital context also entails an additional layer of uncertainty about the limits of what can and should be reported to the police, and if the police file a report, what information counts as evidence in these cases.

Based on interviews with legal experts, we can also begin to discern the difficulty of fitting hate speech or gender-based digital violence into police work:

Either you have police forces that are specialised on gender violence, so they are always thinking about the sphere of violence or abuse by your partner or ex-partner, a purely physical issue, or you have a team specialised on hate speech or cybercrime, so they are thinking about cybercrimes such as big tech scams. They're not thinking about the attack on a councillor or a gamer through social media, those are breadcrumbs. I mean, it doesn't make any sense to them... (EX-15).

As Powell and Henry have highlighted, law enforcement responses tend to minimise the harms of gendered violence when it takes place online or via e-mail and text-based communications:

the minimisation of digital forms of abuse, violence and harassment perhaps stems from a fixation on corporeal or bodily forms of harm, as opposed to harms that are structural, social, emotional or psychological. Conventional conceptualisations of criminality require an overhaul in both theory and practice in order to address relations of power and exploitation that are reproduced in digital contexts, and which can potentially serve to destroy another person's life (2016: 11).

Here, then, two issues come into play that create difficulties in categorising online violence or hate crimes within police work. On the one hand, the online dimension, and, on the other hand, the lack of recognition of gender-based crimes as hate crimes (beyond anti-LGBT hate crimes) or as gender-based violence (beyond

intimate partner violence). This is also made evident by talking to a police unit that specialises on hate crimes. According to representatives of this unit, they can only recall one case of online hate speech against women, which was a hate website against women, reported to the police by a women's rights institution:

An online event and that it is classified as a crime, and that the Cyber Crime unit lends a hand, we have had cases, especially regarding gender, there were websites, people who had created a website with the aim of denigrating to women or promote the supremacy of men in relation to women, and then there was an obvious case that it was criminal (EX-10).

In other words, and in lines with the theoretical arguments by scholars such as Igareda (2022), exposed in Chapter 5, the practical application of the framework of hate speech based on gender is expressed here only in cases of hatred against the entire group of women, and not on attacks against individual women even if the attacks are grounded on gender-based hate; that is, in cases when the attack against an individual integrates hate speech towards women as a collective. On the other hand, following the Spanish criminal code, the police categorise as "gender-based violence" only those events that have taken place within a heteronormative relationship.

In the case referenced above, the interviewed police officer is very certain that this web could be classified as a gender-based hate crime: "Yes, that would be a hate crime, it would be sexism, it would be sexism for promoting this hatred of women, this hatred of women and this feeling that some are above others and this promotion of intolerance" (EX-10). However, it is often difficult even for the police to draw the line of freedom of expression, or to know whether an online incident amounts to a crime. For example, talking about online incidents related to gender-based or anti-LGBT violence, a victim support officer from the police reflects that:

Sometimes there is a doubt, I don't know whether to approach this as administrative or if it's almost there, is this already criminal? Well then, it depends on what it is, yes, tell them to come, we'll report it (...). We also have cases that are sometimes targeted, especially with people who are migrants, where the police in their country have a different temperament, and the issue of a uniform, or someone who is in an irregular situation and

says, look, 'I don't know whether to report or not report'. Then just as often we do an escort and the complaint, instead of reporting it at the police station, we accompany them to the support service of reference and record the report there. It's all a bit, I mean, the procedure has its own standards, but it's flexible and then we adapt a little (EX-11).

Nevertheless, the general perception on where to draw the line in terms of criminal behaviour is that the facts must be rather serious to be able to open an investigation:

Now I make it up, but some insults to a person who is either from the LGBTI sphere or a woman of recognised prestige, some insults for being a woman and so on, that would still be a more administrative matter (...). For it to become penal and counted as penal it has to be something a little beyond some insults or so. In the case of LGBTI people, some insults would be administratively sanctioned based on civil laws, but in the case of insults to a woman for being a woman, I don't know... (EX-10).

As expressed in the quote, attacks against LGTB+ people can be administratively sanctioned in some autonomous communities in Spain, such as Catalonia. In other words, it depends on the intersections of different factors as to whether a victim can find ways beyond the criminal justice system to report an online attack that they perceive as identity-based. However, as I have previously alluded to, the online context presents additional obstacles for administrative processes, related to territoriality and the powers of administrations, for example, in terms of being able to identify the offender in cases that involve anonymous accounts.

Nevertheless, also professionals who are active within the justice system perceive online cases as highly challenging. For example, a prosecutor describes the difficulties in finding out the identity behind an anonymous profile:

In principle, the difficulty we have with false or anonymous profiles is knowing who is behind that profile. Most of the social media platforms are legally located in the United States (...) So, if you want to know who the owner is, you have to do it with a rogatory letter through the United States. You already know that the US judicial authorities are going to tell us that this falls under the first amendment and therefore they cannot provide the information (EX-09).

This has been previously noted by Powell and Henry (2016) who point out that the lack of cooperation by internet service providers was a common theme in many of their interviews with police and other stakeholders, and how this represents obstacles to obtaining evidence to support legal action, as well as cost and time barriers to achieving justice for victims. Within their research, police also noted the increasing demand for forensic services for analysis of electronic evidence and hardware, highlighting how lawmakers do not fully comprehend the extent and types of cybercrime. In the Spanish context, when a conduct constituting "sexist cyberviolence" is reported, scholars have previously highlighted that legal operators frequently do not understand the magnitude of the evidence collected in digital spaces or lack the specific technical expertise to be able to use all the evidence (Igareda et al, 2019). If the case finally reaches court, most likely, it will not be brought before a court that is specialised on gender-based violence and, therefore, the sensitivity of the judge to the various forms of gender-based violence is highly arbitrary (Ibid.). Often this ignorance ends up re-victimising women, blaming them for what happened. Added to this, victims themselves often downplay the violence and therefore either do not report or do not keep the evidence (Ibid.)

It is also important to note that policing and criminalization of behaviours is not the answer to all types of online violence. At the international level, scholars (e.g., Dodge and Spencer, 2017; Coburn et al, 2015; Hasinoff, 2015; Shariff and DeMartini, 2015) have expressed concerns that criminal law approaches provide scarce opportunities to change the cultural beliefs that influence sexist, transphobic, racist, and homophobic harassment more broadly. This has also been recognised in relation to technology-facilitated sexual violence (e.g., Powell and Henry, 2016). However, the primary incentive for women to report misogynistic hate crimes to the police might not necessarily be their strong belief in conviction. Rather, what holds significance for the victimised individuals is the assurance that their complaints are being treated with seriousness.: "Officers need to ensure that they continue to give empathetic and sympathetic responses to female victims and take them seriously" (Mullaney and Trickett, 2018: 55).

This was echoed by some of the professionals interviewed in this thesis research, reflecting on the importance of considering the needs of the women in terms of reporting or not reporting:

Sometimes we also find that what a woman wants is for the violence to stop, what I want is for this guy who is harassing me on my mobile phone to stop. But she doesn't really know if she wants to get involved in a complaint, so how do I balance the right that I have to live in peace without anyone harassing me, with the fact that if it is a crime I have to report it so that certain resources are activated? (EX-14)

In other words, certain professionals do not necessarily regard notifying the police as the most effective method for curbing online violence against women. Simultaneously, some protocols or the activation of support services may necessitate a police report. Nevertheless, a considerable number, if not the majority, of gender violence support services accommodate women without requiring a police report:

Of course, here, for example, to care for a woman who is a victim of violence, we do not need her to have filed a complaint, it's not necessary. That is, we are working with the woman to empower her, to reduce the damage and if she in the future in an autonomous way, is eager and capable of opening that chapter again and place a report, then it's necessary to provide an accompaniment, because it has a greater impact, she may relive things again (EX-14).

I will further address the issue of reporting or not reporting to the police from the perspective of targets of online violence in Chapter 9. In the next section, I will analyse the gap in training and possible ways forward to better equip professionals with knowledge on digital spaces, and to better support victims of gendered violence in digital spaces.

6.5. Training needs and possible ways forward

Examining the interviews with professionals reveals a pervasive deficiency in training across various sectors, encompassing both public and non-governmental entities, in terms of safeguarding individuals potentially susceptible to online attacks. Additionally, professionals express a notable absence of guidance on appropriate responses in the event of an online incident or crime. Some are highly

aware of this and express this deficiency explicitly, as conveyed by the representative of a public information service for women: "A lot of training work will have to be done, both to the services that provide attention to women victims of this violence, such as the legal field, the police field, the lawyers on duty" (EX-14). The deficiency in training and technical knowledge is frequently articulated in the context of a generational gap:

I, for example, am older, and the whole digital issue is not my forte. There really is a significant gap in the age that I am, of some women that we attend to who are perhaps 18 or 22 years old and I think that we are lacking there, we try to change the mindset that it's something that we have to take very much into account (EX-16).

Part of this gap could be overcome with the provision of adequate training. Presently, lacking established protocols or formal training, many professionals rely on informal, personal knowledge derived from the experiences of team members. Typically, these team members are younger individuals, recognised for their perceived proficiency in social media. For example, in the case of a public service that provides support to victims of gender violence, their representatives acknowledge that they have not received training on digital protection mechanisms. Instead, they depend on young team members' knowledge based on their private use of social media: "Yes, we have a person in the team who is young and she's the one who tells us a little 'hey, remember that...'" and I think we need a bit of an update in that sense" (EX-16).

Training, however, is necessary not only to better address the online context, but also to broaden the approach to other types of violence, formulated here by the representative of a local institution that works with feminism and LGBT+ rights:

One of the issues that came up in a meeting between several services was the need for training for professionals so that they have a broader view of violence, because they continue to be trained in the view of intimate partner violence, so to open up the perspective in that sense, that's a challenge, huh (EX-12).

As discussed in Chapter 5, the regional Catalan law against gender violence was recently modified in *Law 17/2020, of December 22, modifying Law 5/2008, of the*

right to eradicate male violence.⁵⁹ Beyond purely criminal matters, this law could provide victims of online violence with access to the support mechanisms currently available for victims of gender violence.

6.6. Conclusion: A hierarchy of violence, challenges of technological development and protection gaps

While the primary objective of this research is not a comprehensive analysis of the broader issue of online gender violence, but rather to concentrate on online gender-based hate speech, the consistent redirection to services catering to victims of gender violence has prompted me to approach the topic more broadly in discussions with professionals. This adjustment is necessitated by the fact that existing services are tailored to address gender violence as defined within the Spanish legal framework.

The nature of technological change creates pressures for renewal and updating of knowledge to reflect the rapidly evolving nature of digital spaces. The qualitative research showcased in this chapter has shown how professionals in Spain are largely unprepared to provide support to victims of gender-based online violence, as they tend to lack specific training and protocols and instead need to rely on informal knowledge. Nevertheless, as has been noted elsewhere (e.g., Burruss et al, 2020; Dodge and Spencer, 2017; Powell and Henry, 2016), it is indeed challenging for professionals to keep up with developments in technology. The chapter has argued that the online-offline continuum is largely perceived by professionals as a binary division between the "real" and the "virtual". This contributes to the perception of violence suffered online as "not real", despite its often deep impact, accentuated by the pervasiveness of digital spaces through the use of mobile phones and the continuity of the violence. Within this paradigm, gendered online violence is at the same time perceived as inevitable (and normalised), whilst it is largely invisibilised within the hierarchy of violence (as compared to physical violence). In more concrete terms, this perception of online

⁵⁹ LLEI 17/2020, del 22 de desembre, de modificació de la Llei 5/2008, del dret de les dones a eradicar la violència masclista, available at <https://portaljuridic.gencat.cat/ca/document-del-pjur/?documentId=889760>, last accessed on April 19, 2024.

spaces and online violence has an impact on the support available for those who are targeted by gendered online violence.

However, this difficulty to keep up with technological advancements is not only due to the rapid technological developments, but also linked to the lack of resources. The lack of resources allocated to this problematic can at least in part be explained by the low level of importance given to events – however violent – that take place in digital spaces unless these pose a risk of immediate physical violence to those targeted. Within the predominant view of the online and offline as binary and separated spaces, online violence is placed on the lowest steps of the hierarchy of violence. Despite a certain level of awareness among professionals of the impact that online violence may have on those targeted, in practice little is done to provide support to women in digital spaces.

Further, there is a clash between what is legally included under the hate crime umbrella and the categorisation and implementation of the legal framework on the ground. Barbara Perry (2001) analyses gender violence as a "classic" example of hate crimes, in the sense that this violence aims to reproduce heteropatriarchal structures and put an entire group of people in their place through the victimisation of an individual. In theory, any violence aimed at attacking a group could be analysed within this framework. International organizations such as the OSCE Office for Democratic Institutions and Human Rights have reinforced this perspective on gender-based violence in recent years, and we have also seen examples of efforts on the ground, such as within the police force in Nottingham, UK, which has explicitly included misogynistic crimes in its work against hate crimes (e.g., Mullaney and Trickett, 2018). However, despite the inclusion of gender as an aggravating circumstance in the Spanish penal code, this perspective is still far from a reality in Spain. Both police services and other institutions still have a clear divide between catering to victims of hate crime (defined predominantly as victims of racist and anti-LGBT hate crimes) and victims of gender-based violence (defined as victims of violence perpetrated by a partner or ex-partner).

Women who are victims of other types of sexist violence, such as online hate speech or digital violence in general, tend to fall into a gap between categories. These crimes or incidents are generally not labelled as "gender-based" nor as "hate crimes". The ramifications of this classification gap encompass the absence of a comprehensive structural analysis and an understanding of the collective impacts of such violence. This gap further results in a lack of statistical data and specialised support services. Presently, the failure to recognise gender motivation in crimes beyond the realm of intimate relationships in the digital context leads to a deficiency in much-needed expertise, both in addressing the victims' needs and handling technical aspects. This deficiency in expertise and awareness regarding the unique nature of these crimes may have implications for the entire justice and support process, spanning reporting, prosecution, court proceedings, and outcomes, with particular significance for the overall experience of the women who are targeted by this violence.

Services are changing to a certain degree, for example in Catalonia where the autonomous legal framework defines gender-based violence in accordance with the Istanbul Convention. Within this framework there are examples of city councils that make a conscious effort of widening the scope of their services in order not to cater only to victims of intimate partner violence, but also to victims of other types of gender-based crimes, including LGBT persons.

Framing gendered online violence within a hate crime framework can facilitate an understanding of the structural roots and the collective silencing impacts of this violence. Analysing it on a continuum of violence makes us see how it is connected to other types of structural violence, and how different expressions of gendered or patriarchal structures may be lived as violence even if the criminal code does not define them as offenses. To counter this gendered violence, there is a need to establish a multi-level networked response, not only based on policing and criminal code, but reliant upon a range of different resources and strategies, including education and awareness-raising, improved digital protection measures, and enhanced collaboration with social media platforms. In addition to this, different types of support would need to be offered to the individuals who

are targeted by online violence, including legal and psychological support. To construct this networked response, it is essential for professionals and practitioners to be equipped with sufficient knowledge to be able to, on the one hand, analyse and address the violence from a structural approach, and on the other hand, provide an empathetic attitude and realistic advice to the individuals who are directly targeted.

7. Targeted voices: Experiencing gendered online violence

...a lot of guys quoted me with my profile photo saying, 'ok you're not so smart putting this on Twitter and you're hot and I'd fuck you', and some barbarities, and there were people who incited to rape me, incited to... to kill me basically (Daniela)⁶⁰

7.1. Introduction

Women's use of social media technologies for feminist activism has been signalled as indicating a fourth wave of the feminist movement (e.g., Evans, 2015). Over a decade ago some scholars already suggested that "Tweet or perish!" is "an inevitable [yet often unquestioned] axiom for feminist dialogue and organising" (Gurumurthy, 2011: 466). Social media and the internet at large are crucial arenas both for feminist debate and for feminist activists to organise themselves. At the same time, feminists' use of public social media channels

⁶⁰ This is a quote from an interview with a woman targeted by online violence. As discussed in Chapter 4, on methodological and ethical considerations, the names used throughout the thesis are pseudonyms, in order to protect the anonymity of the persons interviewed. For an overview of this sample, please see Chapter 4, section 4.3.2.

renders their activism highly visible, traceable, and open to resistance, as the use of these platforms has facilitated increased access to movement ideas and individual women by feminist adversaries (Megarry, 2017). In contrast to the spaces of resistance used by the second wave of feminism, which were often constructed as separatist spaces free from men, the use of public social media platforms makes feminist activists vulnerable to hostile interventions (Megarry, 2014; Citron, 2014). That is, social media platforms are arenas where feminist voices are both highly accessible and highly challenged; however, as this chapter will demonstrate, they are challenged not only for their activities online, but for their participation virtually in any public space or debate that reaches outside of a feminist circle.

In this chapter, I start telling the stories of the women who are targeted by this type of online violence. The experiences included by no means constitute a complete picture of the experience of being targeted by online gender-based hate speech in contemporary Spain. However, they do bear significance. As previously reflected in Chapter 3 and Chapter 4, the sample of targets aimed to include a heterogeneous group of persons, but with the common point of being women active in debates on feminism and related topics, and with experiences of having been attacked online based on their identity and/or opinions. In addition, as active feminists, they are also aware of experiences of their peers, as well as the wider social context of these experiences.

The analysis exposed in this chapter starts with submerging in the participants' online contexts, that is, their different uses of digital spaces. This is followed by an analysis of their experiences in relation to a "continuum of violence" (Kelly, 1987) that becomes apparent in the narratives of most of those interviewed. Here, I start by discussing the experiences that are directly related to their feminism and other types of activism, from a gender perspective. However, not only gender is of importance, but also other identity traits that interlace with gender, such as sexual orientation, racialisation, and gender identity. Following this, I explore other experiences of gendered online and offline violence that those targeted include in their narratives. The inclusion of this wide range of episodes in

the stories of these women specifically signal the existence of a continuum of violence, with a range of experiences that express and recreate the same heteropatriarchal structures. In the final sections, I examine the women's perceptions of the perpetrators, before reflecting on the normalisation of online violence in the lives of feminists.

7.2. Life online: Between work and activism - Between public and private

In most interviews conducted, the initial inquiries revolved around the participants' online trajectories. Specifically, I asked about their use of various social media channels and the purposes for which they engaged with them. This deliberate approach served as a measured introduction to the interviews, strategically avoiding immediate immersion into the central theme of violence. Furthermore, this method facilitated a contextualisation of their narratives within the digital spaces they frequented. The primary focus of the experiences discussed in this chapter in terms of platforms is Twitter, as this is the main field that I defined for the digital ethnography. Nevertheless, also other platforms were brought up and discussed in the interviews. This is to say that the research conducted is not platform specific. As shall be exposed below, some attacks start on one platform and continue on another or are perpetrated in parallel on several platforms. Therefore, as previously mentioned, this thesis adheres to a perspective of polymedia (e.g., Miller and Madianou, 2013), that is, considering how users can exploit a range of opportunities for communication, as opposed to being constrained by the specific affordances of each medium.

However, although violence may be perpetrated through various channels at once, and spill over into offline threats or fear of attacks, the specific channel may still be of significance in terms of the public or private character of the message. In this regard, the channel often determines the conduct and experience of being targeted (e.g., Calala, 2020). For example, on Twitter, attacks are often public, whereas on channels such as messaging platforms, hostility is to a higher extent invisibilised as it reaches targets through individual messages instead of comments on a visible profile or publication. In this regard, platforms do not just mediate public discourse, they constitute it (Gillespie, 2020). That is, "sociality is

not simply ‘rendered technological’ by moving to an online space; rather, coded structures are profoundly altering the nature of our connections, creations, and interactions” (van Dijck, 2013: 20). Platforms are crafted to encourage and influence our engagement towards specific objectives. This encompasses the structuring of profiles and interactions, the preservation of social exchanges, the pricing or payment mechanisms for access, and the algorithmic organization of information. This algorithmic organization often prioritises certain content over others, including determining what is prohibited and how such prohibitions are enforced (Gillespie, 2020).

Common for all the targets interviewed is that they use social media platforms for activist aims and for debating or expressing opinions on feminism and other political topics. Most of them are highly active on several different platforms. However, they tend to use different platforms for different purposes:

On Instagram I focus mainly on lifestyle. Sometimes through stories I post screenshots of articles I have published on other sites or of activities that I do (...) And then on Twitter, I’m harsher and more straightforward (...) People follow me on Twitter mainly for my activism. I do separate the topics that I address on each social media platform (Amara).

Daniela explains that “well, I’m quite young so I kind of grew up on social media. I think I started using Instagram when I was 13 or 14 years old, and the same with SnapChat. Later on, I got off SnapChat because I didn’t use it much”. There is a clear generational difference between the youngest person interviewed (Daniela, 22 years) and the oldest (Alba and Amara, 42 and 39 years), in that the latter have experience of and still to some extent use Facebooks and blogs, whilst Daniela initiated her social media trajectory in her teens on the newer generation social media platforms.

In comparison to the other platforms, Twitter is to a higher extent used for debate, for activism and for professional reasons, as has been expressed by Amara, above. Likewise, Daniela expresses that “I created the Twitter account because I represent an association so for pseudo professional reasons” (Daniela). Also those targets who are freelance journalists or communicators highlight the crucial importance of Twitter for professional reasons – to gain visibility and thus

more work projects. Twitter offers “an open register of feelings and opinions about issues of all kinds, including hate speech or other signs of rejection that are expressed freely and without the barriers that are often present in offline spaces” (Arcila-Calderón et al, 2020: 23). The platform has a capacity to viralise content with a fast speed; however, it should be kept in mind that it is not a medium that necessarily represents all citizens. All the women interviewed refer to Twitter as the most hostile platform, with the exception of Elena, who has experienced extreme verbal violence on gaming platforms, as I shall discuss further below. Amara explains that “Twitter is the jungle, and a battlefield, and it’s the least friendly social media of them all” (Amara). In the cases collected during the digital ethnography, this view of Twitter was also widely supported. In a video, DE-2 describes this platform as the “most sexist and misogynist social media” and goes on to explain that many of her feminist activist friends experience the same type of cyberharassment as she does on this platform.

Instagram, on the other hand, is described as a platform largely used for updating friends and family on life events. That is, it is experienced and used as a platform much more related to the “private” sphere:

Instagram was like private and more personal, for my friends and so on (...) Since I started having a bit of a boom on Twitter, when I started getting a lot of followers, I started getting a lot of requests to follow me also on Instagram, so I made it public, and then I kind of accept that those two social media channels are accessible by anyone that has an internet connection (Daniela).

At a first glance on the narratives, Instagram is generally described as a more “friendly” platform in comparison to Twitter. This is particularly true for those who reported not using Instagram primarily for activism:

The Instagram audience is a super-friendly audience, all the comments are super-friendly. Everything is beautiful and everything is good. I think since I started the account, I have never seen a negative comment (Amara).

Nevertheless, several of the women disclosed that during instances of being subjected to extensive attacks, they encountered a comparable volume and nature of hate messages and comments on both Twitter and Instagram. This is

illustrated by the cases of both Sara and Lucía, who even though they report using the two platforms for different purposes, have been targeted by violence on both platforms. Similarly to the other women interviewed, they use Twitter as the more professional platform for disseminating articles and other content they produce as communicators and journalists, whilst Instagram is used for longer, less formal and more personal reflections. Lucía highlights that the channel where the attacks take place largely depends on where the original message that targets her has been published. Nevertheless, despite this, she describes Instagram as “much more friendly”. That is, there is a certain ambivalence to describing Instagram as hostile, whereas this is the main descriptor for Twitter.

Some of the targets interviewed also debate politics, feminism, and other topics on Instagram. In these cases, they tend to have two profiles: an open profile that anyone can see and follow, and a private, closed profile where the owner must accept those who request to follow them. In the narratives of those interviewed, this is primarily linked to precaution, and being able to publish personal or private information to a smaller circle:

I do watch what I publish in the sense that I don't publish my address or where I live or things like that because I get a bit of a bad vibe and the world is a bit like... bad. I do have another Instagram account where I do talk about my life to my friends but it's private and the great majority of those who follow me on other channels don't know that it exists (Daniela).

Here we can see how precaution, risk and fear sip into the narrative already before entering the topic of violence. That which can be shared publicly without risk is clearly separated from that which needs to be kept to a smaller, private, audience, to avoid harm.

Other channels used by the interviewees are YouTube, Discord and Twitch. In this study, the latter two have only been referred to and used as platforms for discussing and streaming gaming by the gamer interviewed (Elena). At the time of the interview, Elena was trying to build an audience on Twitch, having previously had quite a successful YouTube channel with an audience largely consisting of lesbian women. Talking about this channel, she mentions that she hasn't “suffered harassment or anything like that, because it's for women,

targeted to women, at least because of that I didn't suffer from any of that" (Elena). She perceives the Twitch context, though, as much more unpredictable, and male-dominated, as is the gaming world in general:

Well let's see now on Twitch. They're going to... Women in this world of videogames and so on are superattacked. It's a men's world and we can't get in. You know, it's theirs. And I know that I'm going to receive threats and insults, but I don't care, I'm going to do it [she laughs] (Elena).

That is, due to her perception of a high level of participation of men on Twitch, she already anticipates that her experience there might be more violent, whilst the predominantly female and lesbian following that she had managed to create on YouTube made her channel there feel like a safe space. As we shall see, it is predominantly when women try to enter male-dominated contexts or discuss feminism or other equality-related topics outside of a circle of like-minded that they are attacked. In the next section, I will discuss this type of attack.

7.3. Being targeted as a feminist-, anti-racist or LGBT+ activist

As I have explained above, the common factors for all the persons included in the fieldwork are, first, that they define themselves as women; second, that they publicly debate issues related to feminism and gender equality, LGBT rights, and anti-racism; and third, that they have suffered online hate speech or other types of attacks. That is, the sample has an important impact on the types of experiences of online violence that these persons relate in their narratives, as their role as feminist and political voices in online debates was the very basis for their inclusion in the study.

The incidents prompting my interviews with these individuals or their inclusion in the digital ethnography were directly associated with their roles as activists in the realms of feminism, LGBTQ+, or anti-racism. Consequently, attacks, harassment, and abuse rooted in these activist roles constituted the primary focus of discussion during the interviews and served as the central theme for data collection throughout the digital ethnography. It is crucial to acknowledge this delineation in scope when examining the experiences presented in this empirical

segment of the thesis. Subsequent sections will delve into diverse forms of experiences along the continuum of gendered violence.

7.3.1. Triggers, waves and getting into the “mainstream”

When discussing online violence targeting these women based on their activist or professional profiles (as opposed to targeting them “only” based on their gender and/or other intersecting factors), the instances of violence were often described as having taken place in relation to a specific appearance in media; a tweet or other publication on social media made by those targeted; or a publication by a specific anti-feminist or transphobic public figure that triggered a wider response in the shape of attacks by the followers of the attacker. That is, most of the persons interviewed clearly perceived a “trigger” that unleashed the attacks.

Women subjected to attacks due to their feminist positions have awareness of the specific topics that typically provoke a backlash against them within the public discourse in contemporary Spain: Feminism and gender equality in general, as well as issues related to trans people’s rights, and any content that criticises the so called “manosphere”. In relation to the last point, Daniela explains that a humorous meme that she tweeted about Jordan Peterson⁶¹ led to a wave of attacks from his Spanish-speaking followers: “What I didn’t take into account was that the fans of Jordan Peterson, who is an openly fascist person, well they said some barbarities in some messages, and in quotes on Twitter, telling me any kind of thing really” (Daniela).

However, also seemingly innocuous topics such as astrology have been reported as triggering attacks, particularly when discussed by women who define themselves as feminists: “Astrology is a sensitive topic because it stands out a lot and I think it triggers certain discourses and certain violence” (Aurora). The same is the case for any discourse going against an interest traditionally coded as “male”, such as football: “The problem is that I messed with football. Or I didn't mess with football, I said 'many people let football condition and determine their

⁶¹ Jordan Peterson is a psychologist, author and media commentator, who has acquired a large social media following of men with anti-feminist stances, through a YouTube channel where he uploaded lectures linked to so called “free speech” on topics related to gender, trans rights etcetera.

lives', the schedules, the moods'" (Lucía). That is, the acts that supposedly trigger misogynist attacks can be understood as interchangeable, "in that a woman could receive a near-identical spray of rape threats for commenting about basketball games, bike riding, comic book covers, soft pretzels recipes, DIY fishtail braid videos on YouTube, and on and on ad infinitum" (Jane, 2017: 11). Lucía links this manner of certain persons jumping on basically any topic to her condition as a public feminist and adds that "there are people who are quite angry with me" (Lucía). That is, she perceives that people with strong contrasting opinions take any chance to insult or ridicule her.

Further in terms of triggering topics, feminism itself is a major trigger. Sara highlights how even before she appeared on a social media channel of a national-wide TV show, the mere announcement of a feminist talk triggered hate messages:

The fact that I talked about feminism was the trigger for this kind of hate (...) When it was announced that I would give a talk on feminism, there were already hate messages in relation to that announcement. Even before the talk... (Sara).

It was not necessarily the *content* of the talk that infuriated those attacking her, but instead the very fact that the concept of feminism was brought into the mainstream view. In this regard, online abuse against feminists seeks to contest both their expressed views and their fundamental right to participate in debates and in public spaces (Lewis et al, 2019). As both the interviews and the digital ethnography of this thesis reveal, an important factor of whether a tweet, a publication or a media appearance leads to attacks seems to be that it reaches a wider circle and "go into the mainstream" (Sara) or as Daniela puts forward, "reaches the incel timeline". Once a target has reached that wider circle once, they are also more exposed to subsequent attacks, as illustrated by the case of Daniela: "I got to a point where I started to become quite overwhelmed because anything I published, if a tweet gets a bit viral, you start receiving [attacks]" (Daniela).

Massive harassment often starts when the targets have discussed feminism – or simply appeared as feminists – on a mainstream platform, as is illustrated by the experience of Sara:

I went to do a sort of workshop during the week of the International Women's Day (...). And that generated a mob I think without precedents and that was the first time I became 'trending topic'. The most paradoxical and ironic part of this is that the feminist talk wasn't the subject of debate, but it was [an aspect of her physical appearance] (...) So on Twitter there was a great number of messages that were directly targeting my physical appearance. And this was the biggest mob I've had because it was on a national level. Conservative political parties entered the debate and I think it was my big leap into the mainstream, and into this hate on a very massive and very organised level (Sara).

However, as expressed above, the attacks started even before the talk, when it was announced that she would talk about feminism, so attacking her physical appearance was clearly a way of channelling and expressing resistance towards feminism. In the narrative of Sara, she perceives herself as quite naïve before this event, having managed to discuss feminist topics previously within a safe community as a journalist, and even on Twitter:

I also think it may be that I realised the great exposure that the mainstream implies. I had been very protected in a community of readers, who more or less sympathised or were active in similar issues. I also discovered that vastness and that total divergence of opinions and I think it made me a little more aware of the repercussions that certain comments could have (Sara).

Sara reflects further upon why she was targeted:

I think there's like an impulse fostered by the technologies themselves that allow us to quickly say what we want to that specific person and thus this enemy is generated that we can go and insult with relative ease or with relative impunity. And what are they after? I just don't know... (Sara).

Technologies are perceived by those targeted to have one part of the blame, as they give people the opportunity to immediately react to a message or disseminate hate without risking any greater repercussion. Nevertheless, this perspective emphasises the role of digital technologies in granting access to these visible targets and enabling online attacks to occur with impunity, rather

than attributing hate creation or propulsion directly to the technologies themselves.

Further, Sara describes how the harassment continued for quite a long time. Even when the original attacks had worn off, men kept sending her private messages on Instagram asking whether they could ask her a question, and when she said yes, they insulted her body in the same way as in the original attack. As we shall see in the last sections of this chapter, this repeated and continuous behaviour by many different men gives rise to a perception that attacks are often much more organised than they may seem at first glance. It has not been until recently, a few years after this attack against her, that Sara is finally “out of the mainstream” and has not received any hate messages in a while:

Now it has calmed down. It’s true that I was there until not so long ago. About a year or two had passed and I kept receiving messages from some kid, but now I don’t. But I participate in a section of public television, and I never share the videos, and I have asked myself this many times and I have no answer, if I don’t share the videos because I’m terrified that it will happen again, or I just don’t share them because I don’t feel like it? And the truth is, I don’t know (Sara).

That is, years after the attack she is still on guard and wary that media appearances may unleash new attacks. Along the same lines, Julia who is an LGBT+ activist, has come to expect that any media appearance will lead to some kind of negative reaction:

You like already assume that... For example, the other day when I did the interview [on TV] I knew that they would put it up [on the channel’s social media] and that there would be bad messages, so you already expect it, and it’s like, let’s see where they get you today. And that’s quite crappy now that I think about it (Julia).

This is similar to the experience of DE-8⁶², who comments that every time she appears on mainstream television, she receives hate messages. She specifically mentions that every time that she participates on mainstream television, “things that she hasn’t said” go viral on channels and profiles of the extreme right. As a

⁶² The cases from the digital ethnography are coded as DE-n^o. For an overview of this sample, please see Chapter 4, section 4.3.1.

visually Muslim woman, the very appearance on mainstream platforms seems to trigger online attacks from certain elements of the population, which she perceives as being men from the extreme right of the political spectrum. Also DE-15 and DE-17 report having been targeted by hate speech and racist discourses just by being racialised women appearing in mainstream media. Further, as a Roma feminist, DE-17 has been subjected to repeated gendered and Roma-phobic abuse across various platforms. In a tweet, she denounces that following a TV appearance, she has been targeted by more than 100 abusive messages in the comments to the post about the appearance on the TV channel's Facebook page. Through the tweet, she not only denounces this abuse, but also publicly asks the TV channel to please moderate the comments, which they subsequently start doing.

Similarly, among trans women, in many cases their mere presence on social media triggers attacks against them, particularly in the current societal climate in Spain (and beyond) where trans people's rights constitute a highly polarised arena for debate (e.g., Willem et al, 2022). The digital ethnography illustrates this very clearly. Both DE-6 and DE-9 have been targeted by transphobic discourses after merely sharing photos of themselves, in which they do not adhere to normative constructions of gender. In the case of DE-6, she posted a photo where you can see a slight shade of beard and moustache, with her own comment "With beard and tits, a whole woman". This tweet was cited by several profiles adding extremely transphobic comments. Along similar lines, DE-9 shared a screenshot where a person defining themselves as a radical feminist has taken one of her tweets where she appears in a photo and has cited the tweet along with explicitly transphobic discourse.

Another aspect of gender-based attacks is their often wave-shaped characteristic, particularly when the attacks are perpetrated by many persons. In this regard, the case of Lucía can be used as an illustration. Lucía specifically recounts three distinct waves of attacks directed at her over a brief duration. In two of these instances, notable individuals with opposing perspectives initiated the attacks by negatively mentioning her on their social media platforms.

Subsequently, their followers engaged in targeted harassment against her.⁶³ In this context, understanding the pivotal role of the attacks' origin from two prominent profiles becomes crucial. During the interview Lucía points out - in passing - that she has always to some extent been targeted one-off by hate as a feminist with a public profile; however, it was the massive extent of the three waves that she describes as having a major impact on her. When referring to wave-shape I mean that these events had a clear starting point and kept growing until they reached a peak, after which they slowly decreased. As we have seen above in the case of Sara, the decrease is sometimes very slow and extended, as the abuse continued for a couple of years after the original attack and peak. In Lucía's experience the starting point was the mention on social media by a big profile. The attacks then kept growing as the followers of the original perpetrator joined the attack. In her case, the attacks went on for 6-7 hours in one case and several days in another case, before they clung off. In the case where the attack lasted "merely" 6-7 hours, it only wore off because she decided to make her Instagram account, where the main attacks took place, private. The implication of this is that a seemingly isolated comment by a large profile may quickly turn into a growing attack, out of control of the original poster. The original attacker can thus take their hands off the attack, whilst the target may suffer from continuous attacks by followers of the original attacker for a prolonged period.

In the next section, I shall look further into the discourses and content of the attacks, from the perspective of those targeted.

7.3.2. Discourses and content of the attacks

Sexual assault can be analysed as an institutionalised rather than aberrant means by which (heterosexual) men can perform their masculinity while symbolising and actualising women's subordinate position (MacKinnon, 1993). This becomes relevant in relation to the language used in online hate against women and particularly against feminists, as this is oftentimes sexualised and frequently contains rape threats. Scholars such as Jane (2018) has referred to this

⁶³ I will delve deeper into these attack patterns in section 7.6 of this chapter, where I explore the targets' perspectives of the perpetrators.

discourse - which shares many characteristics across countries and online cultures - as “rapeglisch”. The striking homogeneity of the rhetoric of gendered cyber-hate throughout different contexts presents what Jane highlights as “a diachronic perspective on gendered e-bile by showing that it has been circulating for many years, and that (...) its rhetorical construct has remained remarkably stable over time” (2014: 560). In this regard, she puts forward explicit examples to illustrate several recurring characteristics:

They target a woman who is, for one reason or another, visible in the public sphere; their authors are anonymous or otherwise difficult to identify; their sexually explicit rhetoric includes homophobic and misogynist epithets; they prescribe coerced sex acts as all-purpose correctives; they pass scathing, appearance-related judgments and they rely on ad hominem invective (Ibid.).

These characteristics can still be found in the online hate expressed against feminists and other women activists who raise their voices in the public debate, almost ten years on from Jane’s 2014 article and in another social and geographical context. The main exception compared to the situation in 2014 is the anonymity of the authors. As I shall discuss in later sections of this chapter, anonymity does no longer seem to be a necessary condition to perpetrate gendered online attacks.

In the research underpinning this thesis, the discourses and content of the attacks do not primarily contest the feminist discourses expressed by the targets. In fact, the content of the targets’ feminist discourses was rarely debated by the attackers – further than expressing offense at feminist opinions at large:

My feeling from the comments I received was like they wanted to call me unrepresentable, understanding feminism as a great setback, they felt particularly offended by the speech I had made, in some cases. Others directly had no greater intention than to humiliate and denigrate me (Sara).

That is, the perception of those targeted is not that the attackers are interested in debating the topics under discussion, but rather that they aim to ridicule, humiliate and denigrate women who express feminist opinions. That is, the

attacks are not based on arguments or ideological debates, but rather on insults and discredits (cf. Morena-Balaguer et al, 2021).

The discourses and content of the messages targeting the persons interviewed can be located along a continuum with death threats on one end, and seemingly less serious insults on the other end (cf. Lewis et al, 2017), such as insulting the targets' professional roles, including repeatedly telling a comedian that she is not funny, or telling a communicator that she is bad at communicating. Although this suggests that some experiences are not directly threatening, this does not imply that they were necessarily experienced as less impactful or harmful. In this regard, it is important to keep in mind that online violent expressions may not reach the threshold of being criminal, however, they may still be *lived* as violence. The mere repetition of a message enough times, and by different persons, may make a person question themselves, both personally and professionally As highlighted by Harris and Vitis (2020), "some behaviours may be viewed as typical and normalised, others problematised or criminalised, but all have the potential to have an impact on women's wellbeing, sense of safety, rights and freedoms".⁶⁴

Somewhere between these extremes of the continuum, and in lines with much of the international literature referenced in Chapter 3 (e.g., Jane, 2014, 2017; Megarry, 2017), we find sexualised messages targeting or discussing the women's physical appearance, rating them as sexual objects or posing sexualised threats:

In ForoCoches I remember that they made a thread for me where they put an image of me from Facebook and put "melafo?" that is "do I fuck her?" and different men had to say yes or no and argue for their answers. And this was like my first experience in these ultra-masculine, ultra-heterosexual forums, and where I was already beginning to see a very real violence in which the physique was basically checked and from then on, a whole series of atrocities were said (Sara).

⁶⁴ I shall discuss this further when addressing the consequences and impacts of this type of speech in Chapter 8.

These sexualised discourses are often coupled with direct threats of rape or even murder, and may move quickly from a comment on a woman's appearance to rape threats or fantasies:

I don't remember exactly but I got to 'Incel Twitter' again and a lot of guys quoted me with my profile photo saying, 'ok you're not so smart putting this on Twitter and you're hot and I'd fuck you', and some barbarities, and there were people who incited to rape me, incited to... to kill me basically (Daniela).

Rape can be understood as a punitive and disciplinary act against women (Segato, 2003), that is, a rape threat can be interpreted as an explicit intent to put a woman in her place, and thus as a reinforcement of a power dynamic. Another common sexualised expression is to point out that a woman has not been "fucked well enough" such as in the case of DE-13: "I know you, you're the typical poorly fucked [malfollada] who never got dick" (Comment on DE-13's Twitter post). As mentioned above, men's threatening and sexualised discourse against women is highly repetitive across time and culture. The overwhelmingly impersonal, repetitive, and stereotyped characteristics of the abuse show that the discourse is not about individual targets but about gender (Jane, 2017). Rita Segato (2003), in this regard, when referring to rape as a disciplinary and vengeful act against a generic woman who has "left her place", highlights this generic character of the woman. That is, the act of rape, similarly to rape threats, is more linked to gender and power than to the specific target. The mere displacement of a woman to another position where she is not destined according to a traditional hierarchy, calls into question the position of the man in this structure, or in other words, the power of one does not exist without subordination of the other (Segato, 2003: 31).

In addition to sexualised content, quite common in the discourses described are references to a lack of mental health or intelligence, that is, expressions that state that these women are either crazy or stupid. For example, Naomi recalls how a man with a public profile in the cultural sphere referred to her on Twitter as "mentally destitute", and DE-19 shares screenshots of a man calling her "rude,

uneducated, unstable and ‘loca del coño’⁶⁵”. This can be understood as another way of subordinating women who express feminist ideas; arguing that they are not in their right minds, and that thus their discourses can –or should- be dismissed.

Finally, almost all the targets interviewed as well as many of those included in the digital ethnography have been called the catch-all insult, “feminazi”:

Suddenly a comment 'feminazi, communist'. The feminazi thing has already become fashionable before, and it is still holding up. Out of nowhere a comment that 'you are a feminazi'. I don't feel insulted... More proud. In my profile it says that I'm a feminist. And of course, when they get into my profile and read that, well, they attack me with feminazi and so on (Elena).

Several Spanish scholars have previously highlighted the use of the concept “feminazi”, as a derogatory term to address feminism in general and feminist women in particular (e.g., Crosas and Medina-Bravo, 2018; Megías et al, 2020; Villar-Aguilés and Pecourt Gracia, 2020; Nuñez-Puente et al, 2021; Vergès-Bosch and Gil-Juárez, 2021; Piñeiro-Otero and Martínez-Rolán, 2021; Morena-Balaguer et al, 2021). The use of the word “feminazi” can be understood as a way of undermining the reputation of feminist women and accusing them of crimes and reprehensible actions, through the association that is made between feminism and Nazism through this concept (e.g., Crosas and Medina-Bravo, 2018). “The use of the lexicalized term ‘feminazi’ is one of the ways in which the concept of ‘gender ideology’ has been mobilized in media discourses” (Nuñez Puente et al, 2022: 52), for example, through the circulation of the hashtag #StopFeminazis on Twitter. By employing a mirrored discourse featuring the term “feminazi,” users primarily affiliated with the organized far-right seek to appropriate the victim narrative previously held by the “other,” namely the so-called “feminazis” (Ibid.). Consequently, feminist activism is portrayed as synonymous with violence, propaganda, or depicted as a “tyranny” or a “lobby,” thus portraying it as a threat to a masculinity perceived to be in crisis.

⁶⁵ A Spanish expression to refer to women as hysterics, too emotional, and prisoners of a hormonal volatility (Mauri, 2022). <https://letraslibres.com/cultura/las-locas-del-cono/>, last accessed on April 12, 2024.

Nevertheless, whilst this term is intended as a pejorative, over the last few years, feminists have jokingly taken to this concept, and have to a certain extent reappropriated it. We can observe an example of this in the quote above, by Elena, who says that she doesn't feel insulted, but she rather feels proud. Similarly, there are examples of this humorous reappropriation in media and culture, such as the video series by the national newspaper El País, "Feminismo para torpes" ("Feminism for dummies") where the first video is entitled "Am I a feminazi?"⁶⁶; or the book "Put a gorda, feminazi" by the comedian and feminist Penny Jay. This can be compared to the reappropriation process of other labels that were originally intended to be insulting or pejorative, such as "queer" (e.g., Butler, 1997; Eichhorn, 2001). The revaluation of terms such as "queer" or "feminazi" suggests that speech can be "returned" to its speaker in a different form, that it can be cited against its original purposes, and perform a reversal of effects (Butler, 1997). To be interpellated, even through these types of labels, carries with it a potential to empower individuals to challenge the assumptions on which they've been defined and possibly reclaim labels that historically contributed to their marginalised position. (Ibid.).

Finally, whilst I have previously stated that misogynist discourse is strikingly similar across countries despite the enormous variation in the cultural, political, geographical, and temporal contexts of the women who are targeted, this statement needs to be nuanced. Racialised, Muslim, and LGBT+ women, are targeted for particularly vicious attacks where threats of sexual violence are routinely combined with extreme racism and homophobia (e.g., Jane, 2017). That is, intersectionality plays an important role, which I shall discuss further in section 7.4 of this chapter. In the next section, however, an examination will be undertaken of instances of online violence that do not directly align with the subjects' identities as feminists or activists. These experiences are positioned along a continuum of gendered violence and intricately interwoven within the narratives of the affected individuals.

⁶⁶ El País, "Feminismo para torpes" (March 5, 2018), available at https://elpais.com/elpais/2018/03/05/mujeres/1520276469_431670.html, last accessed on April 19, 2024.

7.4. Gaming, dick pics and unsolicited contacts: Other gendered experiences of violence

As exposed above, the common factor for the women interviewed is that they have suffered attacks because they speak publicly about political topics or define themselves as feminists. Nevertheless, despite this alignment between their profiles, the spectrum of violence that they refer to is broad. When asked about attacks, harassment or violence online, many of the targets expose a wide range of experiences that can be placed on a continuum of gender-based violence, often with further factors than gender intersecting these experiences. This parallels findings from surveys like the one conducted by Lewis et al (2019), where most women who participated had experienced more than one kind of online abuse; and the Spanish survey by DonesTech (Igareda, 2019), where 70.0% of respondents reported that they had been targeted by more than one type of online ⁶⁷

Along this spectrum, the individuals described also recounted a variety of experiences that aren't directly tied to their roles as activists, but rather to their identities as (LGBT+, racialized, trans...) women in an online setting. First, there is the experience of the gamer, who addresses distinct instances of violence across different contexts; nevertheless, she describes the world of gaming as the most violent for her as a woman, although she does not necessarily have an activist agenda when she engages in online gaming:

I remember that I was getting into a game and receiving, receiving, receiving. There wasn't a game in which they didn't insult me when they discovered that I was a woman. And it didn't matter if they were Italians, French, Germans, they knew four insults in Spanish; I tell you, the gaming world is a very violent world (Elena).

However, not all the gendered experiences she refers to in the online context are episodes of explicit verbal violence:

Normally there are three things that happen to you when they discover that you are a woman. One, that they want to help you a lot and are very

⁶⁷ The most prevalent sexist violence online reported in this latter survey was related to insults or discriminatory or derogatory expressions for being a woman (54.6%).

slimy in case you finally give them your phone number, another that they directly send you to suck dicks, and the other that has also happened to me a lot, especially with Latin Americans, is that they send you to clean, to stop playing and go clean. Very nice huh... (Elena)

As this quote illustrates, there is also more subtle or symbolic violence involved, either through downplaying women's competence, offering to help; through referring to them as exclusively sexual objects; or "putting them in their place" by referring to traditional gender roles and activities such as cleaning or doing dishes. These more subtle expressions are not exclusive to the world of gaming, but there are also several examples of references to women staying within traditional gender roles in the digital ethnography. For example, in the case of DE-7, when she publicly denounces a case of online hate, a Twitter user tells her that no bad things would have happened to her if she had dedicated her time to cleaning instead.

The women interviewed also refer to other types of harassment or episodes of sexual or violent connotations online:

Well, Instagram is crazy. I've been on Instagram for a year and you can't imagine the number of messages like 'send me a nude and I'll send you a dick pic', and I tell them that I'm a lesbian and they say 'it doesn't matter, I've changed a lot of lesbians' (...) And it also happens a lot that accounts of girls, who I don't think are really girls, talk to me and send me three photos or so of tits and so on and then say 'now it's your turn'. And I'm like 'what, you don't think I know you're a man' [laughing]. So they create a female profile to send nudes and they think you will fall for it, well someone probably does, who knows... But what woman would send nude photos to someone she doesn't know? Well, it's a man. That's a man. I didn't expect that when I got on Instagram (Elena).

Although Elena in the quote above laughs it off, unsolicited contacts from men on social media can also be experienced as violence. For example, Daniela brings up examples of men asking whether she would like a "sugar daddy". In her narrative, another intersection is highlighted in terms of power relations, namely age and gender. Daniela, who is in her early 20s, has on several occasions been contacted by men in their 40s:

I feel very awkward when they write to me at 12 on a Saturday night because they've read a tweet, guys who are like 45 and with a public position. I tweet some shitty thing about love, like just today's reflection, you know? And they write to me at 12 at night and tell me 'wow, I love your tweets' and I'm thinking 'but do you understand the position you're in and the position I'm in?'. Like, you know that I won't respond in a bad manner because we're all in the same political bubble and I can't say anything... I think it's surreal (...) And I'm not going to expose anyone because I would just lose, but I'm just saying it's quite violent. It happens all the time with men, heterosexual men of course, well it could happen with bisexual men, but it's the typical profile of heterosexual men (Daniela).

These types of experiences do not fall under any criminal offences – unless they are repeated and continuous, in which case they may amount to stalking. However, Daniela explicitly relates to these types of unsolicited contacts by men who are substantially older than her as “quite violent”. The episode brought forward by Daniela is not explicitly violent – however, she experiences it as violent and awkward because of the power differences between them. That is, he might not approach her with the aim of exposing her to an awkward or violent experience, but his position of power protects him in a way that gives him certainty that he can approach her without any interest shown from her side, or any consequences. At the same time, the affordances of social media platforms provide the context for easily approaching and being approached. As Kelly (1987: 76) argues, discrete phases of “violence” and “non-violence” are not easily located; rather, there is “a continuous series of elements or events that pass into one another and which cannot be easily distinguished”.

In this regard, one of the unexpected findings that the fieldwork revealed was the broad spectrum of gender-based violence that the women brought up, when reflecting upon violence based on their own experiences. The events described above cannot be filed under the category “gender-based hate speech”. However, they are clearly related to their role as women online. In the cases of many of these women, their professional or activist roles render them more publicly visible and targeted not only as feminists but also as sexual objects (and often, in explicit hostile attacks, a combination of both). In the example of Elena, even

when she mentions that she is a lesbian and not interested in sexual interaction, the man contacting her on Instagram responds that “it doesn’t matter”, because what counts is his perception of her as a sexual object. Whilst the specific persons interviewed in this research may be more exposed to this type of harassment because of their public profiles and/or their big following on social media, this type of sexualised violence may be directed to anyone perceived as a woman in online spaces. In the words of Elena, the gamer:

It's just that it's gratuitous, I remember it was getting into a game and just receiving, receiving, receiving. The gamer world is a very violent world, very insulting. And I find that it's where they insult you the most. And social media is where they try to get you naked the most, you know? It's like sexual harassment and violence. In the gaming world, there is violence and on social media there is sexual harassment, you know? I divide it like this now that I have social media and I've seen what I've seen (Elena).

Finally, some of those interviewed also reflected on gender-based violence and sexual harassment they had suffered in a physical, face-to-face, space. For example, leaving the online context, Elena highlights an episode that took place when she worked in the advertisement industry and a man tried to drug and abuse her. Daniela mentions gender violence in a previous romantic relationship and explains how this also still affects her social media behaviour, for example in terms of not tagging her location in posts. Lucía, on the other hand, draws parallels to having been bullied in school and how violent online behaviour often resembles offline bullying. The references to these face-to-face experiences points to how, when reflecting on violent experiences linked to being a woman, online and offline experiences are highly connected in the narratives. Although online violence is often normalised and brushed off, when women tell their stories, this violence is still related to offline experiences of gender and sexual violence, as experiences that have happened to them because they are women.

An in-depth analysis of this broader type of violence goes beyond the scope of this thesis. However, the fact that the targets brought up also these kinds of experiences, tells us something about how, from the lived experiences of those targeted, all these experiences are perceived as related. By examining gender-

based violence within the dual frameworks of the violence continuum and gender-based hate crimes, we can observe how these incidents serve as both a reflection and reinforcement of patriarchal structures. They perpetuate the notion of "putting women in their place", defined them as either desirable or undesirable sexual objects, and dictate whether they are permitted to raise their voices publicly. That is, all is "part of the spectrum of harm to which women are exposed throughout their life-worlds" (Harris and Vitis, 2020). In the next section, I shall discuss how different intersecting factors, predominantly racialisation and sexual orientation, shape the experiences of online harm.

7.5. Intersectionality matters

As Lewis et al highlight (2019: 132), "the nature of intersectional identity meant that respondents sometimes reflected on their experiences of gendered abuse as a sort of hate speech but noted that their racialised identity or perceived sexuality was connoted in the language and terminology used, so the focus was not misogyny alone". In this sense, women may be subject to sexist hate speech, but in ways that at the same time draw upon other offensive tropes (Ibid.). It has also been highlighted by previous studies by entities such as Amnesty International how racialised women are disproportionately targeted by online abuse on Twitter.⁶⁸ For Black women, this has been referred to as "misogynoir", defined as "the anti-Black racist misogyny that Black women experience" (Bailey and Trudy, 2018). Digital misogynoir, then, is the often-violent dehumanisation of Black women on social media, as well as through other digital forms, such as algorithmic discrimination (Glitch, 2023).

Both the interviews and the digital ethnography conducted for this thesis point to sexual orientation, gender identity, racialisation and religion as factors that aggravate the abuse, as LGBT-phobic, racist, transphobic and islamophobic discourses are often mixed together with misogynist hate speech in the attacks against a person. Previously, I have drawn upon examples of abuse against

⁶⁸ "Toxic Twitter" (2018), available at <https://www.amnesty.org/en/latest/research/2018/03/online-violence-against-women-chapter-1-1/>, last accessed on April 18, 2024.

racialised women, trans women and Muslim women, taken from the digital ethnography, suggesting that the mere appearance of these profiles in mainstream channels or reaching mainstream audiences on social media often trigger attacks against them.

In this regard, the interviews allowed for deeper reflections on the intersection of misogyny with racism and/or LGBT-phobia in the abuse, which I shall discuss in the next paragraphs. The narratives of Naomi and Amara, as racialised activists, speak of how misogyny and racism are often expressed interlaced in the same attacks: “It started with questioning whether I could give that speech, and ended in some brutal, really brutal misogynist and racist attacks” (Amara). Amara perceives clearly how her identity as a racialised woman aggravates the abuse that she receives. This is similar to the experiences portrayed in Kishonna Gray’s (2012) study, which showed how the online violence that Black women gamers experienced was sexist until their racialised identities were unveiled and the violence abruptly turned racist *and* sexist. Similarly, as Sarah Jeong (2018) has shown, racialised women who changed their Twitter photos to that of a white man noticed a drop in hate speech they received. Any aggressive messages they received ceased to be racist and sexist and had a more benign tone (Ibid.).

In relation to the intersection of misogyny with LGBT-phobia, Elena speaks of how the insults expressed by men in the gaming world change when they realise that she is not only a woman, but a lesbian woman: “Once on Discord, I was playing on the server and they got super aggressive, well with insults of all kinds against my sexual orientation”. This was also one of the few attacks that she felt became so aggressive that she reported the individuals behind it both to Discord and to the police.

In both cases, the experience of being targeted by hate speech became exacerbated due to the intersection of several factors. Nevertheless, as *racialised women*, the attacks against Amara and Naomi almost always include a mix of racist and misogynist hate speech. That is, neither the framing of racist violence nor sexist violence alone is enough to understand the violence that racialised women experience (Rajani, 2022). For Elena on the other hand, although she is

openly lesbian, her sexual orientation is not visible in all her online interactions, as opposed to skin colour, which is a constantly visible trait in this current context of social media, when the link between a person's offline and online identity is continuously demonstrated through visual cues such as profile images. For example, in the online gaming world, her opponents are not always aware that she is a lesbian, whereas her identity as a woman becomes apparent whenever she speaks through her mic in a gaming session, or on social media through photos. This means that she also has experiences of being targeted "only" on the basis of her gender, and not always as a *lesbian* woman.

The experiences of Julia, on the other hand, allows for a reflection on intersectionality from another perspective. As an LGBT+ activist, her public appearances are always linked to LGBT+ topics, and she sometimes appear publicly to discuss LGBT+ issues with an activist colleague who is a cis man. She can directly compare the reactions that she receives and those that are directed towards him: "I'm thinking about an interview that we did for [a TV programme] and the comments were malicious, and I find curious the difference between how they attack him and how they attack me. Most of all, I think it's very much related to gender, what perspective they judge us from" (Julia). That is, she clearly links the specificity of the hate messages that target her to the intersection of LGBT-phobia and misogyny, as her male colleague receives less messages and not as aggressive as she does, although they both advocate for LGBT+ rights.

Finally, we must also consider intersectionality in terms of the real or perceived positions of the specific profiles of women included in this research – women with a public voice and profile and the complexity of power relations in this regard. The misogyny expressed in some online abuse may not be extended to all women, but is rather directed at feminist women, or women who speak out, or women who challenge patriarchal, heteronormative structures (e.g., Lewis et al, 2019). The targeting of certain women suggests that this specific type of hate is expressed selectively towards those women who step outside the expected norms, rather than to all women. This specificity is also perceived by some of the

targets interviewed, for example in the cases of Sara and Susana, who explicitly point out the particularity of the violence that targets feminists. Sara remarks that “I think it’s a very very specific kind of hate that targets women who have a public feminist discourse”. Thus, what intersectionality highlights is not classifications in themselves but rather what Catherine Mackinnon describes as the “white supremacy and male dominance, or white male dominance” (2013: 1023). The perceived threat to the power relations posed by LGBT+, racialised, or Muslim women might be perceived as even more aggravated when a layer of activism is added to their profiles. These women do not only exist outside of white heteropatriarchal norms, but they also raise their voices and claim rights for themselves and other persons. Linked to legal debates, this intersectional ground could also be expressed in terms of “ideology”, which is included in the open-ended list of aggravating circumstance in the Spanish criminal code.

7.6. Exploring the perpetrators through the eyes of the targets

A further critical aspect revealed during the interviews concerns the origin of the attacks, that is, the perpetrators. Although my focus throughout this research has been the lived experiences of those targeted by gendered online violence, the perpetrators, attackers, or “haters”, were far from absent in the stories of the women I interviewed. Rather, they were often at the very core of the experiences, as many of the attacks had originated from one or a few profiles with a big number of followers. Coordinated attacks are usually instigated by one or several accounts that have recognition and a voice within their respective community, in which tens, hundreds or thousands of users target the profile of an activist to harass, insult, threaten her or laugh at her. Both real users -either under pseudonym or with their real name - and bots participate in these actions (cf. Morena-Balaguer et al, 2021).

The women that I interviewed spoke of how these profiles had managed to create a large-scale attack by their followers. This had been obtained through the mention of a woman or citing a video or publication by a woman in their own Twitter or Instagram feed. In other cases, the targets perceived a high level of organisation behind the attacks, for example through online forums. That is,

targets perceive that attacks are often not as spontaneous as they may initially seem. This corroborates what has been shown by previous quantitative research, such as that by Villar-Aguilés (2021), signalling how attacks on feminism and feminists are neither isolated nor spontaneous. Rather, they are related to different currents of networked misogyny, as discussed in Chapter 1 of this thesis. In this section, I will look at the functioning and characteristics of at least a part of the perpetrators, based predominantly on the experiences and perceptions of the targets.

7.6.1. Being thrown to the lions

Harassment can have a public and performative aspect, where perpetrators collaborate with others to intensify their abuse (Lewis et al, 2019). Through the fieldwork, I have been able to observe how this is often a way of functioning of certain youtubers, particularly from the far-right or alt-right spectrum. This works as a double strategy: on the one hand, as a way of punishing those with other opinions, predominantly feminists and LGBT+ activists, and on the other hand, as a way of creating content. As pointed out by the collective Cuellilargo⁶⁹, these youtubers apply a few common strategies: They “blow up” a piece of content or video created by another individual, for instance, by producing a video or hosting a live stream where they provide commentary on the other person's video or content, either aggressively or by mocking them, their appearance, or their ideas. In this way, rather than creating content of their own, they parasite on other people's content. Regardless of the response by the person targeted, the youtuber gains something or “wins”. This is because if the target responds, for example with a video or a tweet, this response is turned into further content by the youtuber. And if the target ignores the attack, it is perceived that they either agree with the comments by the youtuber, or that they give up. This way, these profiles tend to “win” by a strategy of overwhelming, as they can create a lot of content against a target in a very short time, through parasiting on the target's original content. Finally, youtubers who originate this type of attack tend not to

⁶⁹ Video by Cuellilargo, “Monetizadores de odio” (September 27, 2022), available at <https://www.youtube.com/watch?v=rkIZBWHG1Qw>, last accessed on April 19, 2024.

take any responsibility for the waves of hate that they create, as they do not directly incite their followers to attack or spread hate. However, the followers know what to do and what is expected from them. Finally, an important aspect of this strategy is how these perpetrators come to monetise hate. These profiles live from making videos and uploading them to YouTube – or streaming videos on Twitch. They are thus deeply dependent on creating content and gaining followers. By leeching off the content of creators or communicators with contrasting opinions, rather than generating their own original content, they can swiftly churn out videos that resonate with their followers who share similar worldviews. By directing their followers to assail the original targets, they provoke a response that is then utilized to create additional content.

DE-14, who is a young feminist, is a clear example of having been targeted by this type of perpetrator. I first came across DE-14 due to one of her tweets where she publicly denounced the online abuse against her: “They’ve created porn pages with photos from my Instagram. They have threatened to rape me, they have impersonated me, they have harassed me, they have organised hate campaigns” (DE-14). I reached out to her to invite her to participate in an interview; however, I received a quick and polite response that she did not have time. This meant that I did not have the full story of what had happened to her. Then, whilst dedicating a part of the digital ethnography to understanding the modus operandi of some of the perpetrators of attacks against feminists, I spotted her in the list of videos of one of the youtubers who use the strategy described above. His videos are almost exclusively aimed at ridiculing feminists; among these, he has dedicated at least 9 videos to DE-14, where he comments on videos taken from her Instagram and TikTok accounts. Being targeted in this way means that she is likely to receive a constant stream of hate messages from the followers of this youtuber. There is clear evidence from this on her Twitter profile, where every post where she even touches upon the subject of feminism has an enormous number of comments⁷⁰, retweets and cited tweets. The

⁷⁰ When writing this chapter, at the beginning of 2023, the comments range between 1000 and 3000 comments, depending on the specific post.

comments are primarily from men who express contrasting opinion and criticise feminism in aggressive ways. On the post where she publicly denounces the attacks against her, she has directly turned off the comments.

The collective element of attacks – often at a massive scale – is also excruciatingly clear in the experiences and narratives of some of the targets that I did manage to interview, where they use expressions such as “he threw me to the lions”. These are the words of Lucía when she explains how a youtuber with a big following had managed to make his followers go after her, after publishing one of her videos on his profile together with her Instagram handle; that is, following the modus operandi described above. In the original video, she had mentioned his name without directly criticising him:

It wasn't a personal attack against this person. But it doesn't matter. It doesn't matter at all because someone saw it, they mentioned me and then he cited my tweet, and recorded the video and put it on his Instagram and mentioned me. So... He threw me to the lions (Lucía).

This way, this youtuber had managed to redirect his followers to her Instagram profile:

Then all of a sudden, I start to see messages coming to the video that I have posted... to my video. And DMs⁷¹, many DMs. So I start to investigate and I discover that it is this guy who has posted my video on his Instagram and that's already something very strange, it's already very violent because it's like your face and your content in a place where it doesn't belong (Lucía).

For the big profiles using this modus operandi it is just another piece of content. For a feminist with a public profile, it means, on the one hand, seeing their oftentimes personal content being used completely out of context – in an ambience where most of the users do not agree with their opinions. On the other hand, it entails being ridiculed and targeted by attacks that not only take a toll on their mental health, but that also take time to respond to. That is, it is an effective strategy for keeping women both on guard and occupied. Lucía describes how she received around 500 messages, and subsequently spent six to

⁷¹ Direct Messages, that is, messages that are not public and that only the receiver of the message can read.

seven hours blocking people, until she decided to put measures into place so that only her followers could send her messages. Even then, she says “people followed me to insult me, a few, but there were some committed to the project who did it” (Lucía).

The targets interviewed perceive that the big profiles who use this strategy against feminists know very well what they are doing. These profiles are aware of their reach and that their followers will listen to their subtle commands. Still, they can take their hands off it, as they have not directly incited anyone to attack the target. The impunity for these perpetrators is close to absolute:

Of course, they can keep doing it, in fact, they get rich by this and I remember that the youtuber, he asked in his tweet ‘should I make a video for her?’. That means taking my video and commenting on it in his own video. And I spent weeks going into his YouTube, suffering, because if he makes a video based on me, they will come back again, won’t they? I remember the fear, ‘it’s going to happen again’. So all of a sudden that happens, and you want to disappear (Lucía).

This way of functioning has important similarities with what Karla Mantilla (2015) refers to as “gendertrolling”. This type of trolling is distinct from other types of trolling that may more generally try to disrupt online interactions without attacking a specific group or identity factor. Gendertrolling is

exponentially more vicious, virulent, aggressive, threatening, pervasive, and enduring than generic trolling (...) Gendertrolls, as opposed to generic trolls, take their cause seriously, so they are therefore able to rally others who share in their convictions to take up the effort alongside them, resulting in a mob, or swarm, of gendertrolls who are devoted to targeting the designated person. Because of the numbers of people involved in the attacks, gendertrolls are able to sustain their attacks for an extended period of time—for months and, not atypically, even for years (Mantilla, 2015: 17).

Whilst “gendertrolling” as a concept could be applied also on other types of attacks, it is particularly relevant for the functioning of this type of anti-feminist profiles, as they tend to come back to certain feminists and harass them over an extended period of time. In this type of attacks, “women are typically called ‘cunts’, ‘sluts’, ‘whores’, and the like; their appearance is insulted by calling them

‘ugly’, ‘fat’, and much worse; and graphic pornographic depictions are frequently made of images of the targeted women” (Ibid.). The insults related to women’s bodies can be illustrated by the experience of DE-20, a feminist who is active mainly on Twitter. She denounces how an alt-right youtuber and Twitch profile discusses her body in a public live stream. Once she has denounced this publicly on Twitter, he then goes on to dedicate part of another live stream to talking about her Twitter post. Here, it is clear how her reaction only provides him with more content and how it is very difficult for her to go winning out of this situation. Linked to this, there are clear parallels with debates on surveillance. These anti-feminist profiles keep a steady watch on the content published by feminists, and then quickly react by making videos, whilst they subtly send their followers to attack those who are targeted. Surveillance can be defined as the coupling of information collection and the use of power (Andrejevic, 2015). In this case, the information collection consists of the constant watching over the content produced by certain feminists. The use of power is reflected in the combination of the gendered power relation, the power attributed by a big online following, and the power of the profile over what they publish on their own channel. To this we can add the lack of power of the person targeted, to get this content removed, or to stop these profiles from parasiting on their content. The latter aspect is largely linked to the lack of interference by the platforms to restrict this type of publications. The outcome of this strategy is a sensation of constant surveillance for those targeted, who are thus forced to be constantly on guard, in addition to dedicating time to pondering on and setting up protection or response strategies. In the long run, this may lead to them avoiding feminist discussions in public spaces, which I shall discuss further in Chapter 8. That is, male surveillance or techniques of watching may effectively contribute to shaping the type of feminist activism emerging from social media platforms (Megarry, 2017).

Another group of perpetrators frequently mentioned in the ethnographic material is the anti-trans community. One does not need to observe the Spanish Twittersphere for very long, to notice the steadily increasing tension between what can be denominated transexclusionary/transphobic persons on the one

hand, and transinclusive feminists/trans- or non-binary persons on the other hand. Sergio Arce-García and Maria Isabel Menéndez Menéndez (2023) have analysed attacks on Twitter against (and between) feminists and trans persons from a quantitative approach, highlighting that

there is a strong polarisation that provokes attacks between trans and feminist groups, in which groups of a very suspicious nature can be seen (use of bot accounts, suspended or deleted accounts, relations with outside groups and not inconsiderable origins from non-Spanish-speaking countries), with methodologies seen in other organised campaigns. The use of these techniques would come to define an environment in which the aim is to generate feelings of hatred that last over time, since repetitive discourse and the provocation of strong basic emotions generate the main basis of social hatred (Ibid.:13-14).

Also Diana Morena-Balaguer et al (2021) have identified trans-exclusive anonymous women as common perpetrators of attacks against trans-inclusive feminists. These profiles consciously hide their identity, often under pseudonyms and iconography that refer to first wave feminism. This is the group that attacks the accounts of feminist activists with the greatest virulence, through grammatical constructions that deny the identity of trans people and that brand trans-inclusive women as “traitors” and “mercenaries” (Ibid.). There are also synergies between the trans-exclusionary discourse and the discourse of the extreme right, both in terms of vocabulary choices and in the ideology behind the accounts, given that many of these profiles also send racist messages to racialised women (Ibid.)

Trans exclusionary “feminists” as perpetrators have been brought up both in the interviews and in the cases collected in the digital ethnography. Notably, also this group of perpetrators demonstrates a proclivity for orchestrating massive attacks through their followers. In the following vignette I draw upon the experience of one of the targets interviewed and then contrast this with data collected during the digital ethnography. This relates to the second wave of attacks against Lucía, which was originated by a person that she refers to as

very much against trans people and against inclusive language (...) So they asked me [in a video made by an online TV channel] what I thought

about this person. Well, there I had already learned not to write the person's name but of course in the video I said it (...) I said that it's a person who uses the methods of a bully, who is dedicated to attacking, to harassing trans people, using violence against trans people (Lucía).

Lucía recounts that she had criticised this person's behaviour in a calm manner. However, this profile did not receive it well: "So then this person got all their followers together and they were insulting me for many days, many, also going back to things from my professional past to attack me" (Lucía). The experiences of Lucía put forward an interesting point, namely that regardless of the territorial belonging or political stance, the same methods of attack are often used. In Lucía's case, one of the original perpetrators was what can be described as an anti-feminist neofascist, whilst the other one was a catalanist anti-trans activist, who self-identifies as a feminist. However, the same *modus operandi* was used, that is, they both mentioned or cited the content published by Lucía, and subtly got their followers to go after her, "throwing her to the lions":

'This person, I'm going to give her to you', so that my troupes, my soldiers, because they are soldiers like with adoration of their leader, they call them by their name like they were gods almost. So they put you there so that the dogs come and destroy you (Lucía).

One difference between these attacks and those by anti-feminist youtubers, however, is the monetisation. Whilst the youtuber gains direct economical profit from creating this kind of content through commercial spots ingrained in their videos; the anti-trans activist mainly gains visibility for their ideas and a sense of community with those who have similar ideas. Both, however, aim to put certain kinds of women in their place and silence them, through the use of these methods.

I have identified similar experiences in the digital ethnography, mainly targeting trans- or non-binary persons. DE-5 and DE-11 have both suffered attacks from the same transphobic person who defines herself as a feminist. In one of the attacks, she publishes photos of DE-5 and DE-11 as well as two other trans persons, with the accompanying text

[Names of other transphobic profiles as well as her own name] all agree that these persons who were born with penises and testicles are men. We also agree that we shouldn't be fined for saying so. And almost on nothing else. If you also see four men, re-tweet.

Here, firstly, she calls upon other profiles that she already knows will agree with her and who therefore most likely will amplify the message. Secondly, she directly encourages her followers to retweet if they agree with the message. That is, she uses two different strategies to amplify this transphobic message that directly targets four individual transpersons, in addition to the wider collective. It is also significant that she mentions that the transphobia is almost the only issue that the other persons mentioned in the tweet and herself agree upon. This may say something about how currently in our society, transphobia cuts through different political standpoints and demographics.

7.6.2. The troops of hate

Linked to the modus operandi described above is also the broader organisation of certain attacks. This type of violence is not random or casual but obeys some specific patterns that have increased over the last few years, especially on social media platforms such as Twitter (Piñeiro-Otero and Martínez-Rolán, 2021). Many attacks do not stem from only one original perpetrator or inciter, but are organised collectively through different forums:

There were many series of users who came organized, which also seems important to me because for example a local branch of Vox, I don't remember which one, called on their followers, they put a link to my profile to send all these men to my Instagram. And that did seem like a level of planning to me. In the same way that friends of mine told me that in ForoCoches they were also organizing so that they would come to my profile. That's to say, it was not as spontaneous as it might seem (Sara).

Whilst in the Anglo-Saxon context spaces such as Reddit and 4chan have been signaled as cultivating anti-feminist sentiment (e.g., Massanari, 2015; Krendel, 2022), in the Spanish context, ForoCoches is the forum that is most cited as harbouring the Spanish manosphere (e.g., Diaz-Fernández and García-Mingo, 2022; Caldevilla-Domínguez et al, 2022). ForoCoches follows a grid structure similar to Reddit and presents a central forum as the epicenter and a network of

subtopics or subforums from which thematic threads follow (Caldevilla-Domínguez et al, 2022). In comparison to social media, in a forum, the information about the user is very scarce, with a hermeticism that lends itself to protecting the profiles. This inherent secrecy, coupled with a sense of group affiliation, the development of distinctive jargon, and the employment of specific keywords or euphemisms, collectively serve to fortify the sentiments of belonging, identity, and reference within the forum community (Ibid.). As David Caldevilla-Domínguez et al argue, the morphology of Forocoches as a manosphere - a closed community since its membership and entry is by invitation, and pseudo-anonymous - has provided a space where cyber-macho communities with a clear tendency of chaos, conflict and violence, have been able to grow. In addition to ForoCoches, previous research has also pointed out closed Facebook groups and the web Buburja.info as spaces where a Spanish manosphere is fostered (Calala, 2020). The use of these forums signals a high potential for collective action and the possible role of these spaces in the radicalisation of men surrounding topics such as the denial of gender-based violence or anti-feminism (Ibid.).

Nevertheless, as evidenced both by the interviews and the cases collected in the digital ethnography, much of the abuse stems from profiles with a name and surname. That is, anonymity or pseudo-anonymity granted by a forum is no longer understood as an absolute requisite to spew gendered hate online. As expressed by Lucía when talking about the profiles behind the attacks against her, “there were lots with a name and surname, they’re not ashamed (...). They follow you to hate you. People have the need to let you know that they don’t like you. It’s their way of trying to make you not exist. And it’s the only way they have to accomplish it” (Lucía). Likewise, based on the large-scale attack that targeted her, Sara describes a mix of profiles among those that attacked her: “There was everything, there were profiles with first and last names where you could enter their profile and see their group of friends, their houses, some even had children. There were a lot of hidden, or unrecognizable profiles, where you couldn’t see the images” (Sara).

Reflecting on other factors that characterised the profiles of the attackers, Sara points out how young many of them were: “And then I was surprised that there were a lot of very young people. Many very young children, even minors (...). It’s a bit this user profile of 13-, 14-, 15-year-old anti-feminists...” (Sara). The overwhelming majority of the profiles behind the attacks analysed through the interviews and the digital ethnography were men. The main exception in this regard, seems to be the attacks stemming from transphobic communities, where more women are present, particularly those who refer to themselves as “radical feminists”, and who are often denominated “TERFs” - Trans Exclusionary Radical Feminists⁷² by those targeted. In terms of gender, as Lewis et al (2017: 1572) argue,

online abusers may adopt the discourse of misogyny regardless of their gender. This is not ‘male violence’ so much as ‘masculinized violence’; i.e., violence that is generally perpetrated by men against women and girls, but may be perpetrated by women, and which draws on and generates misogynistic discourses.

That is, an intersectional lens also sheds light on the identities of the perpetrators, and how these are not always necessarily men, but that they may also be women who express masculinised, misogynist, racist or transphobic violence.

7.7. “Real” violence vs “not real” violence

The broad span of experiences described by the targets above reinforces the argument put forward by other scholars (e.g., Awan and Zempi, 2015; Lewis et al, 2019) that online and offline spaces must be considered as a continuum rather than distinct domains. Women do not clearly distinguish their online victimisation from that in the offline world; rather, these are located on a continuum of violence. Further, there are cases where the targets share also physical spaces with persons who harass them online. This is the case of Susana, who received some disturbing comments on Twitter from one of her university classmates; a

⁷² A perhaps more suitable term is “trans women exclusionists” as the values behind their transphobia are anti-feminist and misogynistic and therefore should not be associated with feminism (Rajani, 2022).

young man who nevertheless never dared to speak to her in class. Violence and harassment may also be transferred from the online to the offline, as in the experience of DE-18, who recounts how insults were painted outside her workspace after a long period of online harassment by several perpetrators. These experiences point to the blurred boundaries of the online and the offline, or rather, they show that these boundaries do not exist or that they are at least not as firm as they are often perceived to be (Linabary and Batti, 2019).

At the same time, dominant conceptions of physical violence as “real” violence render invisible a wide range of forms of non-physical violence (Powell and Henry, 2017: 50). This is in part linked to the still dominant perception of digital spaces as somewhat intangible and separate from everyday life, as lived in the physical (non-digital) world. In many of the narratives, as I shall discuss further in Chapter 8, the targets challenge the construction of the internet as a space that is not the real world (cf. Linabary and Batti, 264), through pointing out the *embodied* impacts of the digital violence. Nevertheless, this perception entails a kind of “glitch” or gap, when it comes to how *online violence* is perceived. Whilst recognising that the attacks, the hate speech, and the harassment they suffer online is lived and felt through their bodies, and that it has consequences beyond the online sphere; online experiences are still generally not referred to by the targets as “violence”. That is, in their narratives, there is a hesitance to conceptualising the violent actions that take place online as “violence”. Whilst some of the women interviewed even refer to some online episodes as “violent” (see e.g., Daniela earlier in this chapter), at the same time, they quickly brush them off, as they have come to be a part of so many women’s online lives – and thus highly normalised, as I shall discuss further in the next section,

Among some of the targets there is a continuously present concern that the online abuse could reach their offline – physical – sphere. Through these narratives, we can observe how violent actions that take place offline – instances of physical violence – are perceived as “real” violence. This is also shown in the perceptions of risk: the risk of being targeted by “real” violence – understood as physical violence – does indeed cause fear and often leads to further precautions. For

example, Daniela explains how she makes sure that none of her publications has any tags of geolocation, and that none of her photos contain elements that can show where she lives, to avoid being targeted by violent actions outside of digital spaces.⁷³

Nevertheless, the clearest illustration of the perception of violence comes across when discussing the possibility of reporting online attacks to the police. This is shown, for example, in the story of Daniela. She says that she has not reported online threats to the police as she believes that the perpetrators do not really know who she is or where she lives, so she feels that the threats she receives are “empty threats” that could not hurt her physically. This is similar to the experience of Amara, who wasn’t much concerned about the threats directed towards her until she went into the Twitter profile of the person who had threatened her and realised that he lived in the same region: “In that moment I perceived the danger as much more real and much closer and that if this person wants to, he can find me. And then I had a panic attack... (Amara). This was also, as we shall see in Chapter 9, the point when she decided to report the incident to the police; that is, when she felt that there was a possible threat to her physical safety. However, the threats, in and of themselves, are not substantially perceived as acts of violence capable of causing diverse forms of harm. Regardless of the potential realisation of these threats, these can be categorised as criminal offenses according to the provisions outlined in the Spanish criminal code. Nevertheless, it is important to consider that this tendency not to report threats to the police could also, as I shall discuss further in Chapter 9, be related to the women’s perceptions of and previous experiences with the police, and of how the police is perceived to address online violence.

Even if the acts of violence do not reach the threshold of being criminal, also non-criminal hate incidents may entail a highly negative impact on both victims and society, with victims relating that “by their very frequency and ubiquity, some of the most minor types of victimisation – name calling, verbal harassment, and so

⁷³ These measures of digital self-protection are further discussed in chapter 9.

on – had the most damaging effects” (Perry, 2009: 402). Acknowledging everyday acts of hostility is crucial, as they indicate the presence of a spectrum of violence and a spectrum of hate incidents, some criminalised and some not. Both contribute to the marginalisation and subjugation of specific groups, as has also been highlighted by Jennifer Scheppe and Barbara Perry (2022).

7.8. The normalisation of online violence

As hinted at by the lived experiences explored in this chapter, individuals subjected to online gender-based hate speech often normalise digital violence. The accounts of those targeted illustrate how the perception of online violence as not entirely "real" contributes to accepting these experiences as part of everyday online existence. Attacks are viewed as something one must endure, particularly for women expressing opinions in the public sphere or extending beyond circles of like-minded individuals, or for visibly racialised, transgender, or Muslim women. Particularly for the latter groups, normalisation may be linked to symbolic violence and the very internalisation of discriminatory structures. This has been highlighted particularly in terms of discrimination, hate crimes and incidents towards LGBT+ persons (see e.g., Chakraborti and Hardy, 2015; Paterson et al, 2018).

Normalisation is more salient in the narratives of the younger targets interviewed the younger feminists followed through the digital ethnography. Young people often convey a sense of “living within a matrix of sexism: that is, an environment in which sexism, misogyny, and elements of rape culture merge as a normalized backdrop to everyday life” (Sills et al, 2016: 6). In the interviews conducted for this thesis, the younger women talk about being targeted by different violent actions - somewhere on the continuum of violence - as part of everyday life on the internet:

There are many comments that I come across in my day to day, that I justify for ignorance, and I don't react like 'hey, this is a hate message, delete it'. And maybe it's something that doesn't just happen to me but to more persons within the community, and maybe I can't think of more cases because it's like you just let it go... (Julia).

In this case, the person interviewed found it hard to single out any specific incidents, as they take place continuously and thus blend together. The interview process became a point of reflection for her in this regard, about the fact that she and many with her have normalised experiences of online violence, as part of the game of being political online -or offline- as a woman, and LGBT+ person. Similarly, Elena points out, in relation to the online gaming world that “it’s part of it that they harass you, it’s part of it unfortunately. I know what’s in store for me” (Elena). In the same way, when asked about her experiences of online violence, Lucía talks about three big waves of attacks; and mentions only in passing that she has in fact been receiving insults and smaller attacks throughout her time on Twitter. Given that some women receive voluminous abusive messages, detailing threats of physical and sexual violence, a “simple” abusive message may, in comparison, be experienced as relatively mild (e.g., Lewis, 2016).

In this context, normalisation can be part of a coping strategy, linked to the routine and everyday nature of abuse. It would be related to what Lewis et al (2019) refer to as a “reluctant normalisation” (Ibid: 130) by which women speak about their experiences of online abuse as insignificant. There is also a parallel with the normalisation thesis of offline gender violence, which argues that survivors tend to downplay their experiences because they are part of the ordinary and mundane (Lewis et al, 2017).

Linked to this, another coping strategy that reflects the “normalisation discourse” is for targets to downgrade their own experiences in comparison with the experiences of other women or to minimise the impact of the abuse (Ibid.) This was salient in the ethnography, where targets frequently mentioned other feminists who had received large-scale attacks, comparing the attacks and downplaying the seriousness of their own experiences: “It wasn’t like the case of [name of a feminist] where they wrote her home address for example, that didn’t happen to me. I mean, I didn’t fear for my life” (Sara). That is, even targets who have suffered massive attacks tend to somehow downplay these, as long as the attacks do not include threats to their life that they perceive as realistic. In this regard, most seem to perceive a direct threat as a limit to what can be

understood as “normal” and, thus the limit to what they are expected to just deal with.⁷⁴

Whilst “normalisation” may be an effective strategy for coping with online abuse it also raises important concerns about the longer-term, insidious harm of considering death and rape threats as “normal” (cf. Lewis et al, 2017). In Chapter 8, I will discuss the impacts and consequences of suffering from this type of violence as a part of everyday life as a woman in the public debate.

7.9. Conclusion: Unveiling a continuum of gendered violence across and beyond digital spaces

Acknowledging the perspectives of the persons who are targeted by gendered violence in digital spaces contributes to asserting their agency. Through the in-depth interviews, I have been able to unravel the lived experiences of being targeted by online violence, reflecting on this together with the women interviewed. Their feminist perspectives have often been crucial to their own interpretations of and reflections on what they have experienced.

The concept of a continuum is fundamental for understanding the lived experiences of being targeted by gendered violence. In the first instance, conceptualising the online and the offline as a continuum facilitates an understanding of how different life spheres are affected by violence expressed online and how this violence cannot only be analysed as pertaining to digital spaces. As I have exposed throughout this chapter, women are being attacked online not just because they express their opinions in *online* forums and platforms, but for voicing opinions in any public media that reaches beyond a safe circle of like-minded. Violent discourse against these women is expressed online, because digital spaces offer easily accessible paths for opponents to reach the targets and express their opinions -or disseminate hate against their perceived opponents- freely and with impunity. However, the attacks against feminists do not necessarily target the opinions that they have expressed online.

⁷⁴ I shall further discuss this in Chapter 9, addressing responses and encounters with the justice system, as what tends to make targets feel that an incident could be serious enough to report to the police is a direct (death) threat against them.

As I have shown, some of the trigger events put forward have been appearances on mainstream media, such as TV shows. For some targets, such as racialised or trans women, they do not even need to raise any specific opinions, but their mere presence on mainstream channels gives rise to attacks against them. Online attacks, then, do not only impact on women's willingness to voice opinions on *social media*, but may have an impact on women's participation in any public debate or space that reaches a mainstream audience outside of their immediate circle of persons with similar views. Further, in some cases, violence is also expressed in offline spaces, and in the narratives of the targets, violence expressed online is closely linked to that expressed in offline spaces, through the common structural roots.

Second, and linked to the arguments above, the experiences narrated by the targets speak of a continuum of gendered violence (Kelly, 1987). When discussing their experiences, explicit verbal violence that targets them as (racialised/lesbian/trans) feminist women is not clearly separated from, for example, unsolicited sexualised contacts. From this perspective, gendered or sexual violence, whether in the shape of gender-based hate speech or a physical transgression can be interpreted as both a manifestation of and a tool for gendered power relations. Through acknowledging a broad context of experiences of gender-based harms we can better spot the individual impacts, cumulative effects, and collective implications of these harms (Powell and Henry 2017), as well as the structural roots and systemic nature of gendered abuse.

Third, the content and discourses of the very attacks are spread along a continuum ranging from more subtle remarks linked to traditional gender roles, to explicit rape or death threats. Nevertheless, throughout this continuum, the function of the content is the same: to put women in their place, to keep them on guard, to derail debates and keep feminists occupied with self-protection measures and responses to attacks. Further, all these experiences point towards the need to understand this type of violence as a process rather than a one-off time-constrained event (Perry, 2001).

The experiences of the targets are also a powerful reminder of how technologies and the use of these are subject to gendered structures. The currently prevailing model of "radical transparency" (Kirkpatrick, 2011) on social media platforms promotes the idea of maintaining a single identity online, leading to a greater merging of our online and offline existences, consequently emphasising the significance of our physical bodies within our online personas. This underscores the centrality of our gender identity and other identity traits in shaping our online interactions, a perspective that is embodied through both a gendered and racialised lens. At the same time, violence in digital spaces is highly normalised and tends not to be perceived even by those targeted as "real violence". That is, there is also a clear hierarchy of violence, where physical violence is placed at the highest level, and where a multitude of other expressions of gendered violence and particularly digital violence are perceived as something that women, and particularly racialised and LGBT+ women, just have to deal with online.

8. Digital harm, real impact: Unmasking the impact of gender-based digital harm

...the Internet is a real space, it's not virtual, but real, and the emotions it generates are real, you live them through your body, so the anguish and sadness were very real (Lucía)

8.1. Introduction

Although feminists recognise online abuse to some degree as a facet of online existence, it undeniably affects the lives of those directly targeted. This is particularly relevant in cases of massive attacks from large mobs, or explicit death threats. Nevertheless, it is not solely the most extreme forms of online hate speech or violence that have a direct impact. Also those messages that do not contain explicit threats or incitements to violence, but more subtle verbal abuse, can have a profound impact on targets, particularly when repeated by numerous individuals in a public setting. The effects of online abuse on direct targets can be understood as multi-dimensional (cf. Lewis et al, 2019), with ramifications across different life spheres. The experiences of women targeted by online violence reveal a broad spectrum of consequences, transcending the boundaries of the digital world to affect their emotional, physical, and psychological well-being.

As argued in the conceptual framework exposed in Chapter 2 of this thesis, our bodies are not disconnected from what we experience online. The body is the medium that mediates and modulates relationships between the self and the world, also in digital spaces, meaning that technology-facilitated violence leads to real harm to real victims. Thus, this harm must be analysed as embodied harm (Henry and Powell, 2015). Whilst many impacts are salient, also in more conventional forms of sexual and gender-based violence, we must recognise the unique harm caused when technologies are used as a tool (Powell and Henry, 2016). Many of those targeted by gendered online violence are highly aware of these impacts and are eager to talk about how this violence has affected them.

Simultaneously, hate crimes and hate speech are often termed "message crimes" (e.g., Chakraborti, 2012), carrying significant collective repercussions. One of their objectives is to instil fear and caution within the broader community, undermining their dignity and sense of belonging. While the direct victims endure intense physical and emotional repercussions, they also serve as symbols. They represent the recipients of the message, which is intended for the wider community, society as a whole, and those who share similar beliefs as the perpetrators.

In this chapter, I will first discuss the embodied impacts of gendered online violence on those specifically targeted by it⁷⁵, after which I will analyse the further consequences of this type of abuse both at an individual and a collective level. This includes an analysis of self-censoring and internalisation of the messages, as well as the possible "chilling effect" of this violence.

8.2. Embodied experiences of online harm

Findings from the interviews conducted reveal a broad range of impacts, affecting not only the emotional and physical well-being of those who are targeted by online violence, but also their sense of security and, because of this, their behaviour online. The frequency of the abuse is related to its impact, showing that, as Lewis et al (2019) argue, frequency exacerbates significance. However, it is not only the continuity and regularity of abuse, but also the scale of

⁷⁵ The targets and the various experiences of being targeted by gendered online violence have been discussed in Chapter 7 of this thesis.

the attacks, which create a major impact. In this regard, Lucía mentions that she has received one-off instances of hate from individuals over the years. However, it was not until she was targeted by three big waves of hate over a short period of time that it really took a toll both on her mental health and on her willingness to express herself online. Similarly, Amara had been targeted by insults and smaller attacks for several years, but recounts that she was not impacted to any higher extent until she received a death threat and felt that she was in danger. When she started believing that the person who had threatened her might know where she lives and works, she experienced an anxiety attack: “I had to take a sick leave, I didn’t dare to leave the house” (Amara).

Most women interviewed made a point of explaining the consequences of the online attacks both on their physical and on their mental health. Some explicitly stressed the importance of understanding the continuum of experiences in an embodied sense:

I think that there is a moment when it seems that this duality exists where one world is disconnected from the other, but that’s false. Obviously, the violence that reaches you digitally can end up going through you at a non-virtual level. And in fact, one of the things that I could see more clearly was that apart from fear for my life, the fact that I was sad in a real sense and not in a digital one, the fact that when I turned off the computer, I still had this pain in my chest, this feeling of abuse, is proof that this division is not so clear. In the end the body is you. And the truth is that at that time I didn’t fear for my life (...) But I did feel physically the massive contempt that I had received on a national scale (Sara).

She goes on to compare the feeling of digital abuse to the feeling of being physically beaten up:

I remember that in the weeks afterward I told my friends that I felt as if I had been beaten up, it was that feeling. I couldn't sleep, I was very tired, I was afraid to open the computer. That is to say, my belly hurt a lot, like a total lump in my chest, so I had that feeling, even though in principle I knew that these people were not outside my house (Sara).

This metaphorical and embodied comparison with being beaten up, has parallels with the description by Charles R. Lawrence III (1993) of racist speech as “a slap in the face”. Similarly, Elena shares how, particularly when she was younger, the

online violence that she suffered when she was involved in online gaming created a “knot in her stomach”. In these cases, the women targeted by these discourses report that they felt a somatic impact, an embodied harm.

Further, the frequency of being targeted by harmful speech on a daily or weekly basis creates a cumulative effect, which is noted by targets both in the gaming world and beyond:

Now that I'm older, I know what's out there, you know? But when I was younger, I felt it in my stomach, and I thought 'wow, they're saying these things to me.' (...) Many times, I would be shocked, and I didn't know what to answer and a friend of mine would come out and call them sons of bitches or something, you know? When I was younger, like a teenager, it left a knot in my stomach, and it stayed. Then it went away in half an hour or so, you played two more games, and they insulted you in a different way, you know, a heavy heavy insult. Then it went away... But, damn, these are things that we accumulate, you accumulate throughout your life (Elena).

As exposed in Chapter 7, there is often a continuity in the abuse or attacks, as these are rarely a one-off event. The effects, thus, are accumulative and exacerbated by the frequency of attacks. Thus, more frequent abuse tends to increase the impact of online violence (cf. Lewis et al, 2017).

It is also relevant to consider what Lucía emphasises as the *real feeling* of harm. By using this expression, she places the focus on the emotional part and how the attacks –coming at a point when she also was enduring other difficult circumstances in her life- have taken a toll on her mental health:

It affected me emotionally at the time, because I had anxiety, I was crying. That is, the Internet is a real space, it's not virtual, but real and the emotions it generates are real, you live them through your body, so the anguish and sadness were very real. From there, and it surely has to do with other circumstances in my life, but I have gone through a process, a depressive episode all this year and I know it had to do with other things, but... In the end it's not so much the impact it has on the way you use social media, but the impact it also has on your life (Lucía).

As revealed by these experiences, being targeted by online verbal violence and other types of attacks has a range of impacts at the individual level. To

accurately comprehend these impacts, we need to understand that these do not stay online but are embodied in the sense that they are experienced through the body that accompanies us through all spheres of life. As Butler (1997:4) highlights, the lack of proper words to describe linguistic injury makes it more difficult to identify the *specificity* of linguistic vulnerability as compared to physical vulnerability. The connection between physical and linguistic vulnerability is often metaphorical, as illustrated by the feeling of Sara of being “beaten up”. At the same time, the fact that physical metaphors are used to describe linguistic injury suggests that this somatic dimension is crucial to understanding linguistic pain: “Certain words or certain forms of address not only operate as threats to one's physical well-being, but there is a strong sense in which the body is alternately sustained and threatened through modes of address” (Ibid.: 5).

Online abuse may not leave visible physical marks, but it creates embodied harms in those targeted. It gives rise to immediate physical, psychological, and psychosomatic reactions such as stomach aches and anxiety attacks; a general feeling of having been beaten up. Nevertheless, there are also more subtle psychological and emotional consequences related to the cumulative effects of being under attack, which can also be understood as forms of linguistic injury, as I shall discuss further in the next sections.

8.3. Being on guard

As pointed out in the introduction of this chapter, not all impacts of gender-based online hate speech are directly manifested in the body through physical symptoms. When the attacks have calmed, and when the immediate stomach-ache and anxiety have faded, more subtle emotional impacts often persist. One such impact can be described as a sense of remaining *on guard*. The continuum of sexual and gender-based violence and the fear of being targeted, results in women and gender non-conforming people being "continually on guard" because of the possible threat of male violence (Stanko, 1985).

As Sara explained, she had been targeted by a large mob and during the days following the attacks she disconnected from social media: “In that moment, for two or three days, or four, I went off social media. I couldn’t, I didn’t want to be there. Some friends took charge of being a filter and they cleaned all the hateful content from Twitter and Instagram” (Sara). She eventually returned, but stresses that the impact that this event had on her was extensive: “I think much more about everything that I’m going to say, everything I’m going to do, and... But I’m telling you, I’d really like to know to what extent this episode affected me, it’s something I’ve thought about many times” (Sara). Several times during the interview she returns to the fact that this episode changed her somehow. Evidently, being targeted by hate speech not only has an immediate impact, but it also potentially has a long-term impact on those targeted.

In addition to the frequency of actual attacks taking place, there is also the fear of new attacks, meaning that those who have once been targeted can never fully let their guard down, as it is likely that they will be targeted again. Lucía describes how after the waves of attacks against her, her fear was triggered by a negative article about her:

I remember when this woman wrote the article and I started crying loads like ‘no, not again, not again, please leave me alone’ (...) And then, on top of that, you go in [the social media channels] and you see like traces of people talking about you, also just the fact that they are talking about you (Lucía).

We can discern that it is not only the attacks or the anticipation of future attacks that evoke fear, but also the mere realisation that people are discussing you on social media can trigger suspicion and discomfort in individuals previously targeted by online harassment. Daniela recounts how she is in a state of constant tension, feeling controlled, and how this feeling transcends the online:

I’m very cautious about these things and I know exactly what’s out there, and I think part of it is being in a state of constant tension, because of receiving messages, even from specific people on Twitter that you’ve spoken to a couple of times, but you don’t know. They may tell me: ‘I’ve seen you downtown’. I don’t like it one bit when they say that. It makes me feel super violated and it makes me feel super controlled, you know. And

partly because they're not my friends, I don't know them that well, so it's a bad feeling. It makes me live in constant tension. For example, I always remove my location when I post a tweet or something on Instagram. It's like my paranoia, but I don't like people knowing [where I am] (Daniela).

One of the impacts in this regard is a kind of emotional harm, referred to in the targets' narratives as "taking up of mental space" (Lucía), "being on guard" (Naomi), or a "constant tension" (Daniela), which can be understood as adjacent to a feeling of being under constant surveillance (e.g., Megarry, 2017). This sense of being on guard and under surveillance does not only impact on women's mental health, but it also has an impact on their behaviour, across the continuum of the online and the offline, as I shall discuss further under the heading "Self-censoring". All these experiences point to the importance of understanding that what happens online does not stay online in a virtual realm and that it does not only impact behaviour and relations in digital spaces.

This brings to the fore another set of important consequences of being targeted by attacks, which is the emergence of insecurity and different types of changes in behaviour. These are related to the constant feeling of being monitored, watched, and questioned.

8.4. Internalisation and insecurity

An impact of online hate speech that has been largely overlooked in previous studies on gendered online abuse, is its effect on the very subjectivity of some targets. There are parallels between Mari Matsuda's (1993) analysis of hate speech - where she discusses victims' perceptions of the effects of racist hate messages - and some of the harms highlighted in my research. As Matsuda points out in relation to racist hate speech:

the effect on one's self-esteem and sense of personal security is devastating. To be hated, despised, and alone is the ultimate fear of all human beings. However irrational racist speech may be, it hits right at the emotional place where we feel the most pain" (1993: 25).

Although Matsuda (Ibid: 23) stresses that racist hate speech needs a separate analysis from gender-based hate speech, violent pornography etcetera, "because of the complex and violent nature of gender subordination and the different way

in which sex operates as a locus of oppression”, the harms from the point of view of the victim are similar. Further, importantly, sexist and racist verbal abuse is often mixed in attacks against racialised women.

These internalised impacts can be linked to the concept of linguistic injury, exacerbated through the mere repetitiveness of certain words or messages:

At some level, no matter how much both victims and well-meaning dominant-group members resist it, racial inferiority is planted in our minds as an idea that may hold some truth. The idea is improbable and abhorrent, but because it is presented repeatedly, it is there before us (Matsuda, 1993: 25).

Hate speech is always repeatable speech, speech that can and must be repeated or re-incited to reach its aims (Butler, 1997). Some messages, through mere repetition, may become internalised, both among targets and bystanders, who read the words time and time again. This means that certain terms and words may pop up in our minds against our will. Matsuda describes how, after having read an unhealthy number of racist statements for the purpose of her research, the first word that came to her mind when seeing an Indian woman was a racist slur. Only after putting that aside in her mind, could she move on to her own thoughts. Hate speech in this sense permeates into our minds with a kind of toxicity. This way, hate speech shapes subjectivity and thus experience (Butler, 1997). Those targeted by hate speech may at a theoretical level know that the messages they receive are intended to ridicule, dehumanise, and scare them off from certain spaces and debates, but the messages, through their repetition and sometimes massive level may still seep into their minds as some kind of truths. Nevertheless, it is also relevant to highlight that in Butler’s view, the very repetition of hate speech seems to serve a dual purpose: it reinforces its capacity to cause harm while, paradoxically, retaining the potential to be reclaimed by the very individuals it seeks to subordinate (such as in the case of words like “queer” or “feminazi”, as discussed in Chapter 7). This makes repetition significant both for doing and undoing hate speech (Eichhorn, 2001).

To be injured by speech frequently entails a loss of context - to not know where you are, to be put out of control: “one can be ‘put in one's place’ by such speech,

but such a place may be no place” (Butler, 1997: 4). That is, harmful speech may cause a sense of disorientation in those who are targeted. In this process, generating insecurity amongst targets is a critical function of hate messages: Sara suffered a massive attack that targeted a part of her physical appearance, and explains that “at the level of self-perception of my physical appearance (...) I did deny for a long time that it had any kind of impact, but I do think it had an impact. And I see it in like very absurd things” (Sara). She also explains how she now covers certain parts of her body, even at home, and that she is still calibrating to what extent that episode affected her: “I don’t have the answer, but I do have the intuition that it made me more insecure and that it generated a new insecurity” (Sara). She knows that what the messages said is not true, but still has internalised the content of the messages, to a certain point.

Lucía, on the other hand, describes how the content of the attacks often revolve around her professional profile as a journalist, aiming to create self-doubt in her regarding a very important part of her – her professional career:

So to hurt me, they say that I’m not a communicator. Which is like trying to deny my status as a person. I’m both a person and a communicator, you may not like me, that’s ok, I don’t go and tell people that I don’t like them, but ok, the feeling of dislike is licit, but you want to take that away from me, you want to deny that I’m a communicator? (...) They want to make you doubt that, right? To doubt the most primary thing. (Lucía)

These experiences evidence not only the more extreme impacts, but also the more subtle and invisible effects on the direct targets. Insults against a person’s physical appearance or against a person’s career may be easily brushed off if it is a one-off message from one person or a couple of people. However, the mere repetition of a certain message from a number of persons does have an impact even if the message in itself does not include threats or severe dehumanisation. Lucía has highlighted how she feels that the messages want to take away her status as a person; that is, messages that may not *seem* extreme, can still be *experienced, and felt* as dehumanising, particularly if expressed publicly. The aggressors are not only telling *her* that she is a bad communicator, but they are telling it publicly, also discussing it amongst themselves publicly, for everyone to

see. This means that the doubt is not only created in her, but also in those reading the messages. In this way, also more subtle messages can carry the same function as hate speech through their repetition in a public space, with words or phrases concerning specific collectives seeping into the minds of both targets and bystanders.

This brings us to the importance of the distinction between private and public, which does not only concern the spaces where attacks are perpetrated, but also the very categorisation of the violence into one typology or another. Because of its often private nature, gender-based violence tends not to be considered within the hate crime framework (Lewis et al, 2019). However, this dichotomy between private and public is highly problematic in the context of the online environment. Whilst a part of attacks is perpetrated through private direct messages on social media, many of the verbal attacks against feminists online are expressed publicly and openly as comments to publications, as cited tweets, or as exposed in Chapter 7, as content in public videos ridiculing feminists. Public shaming, thus, is often an intended part of this strategy, particularly in the latter type of attacks. Here we can also observe how, as Waldron (2012) has highlighted, hate speech includes both a dividing and uniting function: *They* are ridiculed and dehumanized as different from us, whilst we are united through coming together around the dividing message. Nevertheless, this public expression of hate speech online at times also leads to the emergence of supporting networks, as I shall discuss further in Chapter 9.

In the next section, I shall analyse the impact of online violence in terms of self-expression, online and beyond.

8.5. Self-censoring

Over time, online harassment may change how and what women write about and how they interact online (Linbary and Batti, 2019). The “emotional labour” of defying abuse, progressively takes a toll on those who experience continued online abuse (Ibid.). Some women choose to remain online, but, after observing what happens to the more outspoken women who are targeted, decide to tone

down their opinions or avoid certain topics (Mantilla, 2015). This critical dimension of hate speech is also evidenced in this thesis. Many of the women interviewed expressed that they think twice about anything they post, particularly when they had been targeted by several larger attacks. For some, this change in behaviour was somewhere on the continuum between self-censoring and self-protection⁷⁶.

Daniela describes a certain self-censoring, not only in relation to possible attacks, but also in terms of maintaining a professional profile online:

I would like to talk about other things, but I don't want to get involved because I know that I'm a young girl on Twitter and that everyone comes to quote the tweets and that everyone comes to give lessons. But on some occasion, I've posted something, and I didn't consider it crazy, and also, since I know it's public, if I do it, well, if I want a certain job for a certain company, I keep a close eye on what I post. Like, I can have a political opinion, but be careful. It's not the same as what I would say in a bar with my friends, obviously. It's something that I assume is public, right? (Daniela)

Talking about self-protection, she later mentions that "I'm very cautious when it comes to what I publish and that's it. It's sort of the only thing I've done or see I can do" (Daniela). In this sense, she has internalised the limits of what she can publish without a backlash and what she cannot publish. At the same time, she describes a couple of tweets that have triggered attacks, such as a very general observation about gender roles, which led to a large-scale attack from what she describes as the "Incel timeline". This speaks of how self-censoring to avoid attacks is a very unpredictable endeavour. Attacks may be triggered by virtually any publication by a woman who defines herself as a feminist, as the attack tends to be more about her giving voice to opinions or visibly occupying public space than about the message itself. Self-censoring thus often leads to self-blame and gendered victim-blaming in the likes of "I published x so that's why y happened to me".

⁷⁶ Self-protection is discussed in Chapter 9, as a response strategy.

Susana describes self-censoring from a different perspective. After a couple attacks over a short period of time, she stayed off Twitter almost unconsciously:

After a few months I realised that I hadn't posted anything. Sometimes I thought I'm going to tweet this and I would still write the tweet and not post it. I had no impulse to write (...) Twitter stopped seeming like a safe space to me. And it had never seemed that safe either because things kept appearing, but it was really like 'you can't be here'. Then the very impulse to say 'I'm going to put this nonsense on Twitter' disappeared from me because you never know what's going to happen, where it's going to take you, so... (Susana).

As this quote highlights, it is not necessarily referred to as fear, but rather as a lack of impulse to tweet or interact on Twitter. What had previously been a very seamless and spontaneous way of communicating with people had now lost its spark. This is echoed by DE-8, who describes the lack of impulse or desire to tweet after having been attacked as, “It's not fear, it's like idleness and not wanting to look at notifications” and DE-7 who says that “It's like a disconnection and apathy towards something that used to be normal for you”.

Thus, as the ethnography reveals, it is not necessarily always fear or explicit self-censoring that keeps feminists who have been targeted by gender-based online violence from being active on Twitter or online in general, but it may also be a lack of impulse or urge to express themselves that has been taken away as a consequence of being continuously targeted. The fun and engaging part of digital spaces has been diminished when you are aware that anything that you post publicly may pose you under attack. This means that the hate speech and the attacks at least to a certain extent reach their aim: to shut women up and shut them out from public debate in certain spaces. What is left for them is to make the profile private (“poner el candado”) and discuss political and feminist issues in smaller circles – with the linked disadvantages of losing visibility. Finally, there is leaving social media, which I shall discuss below, with a focus on Twitter.

8.6. Going offline - Leaving Twitter

An evidenced impact of online abuse is for women with public feminist profiles to leave Twitter or online spaces at large. Several cases have been covered by

Spanish mainstream media over the last few years.⁷⁷ One of the most emblematic cases that took place during the fieldwork for this thesis was that of the then Mayor of Barcelona, Ada Colau. In April of 2021, she announced through a tweet that she was leaving the platform indefinitely, also explaining her reasons for this decision. Although she did not put the full blame on hate speech and attacks against her, these were woven into her discourse:

That they critique me, ask me questions, or discuss with me not only doesn't bother me, but I accept it and I like it. Politics is dialogue and debate, and thanks to the exchange of information and opinions we can improve. When Twitter started, it had a lot of that. Unfortunately, over the last years, it is known by everyone that the network has been filled with false and anonymous profiles that poison and incite to hatred. Many of these are even bought with money (bots) by the extreme right (Tweet by Ada Colau on April 11, 2021. Own translation from Spanish).

During the interviews, this and similar cases were often brought up, and proved to be present in the minds of the persons that I interviewed. Several women mentioned the case of Ada Colau, as well as other feminists who have left Twitter over the last few years, such as Brigitte Vasallo.⁷⁸ In a sense, these cases serve as proof to those feminists still active on Twitter that the violence has a real impact; that the abuse is not personal but rather structural, and that it strikes big profiles with institutional support to an extent bad enough for them to leave the platform. This seems in a way to be comforting: It is not strange that other women feel the impact, that they feel the harm. At the same time, however, it may serve as proof of the impunity of online abuse. If well-known established politicians cannot be protected from this abuse, is there any chance for a fairly unknown feminist to obtain protection?

⁷⁷ See e.g., El País (November 4, 2021), "Colau deja Twitter con 'carácter indefinido' pero seguirá en Facebook, Instagram y Telegram", <https://elpais.com/espana/catalunya/2021-04-11/colau-deja-twitter-con-caracter-indefinido-pero-seguira-en-facebook-instagram-y-telegram.html>; La Vanguardia (April 14, 2021), "El adiós de Ada Colau de Twitter, ¿víctima o visionaria?", <https://www.lavanguardia.com/vida/20210414/6684418/ada-colau-twitter-politica-redes-sociales-impacto.html>, last accessed on April 18, 2024.

⁷⁸ See e.g., El Crític (June 4, 2021), "Brigitte Vasallo: El autoodio de clase provoca que nos den vergüenza nuestra clase y nuestros orígenes", <https://www.elcritic.cat/entrevistes/brigitte-vasallo-el-autoodio-de-clase-provoca-que-nos-den-verguenza-nuestra-clase-y-nuestros-origenes-94733>, last accessed on April 18, 2024.

As Nathalie Casemajor et al (2015) have argued in their model of participation and non-participation online, the latter can be defined as a form of political action rather than as mere passivity. In this sense, leaving platforms such as Twitter may be conceived of as “active non-participation”. This highlights how in digital spaces, leaving a platform is often a reaction to perceived instances of surveillance either by the platforms themselves or by other users. Nevertheless, exiting digital spaces is a double-edged sword, as the stakes of opting out of a platform can be high for those who are more dependent on the processes of interaction on social media platforms (Ibid.), for example, for their livelihood. Indeed, the dependence on being active in digital spaces in a public way was a salient theme in the interviews. What women feminists, activists or journalists do online is an important part of their life, both in terms of their private life and their professional life, which are often intertwined on social media. When discussing the possibility of leaving certain platforms, the importance of social media for some professional profiles, such as journalists, becomes evident. That is, the *impossibility* of going offline – or even going off Twitter – is very clear to some of the targets in this research. Aurora is a freelance journalist and feminist, who needs to be visible online to achieve an income: “If I close my Twitter account and disappear, that means that I quit my job”. Sara is in a similar position, and went offline for a few days after a large-scale attack against her, but returned soon after, as she needs to stay visible online to get work:

I can't afford it [to go offline], because I do need these technologies for the work I do, and lately I have sometimes thought about it, but I feel that this disconnection is also a kind of privilege. Right now, my visibility, the visibility of the articles and the jobs that I get depend on my profile on social media. So, well... (Sara)

Amara, on the other hand, went off Twitter for a while following two attacks against her in a short period of time, but explains that she knew from the beginning that it wasn't for good: "I knew that I wanted to return but I had to think about how I was going to return, knowing that the same thing could happen to me again, they could attack me again and how I was going to manage that?" (Amara). In this sense, she knew that she could not afford to leave Twitter for

good. Nevertheless, she needed some time to recapitulate and prepare a strategy for how she would act if she was attacked again.

Having an online presence can be essential for a person's career, meaning that withdrawing from public life risks placing a person's economic livelihood at stake. The implication of this is that online harassment can have a continuous economic impact, whilst also limiting a person's potential for further employment opportunities (cf. Megarry, 2014). This poses difficulties to opting out of social media. Given the increasing importance of the Internet for business, social, and other aspects of life, to suggest that women who are harassed can solve the problem by simply staying offline becomes analogous to telling women to quit their job if they are being sexually harassed at their workplace (Mantilla, 2015: 100). Or, as Citron (2014: 8) highlights,

targeted people who curtail their online activities or go offline incur serious costs. They lose advertising income generated from blogs and websites. They miss opportunities to advance their professional reputations through blogging.

Further, when a person goes offline, the evidence of abuse, harassment, or character degradation does not disappear with them. In a context where employers regularly complete internet background checks of potential employees, going offline does not necessarily undo the effects of the harassment, neither psychologically nor financially (Citron, 2009: 386). The use of a pseudonym, on the other hand, prohibit women from networking "professionally or effectively" (Ibid.: 398) on the Internet and thus reduce the legitimacy and authority of their online contributions, and can thus, in terms of impact, be compared to going offline or making the account private.

Leaving certain social media platforms can be a crucial step toward constructing new platforms that enable active participation; however, crucial to this shift is an exit that enables time, resources and skills to design and build a new online environment (Casemajor et al, 2015). Recently, we have been able to observe this act of active non-participation with many politically engaged individuals leaving Twitter following Elon Musk's takeover and the subsequent changes in platform

policies. Many of these individuals have opted for open source, decentralised social networking sites, such as Mastodon, that have explicitly posed themselves as alternatives to Twitter.⁷⁹ Nevertheless, departing from a platform can also signify a final act, where individuals choose to discontinue their use of a specific service or technology, without transitioning to or engaging with any alternative options (Ibid.).

Due to all the factors exposed above, some feminists with a public profile decide that engaging in public discussions of political issues on Twitter is not worthwhile. They may have other venues for this, as in the case of political profiles such as the former Mayor of Barcelona. In other cases, they may move their discussions to other platforms, with the hope that these are less hostile, or decide to engage in discussions on the same platform, but with a private configuration of their profile. Still others, as we have seen, cannot afford to lose visibility on social media and particularly on Twitter, due to their professional status, often as freelance journalists, or communicators. In the next section, I will look at this in relation to the collective consequences.

8.7. The wider consequences of online violence for the participation of voices in the public debate

Grounded on the ethnographic evidence of the impacts of online violence, what conclusions can be drawn about the wider consequences of this violence for the participation of voices in the public debate (online)? I have written online in parentheses, as it is crucial to consider that these attacks are not only perpetrated when women express feminist opinions *online*, and they do not only have an impact on their online actions. The attacks are perpetrated online; however, women's voicing of feminist opinions, which often triggers the attacks, may take place on any media, including television, radio, newspapers and so on. Further, the impacts go beyond women's participation online. As we can see by

⁷⁹ The welcome note on the Mastodon web (<https://joinmastodon.org/>) states "Social networking that's not for sale. Your home feed should be filled with what matters to you most, not what a corporation thinks you should see. Radically different social media, back in the hands of the people".

the experiences they share during the fieldwork, women may refrain from debate on Twitter, but they may also refrain from participating in political debate in any public space. That is, it is not only the online space that is affected, but we need to analyse the impact on the public debate as a whole. Clearly, to comprehend the full effects of online abuse it is necessary to understand the online and offline as interconnected spaces.

As anthropological scholarship has highlighted, to equate voice with power and voicelessness with silence relies on specific assumptions about speech as the primary form of expression, as the source of individuality and self-worth (Chadwick, 2021; Kunreuther, 2014;). Linked to this, within feminist new media studies, scholars have discussed the politics of *online* visibility for women (e.g., Kingston Mann 2014; Portwood-Stacer 2014). As Larisa Kingston Mann (2014, 293) argues, “feminists ... are well positioned to point out that being visible or accessible to others is not necessarily liberating and that having the ability to say ‘no’ and deny others access to one’s image, words, or creative output can be a requirement for liberation”. Being publicly visible online is not necessarily the right path for the feminist movement, as these authors have also shown. As I have previously noted, in contrast to the separatist organization of women’s discussions during the second wave of feminism, social media technologies bring women’s politics into public spaces structured by male-owned and male-controlled global corporations (Megarry, 2017). In a sense, there is a lack of value attached to women-only activism in digital feminism, a reflection which may make us question how women can examine patriarchal ideologies when they are busy “involving, engaging and ‘educating’ men” about feminist issues (Long, 2015: 149).

Similarly, digital cultures research has challenged the common assumption that social media platforms provide marginalised groups with more power than traditional media (e.g., Gillespie 2015; Megarry, 2017). It is necessary, then, to link this query to the larger debate surrounding the democratic potential of social media, which I have touched upon in Chapter 2 of this thesis. As Christian Sandvig (2015: 2) puts forward, scholars who have conceptualised social media

in terms of “democracy” may have used “the wrong metaphor”. It must be considered that the algorithmic content selection of social media platforms is neither transparent nor neutral. Rather, social media content appears or disappears according to reasoning that is unknown to those outside of the social media corporation that controls the specific platform. In this regard, it is crucial to acknowledge that in their technical, economic, and political design, platforms “pick and choose” what ideas are accessible and socially appropriate (Gillespie, 2015). The platforms’ moderation teams decide what content is allowed to circulate, and what content is deleted. For example, platform moderators often decline to delete posts or comments that women report as misogynist (Megarry, 2017). This means that even without threats, attacks or harassment, online participation is not necessarily democratic, and voices are filtered through an array of different mechanisms.

This analysis of the platform structures has led to a more critical scholarly discussion on the democratic potential of social media platforms, leaning towards pessimistic perspectives on political participation online, particularly for women (e.g., Hasinoff, 2014; Casemajor et al, 2015; Megarry, 2017). Instead of being fully emancipative and transformative, social media technologies may be more accurately understood as “powerful tools of social and political control” (Langlois, 2015: 1). Due to this, in a similar manner to the word “mass” in mass media, the word “social” may be the biggest challenge facing those who study social media (Sandvig, 2015: 3). According to Theodor Adorno, the word “mass” may give the impression of “something like a culture that arises spontaneously from the masses themselves” (1975: 12). In the culture industry, the audience may select programs, buy tickets, or change channels, but in doing so, each person acts as “an object of calculation; an appendage of the machinery” and the demands are produced and shaped by the industrial apparatus itself (Ibid.). Analogous to this, within behaviours on social media there is a genuine human sociality, but what has been built around it is a “calculation machine of dizzying complexity” with users acting “as its input, output, or gears” (Sandvig, 2015: 3). That is, for example, when a user posts content on Facebook, intended as a

public communication with their friends and family, Facebook evaluates the content and may decide not to show it to anyone, or only to a few people. Our intended social act of updating our family, friends or wider community on our lives may in fact not reach the audience that we believe it reaches and, in that sense, may not be as 'social' as we believe. These algorithmic mechanisms are still highly obfuscated to the general public.

Nevertheless, with this preface in mind and based on the argument that to defend fundamental rights, it is desirable to be able to express feminist and/or anti-racist opinions online and offline without being attacked, I will point to some reflections on the wider consequences of gender-based online hate speech for the participation of different voices in these debates. Scholars often refer to the "chilling effect" that online attacks and abuse have on the voices of women in the public sphere (e.g., Jane, 2015; Sobieraj, 2017; Eckert, 2018). However, whilst online violence may indeed have a chilling and silencing effect on some women, it may also have a galvanising function.

In this regard, negative impacts reduce over time, while feelings of being strengthened and encouraged to act on political views may increase in the long term, as previous studies show (e.g., Lewis et al, 2019). Nevertheless, the galvanising effect may be more evident in a sample of feminists than in the wider population of women: "It may be that women who are engaged in feminist debate draw on a feminist analysis to understand their experiences, whereas women less or not at all engaged, might have fewer resources to draw on when dealing with online abuse and might feel silenced by it" (Ibid: 136). That is, we need to be careful in trying to extrapolate findings that are based on a sample of feminists, across all groups of women. This points, again, to the importance of introducing an intersectional perspective that also considers the influence of political stances, levels of activism and professional profiles on a person's identity and resources.

The feminist filter of analysis mentioned by Lewis et al (2019) was distinct in most of the interviews that I conducted with targeted women. There was a clear

awareness of the function of the attacks, and the perception of a need to hold up and stand strong, for example, as recounted by Alba:

You have to be very strong. You have to believe a lot in yourself and put some very clear limits to what they are telling you because if you don't, they get you down (...) The hatred you generate is because you're a woman expressing your opinion freely and on top of that you are giving the counterpoint to their oppressive speech and they are not used to it (...) And you have to do the exercise of saying that all the hate I'm receiving is worth it, basically I'm receiving it because of the message I'm sending. It's what must happen in the sense that the message confronts their message. The message wants to dismantle this system, so the violence it generates in others you can view it as proof that you are doing your work well, right? (Alba).

As this quote evidences, violence is perceived as a response to a message that challenges the norms, and the very triggering of violence is understood as a sign that the message somehow has got through, as it awakes a reaction in those who are against it. This may, in turn, strengthen the feminist views of the targets and their will to continue their activism. This can, to some extent, be interpreted as a stimulating effect of the abuse. Still, this effect exists in parallel to changed behaviours and limited engagement in selected spaces and debates (cf. Lewis et al 2019). This means that it is not so much a matter of recipients of abuse being either silenced or galvanised, but rather, there is a co-existence of both consequences. The multitude of advantages offered by digital participatory media creates a complex dynamic. Some people reject the terms of online engagement, while others, despite recognising the risks, choose to embrace them (Casemajor et al, 2015).

The results of the research conducted for this thesis evidence that the abuse suffered by these targets has a discouraging and tiring effect on them. The overall reading of the added-up impacts is a sense of women being fed up and tired. Several of the women, both in the interviews and in the digital ethnography, express a deep and frustrating sense of fatigue stemming from always being on guard, of having to think about self-protection measures, of self-censoring, and of being afraid of possible new attacks. The word "tired" appears repeatedly. DE-

13, a feminist gamer, describes the collective tiredness of feminists in her response of support to a young feminist who has publicly denounced online abuse on Twitter:

We're tired of living with daily threats of misogynist sexual violence, physical and psychological. No one should have to stand this because they want to be present on social media. We're fed up with having to leave the internet (DE-13).

Another day, denouncing online abuse that she has suffered herself, she comments on her own publication "I'm very very very very tired" (DE-13). Nevertheless, because of their feminist analysis, the common perception of these women is that they must carry on, that someone has to carry the burden. Lucía compares it to taking turns in a boxing match:

But it's very tiring, so we also take turns so that one isn't always in the ring, but the ones who put our bodies out there, fuck [...] It's very exhausting and discouraging (Lucía).

Again, parallels are drawn to a physical – embodied – fight, putting your body out there, in the ring, constantly prepared for an attack, taking turns to be out there in the arena, thrown to the lions. And that is tiring. However, to "take turns" is part of this feminist analysis; those targeted are not alone, but they are often surrounded by a supporting network of fellow feminists. This I shall address in the next chapter, as part of practices and strategies of response and contestation.

8.8. Conclusions: Lingering emotional impacts, ongoing vigilance, and the chilling effect of online abuse

Based on the above evidence and analysis, online abuse prompts women to alter their behaviours, influencing the discussions they engage in, and even the feminist figures they publicly support (Jane, 2017b: 45). The impact of such abuse is multifaceted, manifesting as real harm emotionally, mentally, and physically. The interconnection of online and offline realms reveals the pervasive and lasting effects of online violence, extending beyond digital spaces and affecting various aspects of life.

Repetitive hate messages, leading to a persistent sense of vigilance, can result in the internalisation of harmful language, creating a form of linguistic injury. Despite recognising the intended ridicule, targets may grapple with the repeated messages, contributing to a sense of disorientation. Public shaming is a central strategy, illustrating the dual function of hate speech—dividing targets while uniting perpetrators through a shared message.

Women who have experienced ongoing online abuse often find themselves reconsidering their digital behaviours and interactions, navigating a complex continuum between self-censoring and self-protection. Left with limited options, some women retreat by making their profiles private or discuss feminist issues within smaller circles, incurring the cost of reduced visibility, or, in some cases, choosing to leave the platform altogether. The violence these women encounter is interpreted as a response to a feminist message that challenges norms. Negative impacts may diminish with time, while a sense of empowerment and increased motivation to engage in political action may potentially grow in the long run. However, it is important to note that this empowering effect might be more pronounced within a specific subset of individuals who identify as feminists. This underscores the significance of adopting an intersectional approach that considers factors such as gender, racialisation, age, and religion, but also individuals' political beliefs, activism levels, and professional backgrounds. The galvanizing effect often coexists with changed behaviours and limited engagement in certain spaces and debates.

The comprehensive collective effects of online abuse are difficult to gauge fully. Online attacks are often brushed off with the idea that those who are targeted can just stay offline to avoid attacks; nevertheless, both the individual and the collective effects extend beyond the online realm, influencing public debates and collectives across various mediums. Indeed, leaving digital spaces is a complex decision with high stakes, affecting income, professional opportunities, and networking. Staying online is often imperative for those whose personal and professional lives are intertwined with digital spaces. Precisely those women who need to stay online because of professional commitments – often as “self-

employed” - make up the profiles that more than anyone risk becoming the symbolic prey of anti-feminists, sacrificing themselves in the ring, with consequences not only for online interactions at the individual and collective level, but also with strong *embodied* consequences for the individual. The metaphorical battle echoes the sentiment that feminist voices, despite the fatigue, must remain present in public spaces, fighting against the onslaught of gender-based hate speech. Constructing alternative platforms that facilitate active participation can be a crucial step, but building a new online environment requires time, skills, and resources.

9. Navigating Online Gender-Based Hate Speech: Responses and Contestations

I do believe that we are also defenceless there, we love and care for each other, but that's basically it. We don't have anything else. Sometimes it's not enough. The proof is that women have left Twitter, right? [Lucía]

9.1. Introduction

As has been exposed throughout the previous chapter, women who are targeted by gendered hate online navigate the digital realm with a heightened awareness of potential threats, with a certain level of vigilance as part of everyday life. Resistance to online hate takes various forms, including both individual and collective measures and responses. The range of responses and contestations is extensive: Some are aimed at mitigating the immediate impact of the violence and to protect targets from further abuse, whilst other responses have broader aims such as creating awareness or building community for sharing experiences and exchanging support among peers.

Whilst women who are targeted by gender-based hate speech often identify attacks as attempts to delegitimise their online presence, the attacks may end up

being counterproductive since the targets' resolve to political engagement can be strengthened as a result (Lewis, 2016). Thus, the very mechanisms that create the potential to hurt with words, at the same time open for a possibility to undo the harm with words, as Butler (1997) highlights in relation to linguistic injury. Engaging in discourse that challenges and counters injurious language, thus, may be a form of resistance, that can contribute to altering the prevailing narrative.

Responses to online misogyny can be divided in the two broad metaphorical categories "flight" and "fight" responses (Jane, 2017b). The "flight" category may include distancing oneself from the abuse; restricting internet use; or putting into place self-protection mechanisms. In contrast, the "fight" includes strategies such as confronting the abusers, engaging in digital vigilante - "digilante" - activism (Ibid.), constructing counter-speech, or reporting to the police. In practice, the deployment of different response strategies is contextual, as it depends on the specific situation. Still, women may adopt several different responses at the same time, used in parallel or overlapping.

In this chapter I will examine the array of responses and contestations by those targeted by online gendered abuse or hate speech, encompassing both "fight" and "flight" responses. To this aim, I will use the following main categories, that have emerged from the ethnographic analysis as the main responses to gendered online violence: Digital self-protection strategies; Publicly speaking out; Building community; and Interacting with the justice system.

9.2. Digital self-protection strategies

Throughout the course of the ethnography - encompassing both digital observations and in-depth interviews - a common thread emerges among all the women who are targeted by gendered hate online: their use of digital self-protection strategies. Following Emma Jane's (2017b) proposed categorisation, these responses to digital abuse would be sorted under the broader label of "flight responses".

Prior to delving into platform-specific protection measures or strategies tailored to social media platforms like Twitter, Facebook, or Instagram, it is essential to

comprehend how individuals targeted by online hate seamlessly incorporate general measures into their online social interactions. As I have exposed in Chapter 8, as an impact of being targeted by online hate speech and violence, the persons targeted to some extent change their behaviour online, through avoiding certain topics or being less active on social media. Further, to protect themselves, they also put into place measures linked to the affordances of digital spaces. For example, they avoid sharing private information or any geo-location information that could let followers know where they are physically located:

I'm very careful about what I post in the sense that I don't put my home address or where I live, and things like that because it kind of gives me bad vibes and the world is a bit like... bad (...) I keep a close eye on what I post and so on. A photo where I'm at home, for example, I only post it so that no reflections, no windows, or anything that can be identified can be seen (Daniela).

These self-protection mechanisms are perceived as necessary and are normalised to the extent that they are put into place and described without much reflection. These strategies, thus, speak of the acceptance and normalisation of the presence of violence online. As Daniela puts it, as a young woman with a public feminist presence, she feels that the world is “bad” and it is up to her to protect herself. A generalised protection strategy related to this is the use of several accounts on the same platform. Usually, one account is fully public and thus open for anyone to follow, and the other account is private and only open to friends or other trusted persons: “I have another Instagram account since I've been away [abroad] where I do tell my friends a bit about my life, but this is private, and the vast majority that follow me are unaware of the existence of that account” (Daniela). There is little cross-referencing between the accounts, to keep the existence of the private account hidden from most followers of the public account.

Furthermore, there are also other kinds of strategies on different social media platforms that are related to the protection mechanisms, such as blocking, silencing, filtering or reporting, which I will discuss in the next sections.

9.2.1. Blocking, silencing, and filtering

In addition to the general digital protection measures discussed above, targets tend to use platform-specific measures, which nonetheless are rather similar across platforms. These responses vary depending on the public visibility of the target, their number of followers on social media, their experience in managing online attacks and the intensity of the attack (e.g., Morena-Balaguer et al, 2021). The most habitual and commonly used of these are blocking, silencing, or muting users, as well as filtering out words. These strategies allow targets to stop aggressor accounts from further interacting with them. However, the use of these mechanisms is invisible to those who are not directly involved, that is, to everyone except the person who blocks and the persons who is being blocked (Ibid.).

Blocking abusive or specifically anti-feminist profiles on social media platforms is a strategy widely used among the targets included in this research. These measures are mainly applied in relation to a direct attack; when receiving more sporadic insulting comments; or when a target perceives that a profile is commenting only to derail the debate – a behaviour that has been referred to as “sealioning” (Johnson, 2017). Blocking is, thus, a strategy highly integrated within the social interactions on platforms such as Twitter:

What I do a lot like at the level of protection is silencing and blocking very easily. I also realized that there’s the typical user that everything you publish they come and comment saying any nonsense and for a long time I endured too much and now it gives me much more satisfaction to block them quickly (...) I think I do it more since that thing [the massive attack] happened to me (Sara).

Lucía recounts, in direct relation to an attack against her, how she started blocking the users that were attacking her and that she

started blocking, one after another, all the persons that were insulting me. Maybe they were 500, I don’t know, they seemed to be so many, so I spent like 6 or 7 hours during which the messages didn’t stop, until I decided to cut it off (Lucía).

By “cutting it off”, she means that she turned her profile private and did not allow non-followers to send her messages. This measure, to turn the profile private, is

the only strategy that effectively works to cut off the attack in the case of collective and coordinated attacks (e.g., Morena-Balaguer et al, 2021).

Sometimes the response of putting platform-specific protection measures into place can be categorised under Jane's definition of technological hygiene, that is, aiming to mitigate the impact of attacks through outsourcing the checking of social media during attacks or when anticipating them. Sara relates how during the massive attack against her, she was lucky enough to have a couple of friends who did the blocking and erasing of messages for her: "Some friends of mine took it upon them to act like a filter and clean up all the content both on Twitter and Instagram". This was possible as she had tech-savvy feminist friends who had experienced similar situations and, in that moment, knew what kind of response was needed. There are also cases of targets who have formally outsourced this task. An illustration is the case of DE-12, a trans woman subjected to a high level of abuse, who relates how she has been able to fully outsource this task, through paying a person for reading direct messages, moderating comments, reporting, and blocking.

Blocking is often directly coupled with reporting to the platforms, which is a strategy discussed further below: "When I receive messages with very explicit things and I say 'damn, that's too much', I block the profile and I report it" (Daniela). Some platforms, such as Instagram, explicitly link the two options so that if you choose to block a user, the interface directly asks you if you also want to report them. Some targets, however, believe that blocking may give haters or attackers a feeling that they have reached their goal:

I don't usually enter to confront people. So I silence them. I don't even block them, because depending on which profiles, trolls, haters, as you block them, they grow, because they believe that they have achieved their objective that you have become obfuscated and have blocked them. I don't even want to give them that pleasure. Silencing takes 10 seconds of my time, that's all I dedicate to it and that's it (Amara).

Further, blocking a user who produces or spreads offensive content through comments does not stop the abuse, it merely either makes it invisible to the target or stops that specific profile from sending hateful messages to that

specific target (Smith, 2019). That is, at least momentarily, it mitigates the impact on that individual target. However, a user can easily create a new account, or target other people with the old account even if it has been blocked by several other targets. As Amara expresses, some prefer silencing, or as it is formally denominated on Twitter, “muting”, as it is not possible for a user to know that they have been silenced by another user. On Twitter, this is a feature that allows you to remove an account's tweets from your timeline without unfollowing or blocking that account.⁸⁰ Further, as illustrated by Amara above, some targets have integrated among their protection strategies not to quote tweets or start discussions with anyone who attacks them. Instead, they directly block the user in order not to spread the hate message further, or risk starting an attack by their own followers on that person, as Daniela explains:

One thing I also do is that I never get into controversies like quoting and quoting and quoting, because I don't like it, I'll pass. Besides, I'm very aware of those profiles that have like 30, 40, 50 followers and whether it's a real person or a bot doesn't matter to me, I don't quote it, because I have a thousand people who follow me and they like me and they're going to send a lot of messages and I don't like to get into that dynamic either and then I pass. I simply block them (Daniela).

The decision not to enter arguments or quote hateful content is at least in part perceived as a way of avoiding a dynamic where a counter-attack is produced.⁸¹

Another digital protection strategy detected is that of filtering out certain words. In this regard, Sara describes how her friends helped her to add filters immediately following an attack, so that any subsequent comments that contained words such as “feminazi”, “ugly” or “fuck” would not reach her. This means that the attackers may still send messages and comments that contain these words, but the person targeted does not have to see them. Putting filters

⁸⁰ See <https://help.twitter.com/en/using-twitter/twitter-mute>, last accessed on April 8, 2023. The discussion of support mechanisms refers to the mechanisms in place before the change of the platform's name to X.

⁸¹ Strategies involving counter-attacks, naming and shaming, or “digilantism” (Jane, 2017) will be further discussed in the next sections.

into place or blocking users can also be conducted in an automatised manner⁸² as in the experience of DE-13, who comments on a post that

I would like to add that in my personal case this Twitter violence has reduced considerably since I use automatic blocking tools. I say this because even if you lose a lot of followers, visibility and there are accidental blocks, it is worth it. It's a matter of survival (DE-13).

While these strategies do not necessarily counter hate speech or misogyny at a collective level, they protect the targeted individual and may make participation on digital platforms slightly more amicable, as DE-13 explains.

When reflecting upon the different digital protection strategies together with those targeted, it becomes clear that due to the constant risk of being targeted by abuse, some of these women have developed their own implicit roadmaps of protection and protocols that they follow. Daniela explains her internal rules in this regard:

This is a rule that I have: after 100 likes on a tweet - I have 3,000 or so followers, so that is normal - I silence the tweet and I don't look at it until a few hours later and it's like a protocol that I have. If I see that they have quoted it and there are people saying atrocities or that they are simply spouting hate, then I start blocking profiles without any kind of sense, I block all of them because I don't want to... I don't have to read that, because I know it's a person who doesn't know me, but I also don't want 10 quotes from people who are systematically attacking me and who know absolutely nothing about my life (Daniela).

In a similar vein, Amara recounts that:

If they quote me and put some racist stuff, what I usually do is either I don't answer, or I quote that person to expose them (...) but I don't usually confront, I don't usually go in to confront. So I say that to that person and then I silence them (Amara).

Having these internal “protocols” makes the response strategies more efficient. This implies that targets do not have to dedicate time to ponder on what specific response they will use in each case. Nevertheless, these internal protocols, as well as the use of automatic blocking and filtering tools, also reveal the climate of

⁸² This can be achieved by using tools such as MegaBlock (megablock.xyz), which lets a user block a tweet, its author, and every single person who liked it, through one click.

constant vigilance that those targeted by online hate speech have learned to live with.

9.2.2. Reporting to the platforms

In this section, I will zoom in on the use of the reporting functions of social media platforms. As the focus of the digital ethnography has been Twitter, much of the reflections will stem from interactions on this platform. However, I will discuss the use and perceptions of reporting mechanism at a more general level avoiding a focus on specific platforms, as platform-specific reporting mechanisms are in constant change. As highlighted in Chapter 2, moderation is a complex task and a company's policies often go beyond illegal content since there are categories that are legal, but still harmful. A great deal of the content that targets those interviewed for this thesis would fall under that category. Social media platforms play a crucial role in user protection through content moderation, a fundamental and essential aspect of their role in public discourse, shaping the content users see and constituting the platform itself (Gillespie, 2018). Despite its significance, moderation is often hidden to maintain an illusion of openness and to avoid legal and cultural responsibility. Recognising online content moderation as a governance issue (Udupa et al, 2022), the complexity arises from policies going beyond illegal content to address legal yet harmful categories. Online hate speech adds to the complexity with its diverse and dynamic nature, taking various forms that can shift rapidly (Brown, 2018; Citron, 2014; Delgado and Stefancic, 2014). It often involves rhetorically complex elements, including metaphors, culturally-specific stereotypes, and dog-whistles. Furthermore, moderation extends beyond text to encompass memes, multimedia materials, hashtags, tagging, and other forms (Siapera, 2019).

Frequently, changes in platforms' moderation protocols move towards a perceived improvement. To illustrate this, in June 2022, DE-6, a trans woman, publishes a screenshot on Twitter of a transphobic post against her. She asks her followers "Could you please report?". Many of her followers reply that they have reported to the platform, and several point out that the reporting mechanisms have improved lately:

“Super reported. Also, the new reporting model on Twitter makes it possible to report content in a very precise way. I hope they delete the account...”

“Reported, love. They’ve changed the reporting model and it’s really good now, to be honest”

However, Twitter changed ownership in October 2022 and since then there have been numerous reports in media and on social media regarding decimation of staff in different departments, including those dedicated to content moderation.⁸³ Media reports have also concerned the return of far right- and hate groups to the platform and the impact that this may have on vulnerable groups, minorities and the general level of hate speech.⁸⁴ That is, less than five months after these positive comments, the conditions for reporting had likely already changed for the worse. In early March 2023, Twitter announced a new violent speech policy, aimed to prohibit violent threats or wishes of harm, glorification of violence or incitement to violent acts⁸⁵. Included in this policy - and not present in any of the previous policies - is the mention of "dog whistles" or coded speech often used by hate groups to indirectly incite violence. Nevertheless, given the extensive cuts in the platform’s moderation staff, it is unlikely that this policy is being or will be implemented to any meaningful extent.

As I have previously hinted to, in parallel to blocking users, reporting to the platforms is a strategy that is well integrated in the digital protection strategies of the targets. Nevertheless, their narratives often speak of disappointment with how the platforms handle their reports and thus moderate content:

⁸³ See e.g., “Elon Musk Cuts More Twitter Staff Overseeing Content Moderation” (January 7, 2023), <https://www.bloomberg.com/news/articles/2023-01-07/elon-musk-cuts-more-twitter-staff-overseeing-content-moderation?sref=10lNAhZ9&uuiid=gXnRxrK4kXTXdlfr0745#xj4y7vzkg>, last accessed on April 18, 2024.

⁸⁴ See e.g., “Banned British Far Right Figures Return to Twitter within Hours of Takeover” (October 29, 2022), <https://www.theguardian.com/politics/2022/oct/29/banned-british-far-right-figures-return-to-twitter-within-hours-of-takeover>; “As Elon Musk Opens Twitter, Whither The Digital Town Square?” (November 10, 2022), <https://www.cigionline.org/articles/as-elon-musk-opens-twitter-whither-the-digital-town-square/>; Why is Elon Musk’s Twitter Takeover Increasing Hate Speech? (November 23, 2022), <https://www.brookings.edu/blog/how-we-rise/2022/11/23/why-is-elon-musks-twitter-takeover-increasing-hate-speech/>, last accessed on April 18, 2024.

⁸⁵ See e.g., “Outage-hit Twitter muddies violent speech policy” (March 1, 2023), https://www.theregister.com/2023/03/01/twitter_violent_speech_policy/?td=keepreading, last accessed on April 18, 2024.

There have been so many times that we have reported people with openly racist discourses, like saying that all immigrants or all Muslims must be killed, and nothing happens. You don't trust it anymore (...) At the beginning of using Twitter, I reported more, but now what I do more is blocking. And that is a bit to save yourself, because [when you report] you get the report that it doesn't break the rules and then you see people who haven't broken the rules but got reported and they get the account closed... (Aurora).

This perception is the same regardless of the platform, with similar experiences described across different platforms, including “traditional” social media platforms such as Instagram, Facebook and Twitter, but also video platforms such as YouTube:

Two or three years ago I was subjected to racist hate speech on YouTube and I reported the users because the insults were very explicit and even then they said that the user didn't infringe any norm. In the community norms it says that they don't allow messages that incite hatred on the grounds of race, gender, religion etc. I report it and I receive an automatic response that it doesn't breach the community norms. So what do they need to say to me to break the norms? (Amara).

Or platforms predominantly linked to the gaming world, such as Discord or Twitch:

It's like Discord, you report, and nothing happens. On Discord, I reported loads of times. They said they'd look at it and it's been a year. So I haven't seen any punishment, when I've reported I haven't seen any punishment, so it's normal that they do it with this impunity and they keep doing it because there's no punishment (...) So, they put the code of conduct and that shit just to look good, and because someone asks them to because it's a social network, but they don't do anything. I already told you that they do nothing. It's been the same all my life, reports to the platforms, and they don't cancel accounts or do anything. They tell you that they are going to do it so you will shut up. Then you check and the server is still there, the account is still there (Elena).

There is a strong perception among the women interviewed that platform operators know that they must have a code of conduct, but that these codes are not being enforced to protect those that need protection. The lack of efficient response by platforms has thus given rise to a widespread lack of trust by those

targeted by violent speech. After putting forward a report to a platform, the most common case is to receive a notification stating that the reported user or publication does not infringe the community norms. As an illustration of this, Susana can only recall one report that led to any action by Twitter:

Whenever I perceive that something should be reported, I report to Twitter or to whatever platform it is... But I have only received one from Twitter where... because Twitter notifies you when your complaint has reached a good point. And recently on Twitter they notified me for the first time that this tweet or that this account had been deactivated, but the rest haven't led to anything (Susana).

This leaves those who are targeted by online hate with a sense of impunity, and with a perception that the decisions whether a publication infringes on the community norms are highly arbitrary:

Well, sometimes I get the automatic message that they consider that the reported content doesn't infringe any norms and it's like 'what?', and on other occasions you do manage to get the account closed and that person opens a new account and we're in the same spot again (Amara).

That is, for any platform that includes user created content and thus moderation measures, these measures are perceived by targets of gendered hate as inefficient, non-transparent and arbitrary. The arbitrariness is also experienced by the targets when feminist accounts are reported by anti-feminists or far right profiles and get shut down: "You see people who haven't broken the rules but who get reported and they get their account shut... It's just that the system is rigged" (Lucía). There is, thus, a feeling that the reporting system is not there to protect them, but that it rather as easily can be used against them. That is, the reporting system is interpreted as a reflection of the wider structures that uphold an unequal society. This exacerbates the lack of trust in the platforms and gives rise to a perception that platforms lack real will to protect vulnerable groups. As I have discussed in the conceptual framework of this thesis, the very structure of the internet in terms of algorithms as well as the processes put in place for moderation, bear the same structures as the rest of society and cannot be considered neutral. Whether moderation is conducted by mechanisms based on artificial intelligence or by hired human staff, bias always exists. Nevertheless,

despite this feeling of impunity and arbitrariness, most targets tend to keep reporting, and ask their friends and followers to report profiles that have attacked them, as we have seen above.

Nevertheless, not everyone uses the existing platform reporting mechanisms. Amongst the targets interviewed, one woman replies that she does not tend to report, not even to the platforms. She reflects upon this by relating it to the normalisation of hate messages, within the wider LGBT+ community that she identifies herself with:

What I think happens a lot when we receive criticism on social media is that we don't... or at least to me what has happened is that it hasn't occurred to me to report it. Because, I mean, there are so many messages that I encounter in my day-to-day life that I justify with this ignorance and I don't react like 'hey, this is a hate message, delete it'. And maybe it's something that doesn't only happen to me but more people from the [LGBT+] community, and maybe because of this I can't think of more cases because it's like you ignore it (Julia).

Like most strategies of response, reporting and applying measures of self-protection place an unfair burden on targets (Jane, 2017) - a burden that becomes exacerbated by the low interest of platform operators to establish transparent measures to protect particularly vulnerable groups. Misogyny tends to fall into a gap, where publications reported are not interpreted as hate speech, likely due to the general lack of inclusion of misogyny under the hate speech umbrella. Further, the specific form of misogyny that intersects with abuse on other grounds such as racism or LGBT-phobia, is not captured by these reporting mechanisms, as only one ground can be selected for reporting. For example, Facebook allows to report hate speech with a list of grounds, which includes "sex or gender identity"; however, only one category can be chosen. This means that, for example, a racialised woman must choose whether racism or misogyny is the more powerful ground in the abuse, although to understand this specific strand of abuse against racialised women, an intersectional perspective must be applied. Choosing only one category does not correctly reflect the level of abuse or the protection needs of the targeted person. The lack of possibility to choose more than one hate speech ground, together with the lack of response by platforms in

relation to gendered hate, can be interpreted by targets as a message that they are not deserving of protection.

Digital protection strategies may not address the roots of the issue or even hinder individual perpetrators of spewing hate; however, if mechanisms for blocking profiles and reporting content work well, they will at least contribute to mitigating the impact on the targets.

9.2.3. “Poner el candado”

Blocking or silencing users or reporting to the platforms may not be enough to break off the wave of abuse, during a large-scale attack. Further, it is highly time-consuming, as the attackers need to be blocked one by one. In the case of Lucía the abuse becomes overwhelming, and she finally decides to make her profile private. This is often referred to as “poner el candado” (“put on the padlock”), as Twitter users who do not follow the profile will see the image of a padlock if they try to access the profile. The use of this option on Twitter allows a user to make their posts private so that only those who were following at the time when the account was made private can see and interact with the posts. This is a strategy often used for shorter periods of time, until the attacks against a target calm down. Turning the account private may work both as a protection mechanism and as strategy of subtle denouncement, as this change is visible to other users (Morena-Balaguer et al, 2021). Whilst other users may not know exactly what has occurred, an account that has suddenly been made private is a sign that something has happened to the user.

Nevertheless, not only massive attacks make users turn their accounts to private mode. Also a sudden increase in followers may make some women feel uneasy, particularly if they have been targeted before and do not know the true intentions of their new followers. As Daniela explains: “A lot of people I didn't know at all began to follow me and they did write to me and such and it seemed a bit disturbing to me and I did make the profile private for a couple of days, like ‘I don't know what's going on’”. What had happened preceding her decision to make her profile private, was not a direct attack. Instead, she had gained so

many new unknown followers over a short period of time and decided to “go private” for a few days, whilst figuring out where these followers came from and what their intentions were. That is, the decision to put on the padlock in this circumstance is based on a perception of the need to be on guard, based on previous experiences.

Building upon this discussion, in the next section I shall examine another set of response strategies, which can be sorted under the broader category of “fight” responses.

9.3. Publicly speaking out

Feminist activists have traditionally relied on the sharing of experiences, both privately and publicly. That is, sharing experiences of online violence through a public narrative can in part be understood as a strategy of feminist contestation. In this regard, four main rhetorical strategies have been identified by Linabary and Batti (2019), as used by women within their public narratives online to enact resistance against harassment and hate. The strategies include: First, calling out the abuser(s), including specific examples, often accompanied by screenshots of quotes from the harassment and abuse they have experienced. This can be interpreted as a form of “digilantism”, or “feminist digilante responses” to gendered cyberhate – often referred to as “naming and shaming” (Jane, 2017). These responses run along a spectrum going from drawing attention to the abuse, but without identifying the perpetrators; exposing online identities of perpetrators; and finally, tracking down perpetrators and exposing them to the broader public, often outing them to family members, friends or employers. The second category corresponds to representing shared experience, which tends to involve more in-depth writing about individual experiences to represent the larger network of experiences that women face online. The third category involves disrupting the hegemonic narrative. This includes challenging both hegemonic narratives of the division of the physical and the digital in two separate spheres, and hegemonic narratives of online harassment. This is, for example, through challenging the ways people commonly think that women should respond to

online abuse. Finally, the fourth narrative strategy involves defying the harassers, that is, making direct statements of defiance to reclaim power.

This thesis reinforces this categorisation of strategies of public resistance. The majority of those interviewed have, at some point, publicly talked about the hate and abuse that they have suffered online. Among the targets included in the digital ethnography, most of these have called out harassment or attacks publicly, either on Twitter, or in other media or social media outlets. In the following sections, I will discuss the different strategies used. First, a note: I refer to these categories as “strategies”; however, strategies allude to a conscious and planned manner of response. In the cases included in the present research, this is sometimes true, that is, at times this manner of response is carefully planned; however, often, the manner of response is not consciously applied as a “strategy”. This means that the categories below should rather be understood as different patterns of response, which are sometimes spontaneous and without too much thought or planning.

9.3.1. Calling out the abuser(s)

The main narrative strategy identified in this research is calling out the perpetrators. This manner of response is used for several reasons: to receive support and advice from the wider community; to make other persons aware of the attacks that women who express feminist and/or anti-racist opinions suffer; and to encourage other users to react against the perpetrators, either by reporting the profile or a specific publication, or blocking them to avoid being targeted themselves. That is, the calling out of the perpetrators functions at the same time as awareness-raising, as a warning, and as a call to action.

In May 2021, DE-1, a feminist journalist, publishes a screenshot of a tweet that mentions her Twitter-handle and that includes the text “you’re trying to be good-looking and you’re a whore”. DE-1 tweets the text “What should we do with this kind of creatures?”, accompanied by the screenshot. As we see, DE-1 calls out the perpetrator, as she has not censored their Twitter name, which in this case is a non-anonymous profile with a name and surname. However, she also, at least rhetorically, seeks advice. The comment is indeed interpreted by her followers as

seeking support or advice. In total, she receives 162 comments on her publication, and the great majority replies with advice, which consists mainly of short replies such as “report”, “report and block”. Some of the users who respond add that they have also reported the profile. A few users minimise the event or encourage her to completely ignore the insult: “Nothing, it's Twitter. Block and that's it. We give too much importance to people who don't deserve it”.

In general, on social media, the previous motto “not to feed the trolls” appears to be less prevalent and there is a general assumption that people have sat back and watched for too long and have not reacted enough to online hate speech (Jubany and Roiha, 2018). Nevertheless, as the analysis of comments from the digital ethnography show, many social media users still believe that ignoring hateful comments is the best way to proceed, as the advice to “ignore” is rather frequently given to women who have been targeted by verbal abuse. However, the action to respond to attackers might not be aimed at changing the opinions of those spreading hate speech, but rather at putting forward other perspectives, breaking the spiral of silence, and providing support to those targeted. In this specific case, in the later comments, some users publish a screenshot where it can be understood that the perpetrator's account does no longer exist. They interpret it as if though the perpetrator has deleted the profile himself, however, it is not clear whether this is the case. In my field diary, I made the following observation:

Most of her followers tell her to block and report this person. When I started following the case, the profile had already disappeared, probably because her followers had also reported the profile to Twitter. So, collective response can have an effect, at least temporarily. Then, of course, nothing will stop this guy from opening another profile... [Field diary, June 6th, 2021]

Given the calling out of this person with their real name, we can recognise this as a strategy of “naming and shaming” (e.g., Jane, 2016) or “digilantism”. “Digilantism” can be defined as engaging in extrajudicial practices that are intended to punish attackers or otherwise bring them to account (Jane, 2017). In this case, calling out the non-anonymous attacker seem to have worked, at least

temporarily: he was “named” and “shamed” in that the attention of wide range of Twitter users was called to witness his behaviour and react in a way that was overwhelming enough for the perpetrator to decide to delete the profile. However, the outcomes of digilantism are unpredictable (Ibid.). Counter-attacks may be produced where some participants have little association with or interest in the original offence (Citron, 2014). In this sense, “crowds of angry feminist digilantes who attack and dox individual men risk mirroring the very behaviour being objected to in the first instance – even if their motivations are different and more defensible” (Jane, 2017: 56).

Further, attacks are often perpetrated by large mobs, which would make it extremely time-consuming for women to track down and confront all the participants of the attack. Publicly speaking out and calling out perpetrators also require targets to reproduce and amplify their trauma, making the targets constantly relive threats on their bodies and lives (Linabary and Batti, 2019). In this regard, there is a risk that digilantism backfires in a way that may disadvantage activists - acts of calling out perpetrators may open up to further or escalated harassment, by the original perpetrator(s) or other attackers.

9.3.2. Representing shared experience

A tweet published by DE-7 in the summer of 2022 illustrates the second type of narrative strategy based on shared experience. In a thread on Twitter, DE-7 explains why she has not been active on the platform during the last few months. She talks about several waves of attacks over a short period of time, and how these attacks took away her previous joy of expressing herself publicly on Twitter. Here, she does not aim to call out any specific profile, but she wants to explain her experience and re-connect with her online community. As a response to the thread, she receives 126 comments, of which the great majority express support either by sending hugs, love or strength; saying that they have missed her; or by expressing that they understand her, some also stating that they have been through similar experiences themselves. In addition to this, she receives a few indifferent comments in the likes of “There's nothing wrong with not tweeting either”, as well as a few misogynist comments such as “If you had dedicated your

time to cleaning instead of tweeting, nothing would have happened to you”. Nevertheless, the majority are warm and supporting comments that express a sense of community where people have each other’s backs. Representing shared experience through publicly speaking out can thus be understood as one way of building community, which I will discuss further in section 9.3.

Often, when sharing personal experiences of violence, women add that they are aware that their experience is the norm for many women online. Recounting personal experiences on social media without naming the aggressors can thus be interpreted as a way of speaking out about the commonality of such experiences and, at the same time, for seeking change for women in online spaces through raising awareness. Drawing attention to the pervasiveness of women’s experiences with online violence using individual stories function as a way of enacting solidarity and creating space for women online. Yet, women who are targeted by intersectional abuse may use their experiences also to interrogate whose stories come to count in debates on gendered harassment online, and to demonstrate how racialised or trans women are often erased or excluded from such discussions (Linabary and Batti, 2019). This is linked to the third narrative strategy discussed below, namely disrupting the hegemonic narrative.

9.3.3. Disrupting the hegemonic narrative

The third narrative strategy emerging from the research - disrupting the hegemonic narrative - includes challenging both hegemonic narratives of the division of the physical and the digital in two separate spheres and dismantling hegemonic narratives of online harassment. In terms of the first, as previously discussed, in the interviews, which constitute a private and anonymous setting, the targets reflected upon the “realness” of the harm caused to them online, which in a sense serves to deconstruct the predominant narrative surrounding digital spaces – that these spaces are clearly limited and can thus be easily abandoned, or that harm which is caused online stays online. Through their stories of what has happened to them online and how this has affected them, women challenge the construction of the internet as a space that is not the “real world”. Nevertheless, defying the hegemonic narrative on digital spaces is not to

any greater extent reflected in the testimonies that I have collected online in a *public* setting.

In terms of disruptions of the hegemonic narrative of online gender-based harassment, however, this strategy does emerge in the public narratives collected online: For example, trans women in some cases “flip” or expand the narrative in their public accounts on social media. To illustrate this, DE-6 expresses in a tweet “Cis people receive dodgy messages like this one day and they're terrified. If I don't receive a handful of threats on any given day, I'm surprised". Here, she illustrates the level of hate that she receives, whilst she at the same time puts it in comparison to that received by any cis person. This could also be analysed as a way of representing shared experience with an intersectional lens, through showing how the experiences that are most commonly showcased in broader debates do not represent all women.

Defying the hegemonic narrative can also be performed in terms of disassembling the narrative of how women are often advised to respond to online abuse, that is, to ignore the abuse, “not to feed the trolls”, or to stay offline. In dismantling the unrealistic arguments that women should counter online harassment by removing themselves from the internet, feminists redirect the responsibility of protection from themselves to platform operators and institutions, and the blame is redirected to structures and perpetrators.

9.3.4. Defying the harassers

Finally, to reflect on the strategy of defying the harassers, as the case of DE-14 illustrates, we can refer to direct statements of opposition to reclaim power within the context of online harassment:

They have created pornographic pictures with my photos, they have threatened to rape me, they have impersonated me, they have harassed me, they have organized hate campaigns and they have insulted me every day since I first started using social media. Against this, the message is clear: Not one step back (DE-14).

DE-14 directly confronts and defies the perpetrators, stating that she is not giving up. It is worth noting that she has deactivated the comments on this tweet.

However, looking at the statistics of the post, the publication has been retweeted more than 4.900 times, and it has been cited more than 600 times.⁸⁶ In the comments to the original tweet, DE-14 has later added a screenshot of one of the tweets that cite her original tweet. The screenshot includes the profile of an apparently older, non-anonymous man, who cites her tweet and comments: “Follow guys, this one will end up either schizophrenic or dead”. Later, DE-14 also adds an additional comment where she defies the perpetrators by saying:

For every misogynistic insult that I receive, or that I see against a fellow feminist, the more I want to put my face and my body out there for them [in Spanish “por ellas”- referring to a feminine them] and for Feminism. You will have us against you. Neither your insults nor your harassment will make us step back (DE-14).

This strategy often involves describing the purpose of online harassment to silence women and responding to that by saying that women will not be silenced (Linabary and Batti, 2019). Along similar lines, DE-9, a trans woman, explains how:

The bullying by TERFS doesn't even bother me. They are pathetic and bitter people. I laugh at them and ignore them. But do you know what is real? The people who write to me saying that I'm an inspiration and a model, who thank me for the strength that I give them.

Here, DE-9 launches a message at those attacking her saying that their attacks do not get to her, and at the same time a message to those who support her and who look to her for support, expressing a sense that community is what matters.

Another trans woman, DE-12, uses a similar strategy, however in video format. In a video directly dedicated to her haters, she uses a sense of humour to tell the people who attack her that they need to do something more important with their time. She goes as far as suggesting different hobbies that they can dedicate their time to instead of putting their energy into following her and expressing hate towards her. She also sends a clear message to her haters that she doesn't read the messages or comments but is lucky enough to be able to pay someone to do it for her and finishes by saying that they have actually created a job. This way,

⁸⁶ Citing is often used by persons who wish to contest a publication.

she speaks out publicly, talking directly to the persons harassing her, but through a humorous strategy, which at the same time may serve to knit her community closer together. The comments on the video are largely positive, except for one transphobic comment, by a user who in a second comment asks her to please block them.

As in any theoretical categorisation, reality is more complex; there are overlaps between narrative strategies and different strategies may be used at the same time. For example, in December 2022, DE-13 publishes a tweet with several screenshots with uncensored profiles of the perpetrators who express explicit sexualised hate against her. A comment in one of the screenshots says “You’re the typical person who never got any dick and not because you’re a lesbian but because you’re bitter. Hahaha, try to contest that if you can”. In the text by DE-13 that accompanies the screenshots, she expresses “We are tired of living with daily threats of sexual, physical, and psychological misogynist violence. No one should have to put up with this for wanting to inhabit social media. We are sick of being the ones who abandon the internet”. In a sense, she uses all the strategies at once, that is, she calls out the perpetrators by not censoring their names; she represents shared experience by using “we” and referring to experiences suffered by many women; she challenges a hegemonic narrative on online harassment – that of women being advised to leave digital spaces to avoid harassment; and she also defies the harassers, by stating that no one should have to put up with this to remain on social media.

Nevertheless, it is important to note that not all feminists targeted by online violence choose to speak out publicly. In the interviews, a couple of the targets express that they do not speak out in public about the misogynist violence they suffer online. In this regard, Elena says that “I haven’t denounced it on social media and so on. I’m not the type who does that (...) No, I’ve never started to denounce it like that... I deal with it with my friends”. Further, Daniela, as exposed above, chooses not to cite aggressive quotes or call them out on Twitter, as she does not want to risk sending her – as she describes them, loyal followers, against the attacker and start an attack against them. Here, she shows clear

awareness about the uncertain outcomes of “digilantism”, as she has decided not to call out abusers or speak publicly to avoid creating counter-attacks. Instead, Daniela chooses to block the attackers, and to talk about the violent events with her friends, in a private setting, whether online or offline. In the next section I shall further discuss this strategy, namely the building of community –along the continuum of the online and offline and across private and public spaces - as a response to misogynist attacks.

To conclude, women’s public narratives serve a crucial active purpose of resistance. Their stories can be interpreted as both individual and collective strategies of “counter-storying”, meant to raise awareness, disrupt hegemonic perspectives of digital spaces and social realities, knit their community closer by recounting shared lived experience and to counter the way that masculinised spaces silence women. However, the emotional labour that this involves – particularly in terms of defying abusers and potentially triggering more hate - may take a toll over time. While some women persist in raising their voices, online abuse may have a chilling effect on others, calling into question what narratives might remain untold or unpublished due to the fear of triggering additional abuse.

9.4. Building community

In a context with limited resources for reporting specific types of online violence, and where existing resources, like social media platforms' reporting mechanisms, fail to adequately protect vulnerable groups, the importance of building community and establishing peer networks for protection becomes paramount. As previously discussed, also reporting and blocking are often rooted in collective action, while targets find emotional support from followers and fellow feminists when publicly addressing attacks or calling out specific profiles.

However, there are also other elements in terms of community-building, which I shall discuss in this section. Whilst these sometimes build on publicly denouncing the problematic of gendered online violence, just as often they take place in private settings. There are several aspects of this community-building: Some are

technical, whilst others are emotional or psychological. Based on the ethnography, I have divided these aspects into three main categories: First, community related to collective resistance; second, community in terms of exchange of technical know-how and support; and third, building community based on shared experiences and emotional support from a feminist perspective. These interlaced and overlapping aspects are explored in depth and discussed in the following sections.

9.4.1. Collective resistance

As examined above, women's public narratives and recounting of their experiences of online violence serve a purpose of resistance, with both individual and collective means of "counter-storying" (Linabary and Batti, 2019). In this section I will highlight examples of community-building among feminists online in terms of resistance mostly from a practical perspective, but I will also look at symbolical aspects.

In terms of practical aspects of direct resistance, that is, reacting directly to attacks and abuse, targets tend to reach out either to their close friends or the wider Twitter community to ask other users to report abusive profiles. This can be achieved either through publicly denouncing the attacks on Twitter, or for those who do not wish to go public about the abuse against them, they may reach out to friends and fellow feminists through private channels. As Daniela expresses: "I ask my friends who also use Twitter 'hey, can you report this profile for me' and that's it". Collective reporting is thus a resistance strategy applied in the immediate relation to attacks.

As exposed previously in this chapter, reports often do not lead to any action by the platforms. However, reporting and other response strategies by followers and fellow feminists may still provide targets with a sense that someone holds their back and that they are part of a community. In this regard, Lucía talks about how in the immediate aftermath of an attack, one of her friends showed support through a kind of "counter-storying" to the hate expressed towards her: "A friend and colleague of mine did like a love-bombing campaign on Twitter and that was

quite good for me. She started saying good things about my work” (Lucía). Although this love-bombing did not reach the same magnitude as the attacks against her, it still calmed her sense of her professional reputation being destroyed by the negative content of the attacks and made her feel supported. At the same time, it can be read as an act of resistance in the shape of a counternarrative.

The digital ethnography reveals illustrations of when collective responses, or what we may refer to as “counterattacks”, do work, and reach an extensive dimension, based on the implication of the community of followers of the person who has been attacked. As previously discussed, this can be described as a type of “digilantism” (cf. Jane, 2017). DE-2, for instance, talks about when she was attacked by right-wing bots on Twitter. When this happened, she managed to obtain support from many persons on Twitter – both among her followers and the followers of her followers, who spread the original tweet where she had denounced the attack on her. This set off a campaign that she refers to as a “counterattack close to lynching”. That is, whilst some feminists under attack choose not to publicly denounce abuse due to their fear that the original attacker might get hurt (as in the case of Daniela), others explicitly use the strategy to publicly denounce to call for collective counter-response and resistance.

Another aspect of community resistance can be observed in the way individual feminists express that they sacrifice themselves for the common good. These narratives tend to involve expressions that they can continue upholding acts of feminist resistance because they are feeling held and backed by a community of fellow feminists. To this end, DE-18 posts several tweets and comments where she expresses this sacrifice and community:

"I endure it because you support me."

"I see myself as a tool for the feminism agenda to progress further."

"I assume that visibility is a lesson, but I don't think we have to assume that violence comes with the pack."

This “sacrifice” is linked to her explicit definition on her social media channels of masculinised online violence as political violence, and as a reaction to the

winnings of feminism. That is, her perception of the importance of defending these winnings makes the sacrifice and resistance online worthwhile, whilst at the same time, the support from her community of fellow feminists makes the experience somewhat easier to endure.

Amara highlights how she is encouraged to resist and stay on social media by focusing on the community of predominantly women who follow and support her:

Another thing that has also helped me is to focus on all the people who transform, who learn, who grow by reading me, because when I decided to return [to Twitter after being offline for several months following attacks against her], I posted a tweet and I was very pleasantly surprised by the comments from people welcoming me, saying that they had missed me, how good it is that you're finally back and I didn't expect that at all (Amara).

Community resistance, thus, has several aspects - some of which overlap with community related to emotional support. Feminists who are under attack use their peer communities for collective reporting and for spreading the public denouncement, but their community also provides them with a sense of meaning in what they are doing and provide them with strength to resist, stay online and keep discussing feminist matters. This is sometimes expressed as sacrificing themselves for the common good, and the community that backs them also provides them with a perception that they are there to take turns in this oftentimes tiring sacrifice.

9.4.2. Exchange of technical know-how, protection strategies and support

Community is also expressed through providing support from a technical point of view and exchanging know-how on how to protect oneself as a (racialised, LGBT+...) woman and a feminist in digital spaces. Sometimes this is articulated by the means of immediate technical support following an attack. As previously exposed, Sara received instant support of friends with similar profiles when she was targeted by a massive attack. These friends took a hands-on approach and took over her social media profiles to help her block users and filter content, so that she could go offline and disconnect and, in this way, did not have to read the abuse:

I have a couple of feminist friends who understood perfectly what was happening and acted very quickly (...) And not just friends of good will, but with the quick strategies of putting all the filters on, quickly removing everything from Instagram, closing everything. In other words, they were very fast because they know a lot about social media and technologies, but I wonder what would have happened if I had been someone without that environment (Sara).

Several participants talk about plans to make this technological support more organised. Sara explains how she has discussed this with other feminists in the same situation:

We have talked about establishing a kind of digital self-defence kit that would allow us to be organized. We had detected that this violence by men is organized, whether in forums like ForoCoches, or in other types of social networks. The fact that they come to us in a mob and the fact that they repeated the same joke to me shows that they knew what they were playing at. So we have talked about the lack of protection that we as women or as feminists have when something happens to us, that we don't know what to do. (...). And we asked ourselves, how could we act at a digital level, how could we be more organized, what could we do when "hey, there is a girl who is being attacked today on the internet, how do we all organize?" And maybe that's it, it could be something in parallel or complementary or in some cases a substitute for going to the police, which could perhaps be more efficient (Sara).

Although this specific loose network of feminists has not come to organise themselves in practical terms yet, there are already existing examples of digital self-defence kits created by digital feminists in Spain, with the aim to provide women with tools to better protect themselves in digital spaces. For example, *Red Autodefensa Digital Feminista* has set up a web page with a collection of resources aimed for feminist digital self-defence.⁸⁷ Some of these resources have been elaborated by organizations active in Spain, such as the "Kit contra las violencias machistas on-line"⁸⁸, produced by the digital feminist organization *DonesTech*. However, other resources stem from transnational or Latin American

⁸⁷ Available at <https://autodefensa.online/recursos.html>, last accessed on April 12, 2024.

⁸⁸ Available at <https://donestech.net/files/kitcontraviolenciasmasclistesonline2018.pdf>, last accessed on April 12, 2024.

organizations, such as the guide "Cibermujeres"⁸⁹. This is important because when discussing digital feminist resources, activists often refer to Latin American feminists and laude their organization in terms of collectively assembling against digital attacks.

Still, whilst there are loosely knit networks such as the above-mentioned *Red Autodefensa Digital Feminista*, as well as a desire among individual feminists to organize collectively against gendered violence online, Spain still lacks a more consolidated feminist network in terms of organising counterattacks or providing support in putting digital protection mechanisms into place. When an attack against a feminist occurs, counterattacks tend to take place in an ad-hoc manner through tweets and re-tweets. This support, as well as support to add filters or block perpetrators, then, depends on having a network of peers that are active on social media and who can actively disseminate the call for help, as Sara refers to:

I think it has been very good for me to have two very close feminist friends, who understood perfectly what was going on and acted very quickly and I wonder what would have happened if I had been a journalist or a professional where it is an anomaly or who is in a more isolated environment and you don't have that quick support that I had from these friends (Sara).

The case of Sara illustrates the crucial role of immediate technical support within feminist and activist networks, where friends with similar profiles rallied to safeguard her from a massive online attack. The hands-on approach shielded her from the abuse. However, at the same time she raises a relevant concern about the accessibility of such support for those without a supportive environment or technical expertise. The absence of organized responses leaves some targeted individuals, especially those in isolated professional environments, vulnerable without prompt and efficient support systems.

9.4.3. Shared experiences and emotional support

By sharing and exchanging personal experiences, women connect their individual stories within the intricate network of online harassment narratives. This serves to underscore the collective aspect of gendered encounters on the internet,

⁸⁹ Available at <https://cyber-women.com/es/#seguir-leyendo>, last accessed on April 12, 2024.

validating the credibility of their lived experiences and challenging the dominant masculinity that prevails in many online spaces. (Linabary and Batti, 2019). Building community through shared experiences and emotional support entails several aspects, such as providing emotional support from a feminist lens, sharing with others who have been through the same and understand what you are going through, or simply laughing it off together. In previous sections, I have explored this linked to women's public narratives. However, building community often also entails sharing experiences and providing emotional support in *private* spaces. From the narratives, a strong sense of feminist community stands out, sometimes referred to as "sisterhood" or "sorority":

I believe that sisterhood is something that can be felt and that is there... It's a blessing, and I think that the internet has precisely facilitated this practice of sorority, it's where we have learned to read other women, learn from the experiences of other women, to see that what was happening to other girls was the same as what was happening to me. It happened to me because I'm a woman, I'm a lesbian. So I do believe that we are also defenceless there, we love and care for each other but that's basically it. We don't have anything else. Sometimes it's not enough, the proof is that women have left Twitter, right? [mentions three feminists who have left Twitter over the last couple of years] (Lucía)

Feminists who are targeted by online gender-based hate speech look for support primarily among peers to talk about what has happened to them online. This exchange of support becomes crucial in a context where many, as Lucía in the quote above, feel that they don't have anything or anyone else to care for them or defend them than their own peer community. Elena, the gamer and feminist, highlights the importance of gathering with other women gamers to provide

a little bit of support, especially for the young girls so that they know that it is the most normal thing in the world, that they [men online] don't say things to them at a personal level, nor do they really know what they are like, but that it happens to all of us, you know? (Elena).

A critical element of collective support among women online is, thus, to share from a feminist perspective, to create a shared understanding of the structural elements of the abuse: That the violence is not really personal; they suffer violence that is targeted at the group of "women" or "feminists" and they

themselves as individuals are to a high degree symbolic in this regard. By talking about their lived experiences together, whether online or offline, they notice the similarities in the expressions used against them, and in the modus operandi of the attacks. This makes it easier to understand the collective element and the functioning of hate speech; that it's not really an attack against them as an individual – although it is expressed as such – but an attack against the group of women in general, or often in these specific cases, against feminists and their struggle for equality.

Some participants also point to the importance of gathering with women with similar experiences to “laugh it off”:

Between the girls and so on, most of all with the youngest ones, to put together a group, a support group and when they insult you, and not for the dramatic effects but rather put the insults there to laugh like ‘look what they called me today hahaha’ or ‘this one is new’, you know? To downplay it a bit, most of all so that the girls aren't scared, so they're not afraid and suffer (Elena).

This leads to examining humour as a strategy that may work well to strengthen one's own community as it helps users to better cope with hate (Jubany and Roiha, 2018). The use of humour, however, is also complex as it may reinforce divides between groups and trigger further hate when applied in public forums. This is illustrated in the case of DE-2, who tweets a humorous post about how men are so late in getting into issues such as mental health or self-care. Whilst she receives many replies of recognition from women, this in turn also triggers one of the largest attacks against her, not only in the immediate comments to the tweet, but also through other platforms. On the one hand, she gathered and united fellow feminists; however, reactions from anti-feminist men were aggravated and unleashed both attacks and counterattacks.

Some women also describe immediate emotional support by other feminists in direct relation to attacks. This support is expressed along the continuum of the online and the offline. For example, in addition to the technical and emotional support by her close feminist friends, Sara recounts how she, the same day as the

attack started, talked on the phone with another feminist who has been targeted by attacks on several occasions:

That crazy morning, I spoke with her on the phone because people contacted me to say 'don't worry' and so on and it was also good to talk to a person who unfortunately is very experienced in this, who also helped me to relativise everything and told me 'go for a walk, you will see that the world goes on'. And it's true that there is sometimes this feeling of total duality that on the internet everyone is against me but then you go out into the street, it seems that nothing is happening. So that calmed me down... (Sara).

Receiving support from someone who has gone through the same experience may help the targeted person regain perspective. In this case, the advice to disconnect for a bit and go outside helped Sara calm down and see that not everyone was against her, although it may have felt like that when she was caught up in the social media fury. Although there is a clear continuity of the online and the offline, and online harms affect targets in embodied ways and at several levels, it is also possible to disconnect and recompose before going back to confront the waves of hate in digital spaces, at least momentarily. Peer support based on shared experiences may help remind those targeted of this.

Community-building can, thus, be understood as taking place along the continuum of the online and the offline. Friends who understand the affordances of social media, and how the specific type of violence against women activists or feminist work, are perceived as particularly appreciated for emotional support. However, sharing of experiences between feminists sometimes also gives rise to comparison, relativisation, and an elevated feeling of normalisation. In the case of Daneila, she talks about what had happened to her and then adds that it is nothing compared to other feminists with bigger profiles. It is this that makes Daniela reflect on how this violence not only reaches the biggest feminist profiles but also a profile like herself, which she describes as

not famous or anything, I'm a girl who sometimes tweets and I've moderated some events, that is, I'm not [mentions the names of two feminists], nothing like that. That people are getting up to these things on social media kind of scares me (Daniela).

As a “smaller” profile with less followers and a less established position in informal feminist networks, she may receive less support when attacked. Further, in her case, she does not tend to publicly denounce the attacks, which gives less opportunity for other feminist profiles to provide immediate support. Publicly denouncing violence, as I have discussed previously, not only increases awareness and provides followers with the opportunity to report the aggressor but also invites support from a broader network beyond those in the immediate circle, who may be notified through a message or phone call.

Sharing with one’s family

In contrast to exchanging experiences and support among peers, not everyone chooses to share what happens to them online with their families. Whilst Daniela feels that she is supported by her friends and her partner, she has chosen not to share her online experiences with her family:

I don't tell my parents... My parents know I'm a public person, but I don't think they understand the scope of it. (...) What I do is talk to my friends, with my partner, with people around me. Not with my family. I don't tell them, because I know they're going to worry in the end. If I say to my mother now like 'Mom, I made a tweet' -she won't understand because she doesn't have Twitter and she's going to freak out - 'I made a tweet about gender roles and I have a man from Madrid saying that he hopes they rape me on the street'. Of course she's going to freak out (Daniela)

Whereas Daniela explicitly seeks support among her friends, her parents are not aware that she has suffered threats and violent comments on social media. From her perspective, there is clearly a generational gap. Her perception is that her parents, who do not use Twitter, would not understand the online dynamics, and would worry too much. This means that her parents are absent from a very important dimension of her life. The narrative of Daniela also brings back into focus the issue of normalisation. Daniela understands that her mother would “freak out” if she knew that anonymous men online are hoping for her daughter to get raped. Still, for Daniela herself, this is part of her day-to-day online existence and as such rather normalised.

Sara, on the other hand, explains that she talked to her family and received support in the case of the massive attack against her, because the magnitude was so extensive that it reached well beyond Twitter circles:

In that case, they did know about it, because precisely what scared me the most was that it had even reached those people, because that controversy has already crossed all possible niches. When people interceded, political figures, [a political right-wing party] and such, it became very massive and I started to get scared when my friends who don't have Twitter knew about it 'wow, this has already got all the way there...'. But my family understood what was happening because in this case it was very very... it was quite massive (Sara).

This reinforces the importance of a perspective of interconnectedness, not only of support networks along the continuum of the online and the offline, but also of understanding how information moves through this continuum. Increasingly, traditional media in Spain, such as television news programmes, showcase and discuss events that have taken place online, particularly viral pieces of content. Clearly, online content increasingly reaches those who do not consume it in its original outlet, but also generations and population sectors that are more reliant on "traditional" media. What happens online does not stay online; nor does it affect only that sphere. Likewise, support can be expressed and needed also in offline physical spaces, particularly considering that the violence that targets these women follows them throughout, by the means of mobile phone apps.

Seeking emotional support by professionals

Emotional support may also come from outside the peer community or the family, from professionals. This aspect did not emerge to any high extent in the interviews, as very few of the women had tried to seek any professional support. However, Lucía had experience of going to therapy and reflected on the importance that professionals who provide psychological or emotional support are also prepared to address experiences in digital spaces:

In the end, there aren't many concrete tools... Because if you go to your therapist, maybe it's a person who is very valid, but maybe not in the Twitter world. It's not that it seems silly to them, but they are not able to capture the dimension of when a person is on Twitter how much it seems

that everyone is Twitter and it's also a job to see that not everyone is Twitter, that many people live oblivious to what happens on Twitter or Instagram or whatever and live quietly and it's not so important. But obviously when you are there, when your job consists of exposing yourself, it's also very complicated (Lucía).

Both in terms of technical support, as I have exposed in the previous section, and in terms of psychological and emotional support, there is a strongly felt need for more organised support networks. In an ideal world, as Lucía expresses above, psychological support would be provided by the public health services. Targets perceive that what they experience is not taken seriously enough by institutions, that resources are lacking everywhere and that, therefore, they can only dream of this type of support.

There is a wide sense of abandonment by institutions and digital platforms, and the view is persistent that the only support available is that provided by peers who have lived through similar situations and who can provide technical support, emotional support and a collective sense of resistance, community, and in-depth comprehension. This should also be highlighted in a positive manner: there is a strong sense of community among feminists, in digital spaces – and outside of these- particularly among those who have been targeted by online violence and who have been open about this. Sharing means understanding, a shared burden, and laughing it off, making jokes about the attackers. Without this feeling of a shared community, it is likely that the impact of online violence is experienced even harder.

Whilst so far in this chapter I have explored institutional and professional support briefly in terms of psychological assistance, in the next section, I shall investigate the experiences of seeking institutional support, in terms of interactions with the justice system, through reporting to the police.

9.5. Interacting with the justice system

There is ample evidence that both hate crimes and gender-based violence are much underreported to the justice system; for example, according to the Survey Report on hate crimes of the Spanish Ministry of the Interior (2023), 89.24% of

hate crime victims had not reported to the police. Analysed by the aggravating circumstances, 87% of the victims on grounds of sexual orientation or gender identity and 90% of victims of a hate crime based on sex or gender did not report this to the police. In the digital context, violent expressions are highly normalised and reporting to the police is even more uncommon, with very scarce official statistics. In this section, I will, on the one hand, discuss experiences of reporting to the police and, on the other hand, zoom in on the reasons that women give for not seeking support from the formal justice system when they are victimised by online violence.

9.5.1. Experiences of reporting to the police

As a preface, it is important to note that the data I have collected during the research for this thesis reveals a clear scarcity of experiences from women who have reported online violence to the police. Only two women, amongst the ones interviewed, had attempted to make use of the criminal justice system as a means of responding to online gendered hate, through reporting to the police. This reflects findings from previous qualitative research such as that by Jo Smith (2019), Stina Eckert (2017) and Karla Mantilla (2015). As Smith (2019: 290) concludes, “where such attempts had been made, the experience had been disheartening and had discouraged them from adopting this form of response in the future”. In addition to this, in the digital ethnography, I have found information on reporting to the police in two of the cases. This information comes from sources beyond Twitter, that is, I have participated in online talks and have observed media appearances by these persons, where they have talked about their experiences in this regard. Nevertheless, in the data collected directly from Twitter, I have not collected any data or observations of targets explicitly saying that they have reported or will report to the police. Some of these women might have reported, however, without stating this publicly on their open social media profiles. Still, police reporting is quite often mentioned by Twitter users in comments to tweets where women publicly denounce attacks, as advice on how the women should proceed, particularly in the most severe of the cases.

In terms of experiences of women who have reported to the police, I will first briefly describe these cases and then discuss them: Elena, the feminist gamer, had experienced attacks on Discord, directed towards her as a lesbian woman:

Once on Discord on the server, playing, they became super aggressive, with all kinds of insults towards my sexual orientation. And I did report them to Discord, and I went to report them to the police, but even the kid himself laughed at me and told me 'report, report, hahaha', and it's true because nothing happens, and in fact nothing happened. I went with the image captures and such but it's not like they gave me many answers. In other words, the police really don't do anything about these things (Elena).

Amara, an anti-racist activist, had on two different occasions been directly targeted by a mix of racist and misogynist threats. On previous occasions, the men attacking her had seemed to be in places far away from her hometown and because of this, she did not believe there was any real threat to her physical integrity. During these new attacks, however, she realised that the perpetrator lived in the same autonomous community as she did and that he may know how to find her, as she had been quite open about where she lived. That was the moment when she got afraid and decided to report to the police:

I reported. For both of them I went to the police. The first time they told me that they would pass it on to the Computer Crimes Department. The second time, it was a threat, and even then, they did nothing about it. They told me they would call within 15 days and let me know. Nobody ever called. Nobody ever called. That also makes you feel helpless. But I am still very sure that if it happens again, I will continue to report it. I'm not going to stop reporting. But I understand that there are people who decide not to report because they are not given any importance (Amara).

As evidenced above, the outcome was not what she expected. Although the police filed a report, afterwards she did not hear back from them and as far as she knows, no investigation was opened. Nevertheless, whilst clearly disappointed, Amara claims that she would report again. That is, this experience has not completely undermined her trust in law enforcement agencies.

In this same vein, DE-18, a feminist journalist and activist, had received a mix of harassment online: in addition to direct verbal abuse on social media, someone published her phone number on Twitter; she was added to different WhatsApp

groups, and they placed an advert in her name on an adult contact page so that she received sexual photos from unknown men. The harassment also included offline elements: Someone painted insults outside of her workplace. She reported all these events to the police, and explicitly asked the police to join all the attacks against her in one report, as the different attacks were clearly linked, and based on her role as a journalist and feminist. However, during an online seminar that I participated in as part of the digital ethnography, she explains that she has the feeling that the police do not apply any kind of gender perspective or look at these crimes as systematic and connected: "It's not just Twitter anymore... I don't even know how many crimes I've reported. The police have no idea about how to manage this type of report. I've done all my homework and I'm almost like the handbook victim" (DE-18). Despite having consulted extensively with a lawyer specialised on gender and hate crime, and having reported with the assistance of this lawyer, with all the evidence presented, she explains that the reports have not led to any action. On the one hand, the police do not connect all these crimes, on the other hand, they do not address them as gender-based: "They ask if it's my partner doing this, and when I say no, they flip the page of gender violence reporting and there's no gender perspective or anything that connects all these crimes against me" (DE-18). This is a clear illustration of the discussion about the omission of a gender perspective when the crime is not committed by the woman's partner or ex-partner, in accordance with the Spanish legal framework.⁹⁰ Furthermore, in most of the cases, the attacks against her have been perpetrated by several men. Some of these men are not even aware that they are involved in an attack, such as those who sent her sexual images when her phone number had been posted on an adult contact page. This adds an extra layer of complexity to reporting and addressing these crimes within the criminal justice system.

Further, DE-26 had received a death threat on Twitter together with a photo of a gun, with a text that read: "We're already tired of you, whore, the only solution to your claims is death. We're going to rape your corpse, bitch". Whilst she had

⁹⁰ See Chapter 5 for discussion of this topic.

received numerous messages before this, the gravity of the threat made her decide to report it to the police. With reporting she also aimed to increase the visibility of this type of crime and to go through the process to be able to later inform peers in the same situation on how to proceed. However, like the experiences described above, the report was futile, as she explains that she did not receive any updates on the investigation.

Looking at all the experiences that led these women to report, these instances were attacks or harassment that had the following factors in common: First, in three of the cases there was an intersection between several factors in the motivation of the attacks, such as gender and LGBT-phobia or gender and racism. Second, the women perceived a high level of gravity, making them fear that the online attacks posed a direct physical threat to them, and linked reporting to the need to protect their physical safety. This reveals how reporting is highly coupled with feelings that there is a threat to one's physical integrity or a threat to one's identity. Regarding the latter, LGBT-phobia or racism can be read as factors that are more acknowledged as grounds for reporting, in contrast to misogyny or threats towards a woman because she expresses feminist opinions. This shows that hate crimes and hate speech based on gender are less perceived as such among women feminists and activists. Speech that includes racism or LGBT-phobia, on the other hand, is to a higher extent recognised as "hate speech" or "reportable", even when expressed online. This may be linked to the lack of engagement of the wider feminist movement in placing gender-based violence on the hate crime spectrum. In this regard, the law in a way – even if only in theory, as we have seen in Chapters 5 and 6 – is one step ahead of the civil society movement and society at large.

As exposed in Chapter 5, gender and sex have been included as aggravating circumstances in the criminal code. However, this is only applied on speech against the collective of women, and not against individuals, at least in relation to the online context. This is despite the fact that this speech may have the same function of ridiculing and dehumanising the group of women (feminists) through violent speech against individuals belonging to this group. That is, by attacking a

symbolic individual, the attack targets the collective. Focusing on the group of feminists, the attacks could also be categorised as hate speech based on ideology, which is also recognised as an aggravating circumstance in Spanish criminal law. Nevertheless, as the analysis of Chapter 5 revealed, the feminist movement has not adhered their cause to the broader movement of struggle against hate crimes and hate speech. On the other hand, from a legal perspective in Spain, gender-based violence is limited to violence perpetrated by a partner or ex-partner. This leaves women who are attacked online by someone who is not their partner or ex-partner, or by a large mob, with few resources in terms of legal tools.

In addition, common in these cases is the lack of information or updates by the police. Victimised women have been left without answers, or a follow up, leaving them with a sense of impunity, despite reporting. This reinforces their feeling that these types of crimes are not taken seriously by the police and may contribute to a lack of trust in the law enforcement and judicial authorities. A crucial aspect for women reporting misogynist hate crimes or hate incidents is the significance of being heard, receiving reassurance from the police, and sensing that their concerns are treated seriously (cf. Mullaney and Trickett, 2018). Even when these women understand that the police may have limited possibilities to locate and take action towards the perpetrator(s), it remains crucial for those who report crimes or incidents to feel heard, reassured, and taken seriously by law enforcement. In the next section I shall zoom in on the narratives of the women who did not to report to the police.

9.5.2. Why do women not report online violence?

As the empirical research that forms the foundation of this thesis reveals, reporting of online violence is not common, even in cases where the abuse includes direct threats towards an individual. As exposed above, this is also supported by previous research and official statistics. To explore this further I focused part of the interviews on discovering participants' reasons not to report, with revealing explanations.

The case of youngest person interviewed (Daniela, 22 years), serves as an illustration of several of the reasons not to report. Daniela had received a mix of verbal abuse towards her, including direct threats. When asked whether she had reported any of the online abuse to the police, she said that she had not even considered it. When asked why she thinks that she did not consider it, she explains that

I mean, I'm not that well known, and my threats have not come to the point of intimidating me in the sense of making me change my routine. In other words, they are anonymous profiles that don't know who I am, and that's it. Nothing strange has happened to me in that sense (Daniela).

Along similar lines, Lucía explains that she has not felt that any of the attacks against her have been serious enough to report to the police because:

Luckily, they haven't threatened to kill me, or control me, or these things that have happened to other colleagues (...) It's more of an attack on my self-esteem and my personal integrity, rather than my physical integrity, so luckily, I haven't had to report it. Because you can't report this - I wish you could report that this person has sent me their followers, they have so many followers, they are aware of their reach and this is a responsibility... There must be consequences for the person who incites hatred against a person. I wish it could be reported, although I don't know if I would, because we all know what happens with complaints. I wish I lived in a system where this could happen because of course there is obviously some damage (Lucía).

In both of their stories, the gravity of the abuse and/or threats is minimised, despite the impact of the verbal attacks on their lives. Daniela does not believe that the threats would become a reality, because the attackers do not seem to know her outside of social media. That is, she does not feel that she would need protection from a physical threat. Lucía, on the other hand, refers to not having received any direct threat to her physical integrity. She has been deeply impacted by verbal abuse by many people; however, there is nothing palpable enough (such as a death threat or threat of rape) that she feels that she could report to the police.

The hierarchy of violence resurfaces – a threat that is seemingly “empty” is not perceived as a violent event, unless it carries with it a threat to the physical

integrity of the person targeted. Online abuse is not perceived as possible to report to the justice system unless it includes explicit threats. Nevertheless, as I have exposed in Chapter 5, there are other crime typologies, such as stalking, unfair vexation or crime against moral integrity, that could be applied but that those who are targeted by this violence are largely unaware of.

There seems to be a widespread perception that the only online violence that can be reported to the police is that perpetrated by a partner or ex-partner (that is, within the framework of gender violence) or direct, explicit, and serious threats by strangers. In addition, the targets interviewed feel that this threat needs to pose a risk of being materialised, that is, that there needs to be an immediate risk to their physical integrity. This highlights the impact of the binary categorisation of the online and the offline, in terms of perceptions of violence and subsequent actions. Categorisation does not only serve to perceive, explain, and understand reality, but it also provides guidance for actions. In this instance, relating certain behaviours as pertaining to the “online”, justifies the lack of actions of response.

Furthermore, Daniela also gives other reasons for not reporting, perhaps more significant to her lack of even considering this as an option:

And besides, I don't know if I would report it to the police because I have the feeling that the police would basically laugh in my face if I go there... I mean, I have screenshots of a person saying 'I hope they rape you' and 'I see you on the street, and I'll do something to you.' I could report, but I think they're going to laugh at me. Like testimonies that you see on social media, where women are going to report and they are laughed at (Daniela).

Her explanation speaks of a lack of trust in the police. A lack of trust in authorities and particularly in law enforcement forces has been widely acknowledged by previous research as a reason for hate crime victims not to report (see e.g., Perry, 2001; Awan and Zempi, 2015; Langarita et al, 2018; Paterson et al, 2018; Domínguez Ruiz et al, 2023). For example, collectives targeted by LGBT-phobia or racism may have seen examples of these discriminatory structures expressed by representatives of law enforcement

agencies themselves, and the same is true for misogyny. Daniela believes that the police would not take these events seriously. This is in part based on accounts that she has read on social media of other women who have reported instances of digital violence and the reactions they have received from the police. In this regard, her argument is also revealing in relation to collective responses to and public narratives of online violence. Women who publicly share their accounts of interactions with the police often do this to raise awareness of how law enforcement agencies address (or fail to address) gendered online violence. However, they may not be aware of this side effect: That other women in similar situations may be scared off from reporting to the police. This risk perception may create a sort of catch 22, where the police receive very few reports, do not dedicate resources to this or do not receive training on how to respond to this type of violence, and the few women who do report may be met with attitudes of not taking the violence seriously (or at least made to feel this way). Whilst the research participants of this thesis who did report were not met with laughter or explicit indifference by the police officers they encountered, the lack of action and feedback after the report made them perceive the law enforcement agencies' priority of addressing online violence as very low.

In some of the narratives, not having reported to the police is also linked to normalisation and the day-to-day experiences of online violence. In the same sense that Julia, as we have seen above, does not report online hate speech through the platform reporting tools, she has not considered reporting it to the police either. It is to be expected that if a person is targeted by hate messages on a day-to-day basis, these messages would need to be something very extraordinary for the person to go to the police, or to denounce through other channels. In this sense, the very normalisation of online violence may hinder women from even realising that this is something that could be addressed by the formal justice system. This link between normalisation or repeated victimisation and underreporting has been highlighted in relation to wider debates on the underreporting of hate crimes, particularly among LGBT+ people (see e.g., Chakraborti and Hardy, 2015; Paterson et al, 2018). Normalisation, in this regard,

can also be linked to an internalisation of discriminatory structures, through the functions of symbolic violence. When discrimination and violence is so present throughout different life spheres, this is viewed even by those victimised as something they just have to live with and not worthy of reporting. In combination with a lack of trust in law enforcement agencies and the feeling that these will not take the abuse seriously, reporting is often not even perceived as an option.

Finally, in the explanation of Sara, we find traces of another reason not to report, linked to managing the online context. Sara acknowledges that some of the messages targeting her could have been reported to the police. However, these had already been deleted by the friends who helped her filter messages so that she did not have to read them:

Many of the messages that perhaps would have constituted a crime because there were some messages in which they threatened me with 'we are going to kill you', there were a couple or three that could have constituted a crime. But what happened is that my friends deleted everything, because in the midst of that absolute influx of massive messages, they needed to make a filter, and then I never managed to report anything (...) The truth is that at that moment I didn't see reporting as an exit that was going to be useful to me for anything. As I already told you, what I considered could be most comforting was turning off the technologies, being in my house letting my loved ones take care of me, and at the same time there were some people who were doing a sweep so that I wouldn't see the atrocities. I did not see reporting as an efficient solution to this crisis I was going through (Sara).

In this sense, an immediate strategy of response, which the targeted person felt was what she needed in that moment, ended up hindering another strategy at a later stage. If she had eventually wanted to report, the evidence to support the report was already deleted. Yet, in that moment, Sara did not perceive reporting as strategy that would be efficient in any manner. Looking back now, a couple of years later, she defines some of the messages as crimes that could have been reported and managed by the justice system. Still, in that moment, she had other needs that she prioritised, and she did not perceive institutional support as a path that would help her meet those needs.

As Powell and Henry (2017: 273) highlight, “‘justice’ is given a meaning through the opportunity to tell one’s account of victimisation in a public hearing, to have that account validated by a jury of one’s peers and acted upon in a way that seeks to hold the perpetrator accountable, correct the harm and/or prevent its recurrence”. From the accounts of those who did report, although not explicitly stated, this may in part have been what they hoped for. Instead, the common feeling that victims have whether they report or not is, on the one hand, that reporting to the police does not lead anywhere and, on the other hand, that the police does not take this type of gendered online violence seriously. This means that those targeted are left with a feeling of complete impunity, and the perception that the only measures available are those of self-protection and peer support.

9.6. Conclusion: Digital protection, community and impunity

This chapter has explored, revealed, and discussed the diverse range of responses (or lack of responses) to online gendered hate speech by targets and other actors. This has exposed a spectrum of reactions from digital self-protection measures to public narratives, community-building, and seeking justice through formal and informal channels. The broad normalisation of digital self-protection measures underscores the acknowledgment of the prevalence of online violence against women with a public (feminist, activist) profile. These measures provide individuals with a sense of control and immediate mitigation during or after attacks. At the same time, disillusionment with the response by the social media platforms to reports has led to a lack of trust in these platform-specific measures among targets, emphasising the need for more transparent and inclusive reporting tools.

Further, publicly speaking out, whether through calling out abusers or disrupting hegemonic narratives, emerges as an act of active resistance, which serves both to counter silencing mechanisms and to foster shared experiences. Linked to this, community-building becomes a crucial aspect in the absence of effective institutional and platform-based protection, emphasising the role of peer support. The reluctance to report incidents to the police, due to perceived

inefficacy and a lack of faith in the criminal justice system, highlights the need for systemic changes in legal frameworks and law enforcement practices. In light of the perceived impunity of digital violence and the disproportionate burden on targets for self-protection, there is a call for systemic changes that shift responsibility from victims to perpetrators, and from individual to collective measures. The narratives of abandonment by institutions and platforms and the continuum of online and offline experiences call for interconnected support networks that extend beyond digital spaces. Deprived of institutional or platform support, targets may resort to informal methods for seeking justice, such as engaging in digilantism, which involves speaking out or publicly identifying perpetrators on social media. These approaches are closely tied to discussions on "viral justice" (see e.g., Antoniadis, 2012; Wood et al, 2019), referring to instances where victims share their own evidence of abuse on social media, leading to the content going viral. To some degree, the justice system is not perceived as a crucial, credible, or efficient means to achieve a sense that "justice has been done". Instead, it is often perceived as ineffective or not trustworthy, especially when it comes to addressing online harm.

10. Recapitulation and concluding observations

10.1. Recapitulation

The conceptual and empirical analysis presented throughout this thesis has highlighted the extent to which digital spaces offer easily accessible paths for opponents to reach their targets and express ridiculing discourses and violent, dehumanising speech, freely and with impunity. Through this analysis, the thesis has sought to illuminate the phenomenon of gender-based online hate speech and its intersectional dynamics, focusing specifically on the online violence directed at feminists in Spain. In this violence, hate against different collectives (women, feminists, LGBT+ people, racialised persons, Muslims...) is expressed through individualised attacks. Through in-depth interviews with feminists and other activists targeted by online violence, as well as through a digital ethnography, I have unravelled the lived experiences of a continuous exposition to violence. The firsthand experiences of women subjected to online attacks has served as a poignant testament to the gendered structures and dynamics that shape the use of technologies. Interviews with professionals, on the other hand, have served to shed light on how digital violence is perceived from within institutions and civil society organizations, as well as from a legal perspective.

The thesis delves into how technology-facilitated verbal harms can materialise as embodied harms, thus, impacting not only on digital communication, but also on non-digital bodies, collectives, and interactions (Powell and Henry, 2017). The

prevailing paradigm of "radical transparency" (Kirkpatrick, 2011) on social media platforms promotes the disclosure of one's "real" identity online, involving the revelation of gender, the use of photos, and real names. This trend contributes to an increasing fusion of our online and offline lives, emphasising the central role of our bodies in shaping our online personas. This underscores the significance of gender identity and other identity traits in influencing our online interactions, viewed through both a gendered and racialised lens. The core of this thesis offers a perspective on the intricate dynamics between technology and the lived experiences of individuals, highlighting the multifaceted nature of online harms, whilst emphasising the profound impact of these harms on both individuals and collectives.

Technofeminism views the relationship between gender and technology as mutually shaping, where technology influences gender relations and is influenced by them in return (Wajcman, 2002). Both technology and gender can be understood as products of a moving relational process, emerging from collective and individual acts of interpretation. Gender shapes and is shaped by technologies; in this thesis, the technologies are made up of the digital context of the internet and social media platforms. Nevertheless, gender cannot be analysed in a vacuum, but needs to be considered as interlaced with other factors such as racialisation, sexual orientation, or gender identity. These intersections bring into play specific experiences and distinct discriminatory or violent episodes. In this regard, the examination of violence as a factor in the reproduction of gender norms is crucial, elucidating the interdependence also between violence and gender (Segato, 2003). Consequently, digital interactions necessitate interpretation within a broader socio-political context. This broader view enables the analysis to transcend simplistic notions of technologies either driving changes in gender-based harms or serving as mere tools in the hands of individuals. Instead, it propels us towards a more comprehensive understanding of technology and gender relations as mutually shaping digital interactions, influencing experiences of online harms, including hate speech.

10.2. Along the continuum of violence, spaces, discourses, and self-protection mechanisms

As both the approach and the results of this thesis have highlighted, the concept of a continuum is fundamental for analysing gendered violence. In the first instance, the methodological approach has recognised this continuum and the importance of the ethnographic approach of combining observations of interactions and talking to people, across digital and non-digital spaces. Further, conceptualising the online and the offline as a continuum facilitates an understanding of how different life spheres are affected by violence expressed online and how this violence cannot only be analysed as pertaining to digital spaces. As I have exposed throughout this thesis, women are being attacked online not just because they express their opinions in *online* forums and platforms, but for voicing opinions in any public media. Online attacks do not only impact on women's willingness to voice opinions on *social media* but may have an impact on women's participation in any public debate or space that reaches a mainstream audience outside of a circle of persons with similar views. Further, in some cases, violence is also manifested in offline spaces, and in the narratives of the targets, violence expressed online is perceived as closely linked to that exhibited in offline spaces. These experiences expose the importance of considering online and offline realms as interconnected and overlapping, with continuous mutual impacts. Conclusively, as argued throughout, online violence must be understood as not confined to the digital world but as embodied through the very physicality of those who experience it, resulting in a profound and far-reaching impact that extends into various facets of life.

Linked to the arguments above, the experiences narrated by the targets speak of a continuum of gendered violence (Kelly, 1987). When discussing their experiences, explicit verbal violence that targets them as (racialised/lesbian/trans) feminist women is not clearly separated from other experiences such as unsolicited sexualised contacts that might not be perceived as violent actions from the outside, but that may still be *experienced* as violence.

In this sense, gendered or sexual violence, whether in the shape of gender-based hate speech or a physical transgression can be understood as both a manifestation of and a mechanism for sustaining gendered power relations.

Further, the content and discourses of the very attacks are distributed along a continuum ranging from more subtle remarks linked to traditional gender roles, to explicit rape or death threats. Nevertheless, throughout this continuum, the function of the content is the same: to put women in their place, to keep them on guard, to derail debates and keep feminists occupied with self-protection measures and responses to attacks. Further, all these experiences point towards the need to understand this type of violence as a process rather than a one-off time-constrained event (Perry, 2001).

10.3. Intersectional dynamics

Women who are racialised, Muslim, or part of the LGBT+ community often face disproportionately harsh attacks, where threats of sexual violence are compounded by extreme racism, LGBTphobia, or Islamophobia. An anthropological approach that incorporates lived experiences brings precisely such nuanced perspectives to the forefront. The ethnographic evidence of this thesis uncovers the ways in which multiple identity factors intersect and exacerbate online abuse and thus shape the experiences and impacts of the violence. Participants, reflecting on their experiences of gendered abuse, note that their racialised or perceived sexual identities are connoted in the language used, expanding the focus beyond misogyny alone. Racialised activists who narrate their experiences, reveal how misogyny and racism are often expressed interwoven in the same attacks. Similarly, LGBT+ activists and trans women note how their sexual orientation or gender identity intensify the aggression they face online. The narratives reflect the compounded impact of identity in shaping the dynamics of online abuse, underscoring the importance of an intersectional lens in addressing the multifaceted challenges confronted by targeted individuals.

The intersectional lens can also be applied to the real or perceived positions of the profiles in this study, particularly those with a public voice and profile, thus

delving into the complexity of power relations within this context. The online misogyny observed does not uniformly affect all women but is instead directed at specific groups, such as feminist women, those who speak out, or those who challenge patriarchal and heteronormative structures. The selectivity of this hate suggests that it is expressed towards women who deviate from expected norms rather than being indiscriminately directed at all women. This means that the findings of this thesis surrounding experiences of receiving hate cannot be directly generalised for the broader population of women, which also underscores the significance of adopting an intersectional approach that considers not only factors such as gender, racialisation, age, and religion, but also individuals' political beliefs, activism levels, and professional backgrounds.

The complexity arising from the intersectionality of prejudice adds nuance to debates surrounding whether incidents of misogyny should be categorised as hate crimes, given the multidimensional nature of perpetrators' prejudicial motivations. Throughout history, the legal system has grappled with addressing the intersectionality inherent in certain manifestations of hate and abuse, as argued by Kimberlé Crenshaw (1991). As Ruth Lewis et al (2019: 133) highlight, “Racist, homophobic or disablist hatred is exacerbated by combination with misogyny but in ways that are unpredictable, mediated by context”. Placing a greater emphasis on categorising and addressing this violence as hate speech may provide not only a conceptual framework for better recognising intersectional experiences of violence, but also a wider array of practical tools.

10.4. Embodied and collective harm

The comprehensive analysis conducted in this doctoral thesis reveals that online abuse significantly influences women's behaviour, leading to them “making changes to the places they go, the debates they join, the material they post” (Jane, 2017b: 45). The narratives uncovered illustrate how the detrimental effects of online experiences extend beyond the digital realm, manifesting as genuine harm encompassing emotional, mental, and even physical or psychosomatic dimensions. The recurrence of hate messages, with their potential for internalisation, functions as a linguistic injury, where the relentless repetition of

certain words and messages contributes to a sense of disorientation. Despite recognising the intended ridicule and intimidation intellectually, targets may succumb to the repeated messages, perceiving them as a distorted version of truth, thereby experiencing a dehumanising effect, particularly in public spaces. As we have seen, these more subtle and indirect forms of hate speech play a comparable role to overt hate speech, persistently permeating the public sphere whilst also carrying out public shaming of individuals. This underscores the dual function of hate speech (Waldron, 2012), dividing by dehumanising targets and uniting perpetrators through their shared messages.

As a result, women subjected to sustained online abuse often find themselves reevaluating their digital conduct and interactions, navigating a nuanced spectrum between self-restraint and self-preservation. These experiences illustrate that the suppression of feminist voices targeted by gender-based online hate speech is not always driven solely by explicit fear or self-censorship. A prevailing sense of tiredness of being on the receiving end of abuse, and thus a waning enthusiasm for online engagement, can also impede their involvement. Faced with limited choices, some women opt to withdraw by making their profiles private, or confining discussions on feminist issues to smaller circles, bearing the consequence of diminished visibility or, in some instances, opting to abandon the platform altogether.

The violence these women encounter is often interpreted by targets themselves as a response to a feminist message that challenges norms. Through this we realise how while negative impacts may diminish with time, a sense of empowerment and increased motivation to engage in political action may potentially grow in the long run. This empowering effect might be more pronounced within a specific subset of individuals who identify as feminists. Nevertheless, this potentially galvanising effect coexists with changed behaviours and limited engagement in certain spaces and debates.

Further, the comprehensive collective effects of online abuse are difficult to gauge fully; however, based on the significant experiences of a set of women, the impacts of gender-based online hate speech are likely to yield far-reaching

implications for the participation of voices in the public debate. Online attacks are often brushed off with the idea that those who are directly targeted can just stay offline to avoid attacks. Whilst exiting a platform can indeed be conceived of as a kind of active non-participation (Casemajor et al, 2015), the stakes of opting out of a platform are high for many people. This is particularly for individuals whose professions, personal lives, and advocacy contexts are closely linked to these spaces. Considering the communication between online and offline spaces, separating these realms, and staying offline is virtually impossible, and exiting digital spaces is thus a complex decision. Publicised cases of women who leave Twitter, such as Barcelona's ex-mayor, serve as a stark reminder that violent episodes are not isolated incidents but are systemic, with a strong impact even on those women who have institutional support. Women who must maintain open, public online accounts due to professional obligations, often as self-employed professionals, represent the profiles most susceptible to being targeted by anti-feminists. They find themselves in the forefront, risking not only the impact on online interactions at individual and collective levels but also facing significant embodied consequences. The metaphorical battle underscores the notion that, despite the weariness, feminist voices must persist in public spaces, countering the surge of gender-based hate speech. The perceived sacrifice made by individual feminists for the greater good serves as both evidence of the political nature of online violence and a reflection of the community's support, offering strength to endure the challenges. Finally, considering that the message of hate speech targets whole collectives, the assumption that online violence can be avoided by going offline becomes even more complex: Individualised responses cannot be considered sustainable solutions to structural problems.

10.5. Navigating self-protection, platform dynamics, and institutional challenges

Reactions by those who are targeted by online violence are often interconnected and overlapping, meaning that individual women's responses cannot always be neatly categorised as one thing or another. The normalisation and integration into daily life of digital self-protection measures reflects an acknowledgment and

acceptance of the prevalence of online violence among certain groups. As has been shown, digital protection strategies are often interconnected and include generalised protection strategies and platform-specific protection mechanisms. Both sets of strategies are widely adopted and contribute, on the one hand, to a certain sense of increased control over personal online spaces, and on the other hand, to immediate mitigation of impact during or after an attack.

The role of social media platform operators and their lack of efficient involvement in protecting vulnerabilised groups has been highlighted throughout the ethnography. Content moderation in all its complexity - encompassing legal, cultural, and ethical dimensions - has a fundamental role in the functioning of social media platforms, shaping their identity and functioning within public discourse. Nevertheless, this ethnography has shown participants' disillusionment with the platforms' response - or lack of response - to reports of gendered and/or racist abuse. This disillusionment fosters a widespread lack of trust among targets, who perceive reporting mechanisms as inefficient, non-transparent, and arbitrary. This places a disproportionate burden on targets, and points to the need for more transparent and inclusive reporting measures.

Another set of responses concern publicly speaking out. Women's public narratives function as a form of active resistance, countering silencing mechanisms, and creating a space for shared experiences and collective resilience. These strategies emphasise the potential impact of collective responses in deterring online attackers and mitigating both individual and collective impacts. Nevertheless, not all feminists targeted by online hate choose to speak out publicly, but rather use private strategies such as simply muting or blocking attackers or seeking support from peers in private spaces. In this regard, the exploration of community-building in the context of online gendered violence reveals the crucial role that peer support plays in the absence of effective institutional and platform-based protection. Emotional support, whether through humour or through sharing experiences and collectively pinpointing the structural elements of online violence, provides a sense of belonging and solidarity.

Community-building expressed as an exchange of technical know-how and support also highlights the need for organised digital self-defence measures.

The thesis has evidenced that women targeted by online attacks often feel like peer support and digital protection measures constitute the only existing response. It is through bringing these experiences to the light that we can expose the challenges of seeking institutional support, in the context of a pervasive sense of abandonment by institutions and a persistent need for more organised support networks. In this regard, experiences of reporting online violence to the police speak of a lack of efficacy in the criminal justice system. Despite facing severe threats and harassment, these women encountered minimal response or follow-up from law enforcement agencies, leaving them with a sense of helplessness and abandonment. The hesitancy to report is further compounded by a range of factors, including a lack of faith in the police's willingness to address online gendered abuse, the perception that incidents are not serious enough to warrant reporting, and the belief that even if reported, no meaningful action will be taken. Further, normalisation of online violence and a daily exposure to hate messages may lead targets to downplay the severity of the abuse, hindering their recognition of the potential for justice through formal systems. Ultimately, the narratives reflect a need for systemic changes, encompassing both legal frameworks and law enforcement practices, to effectively address and combat online gendered violence.

The lack of institutional support places a disproportionate burden on women feminists and activists to protect themselves, emphasising that it is their responsibility to prevent, stop or avoid violence against them, rather than on the perpetrators to stop acting violently (e.g., Maher et al, 2017). What is left then for those who are directly targeted are measures of self-protection and community-building with peers, to exchange practices and strategies for self-protection; to discuss and debate in closed, private circles, thus avoiding the risk of attacks, whilst not reaching out with their messages; and to gather with peers online or offline to laugh about it all. Nevertheless, as Jo Smith (2019) highlights, by placing this responsibility on women, there is a risk that women will be blamed

when targeted, for failing to adopt the right responses, for failing to take steps to avoid the abuse, and, for bringing abuse upon themselves for acting provocatively.

10.6. Legal and institutional responses

As has been revealed by this thesis, discerning the gender-based motivation in online violence encounters various challenges in terms of practical application of the legal framework. This demonstrates the extent to which violence against women is still largely perceived as belonging to the private sphere, with insufficient social and institutional acknowledgment of this violence as gendered outside the context of the heteronormative couple or ex-couple. Despite the 2015 reform of the Spanish criminal code incorporating sex and gender as aggravating circumstances, this has been sparingly applied, especially concerning hate speech. The lack of clear recognition results in the absence of a direct and cohesive legal framework tailored for addressing hate speech or violence targeting women online. As argued and shown throughout, the legal framework in this domain is fragmented, which has an impact not only on the legal tools available but also other support services accessible to the affected individuals.

For institutions, public services, civil society organizations and other professionals, the dynamic nature of technological advancements calls for a continuous renewal and update of knowledge to keep pace with the rapidly evolving landscape of digital spaces. The thesis research highlights a level of unpreparedness among professionals to offer adequate support to victims of gender-based online violence. This deficiency stems from a lack of specific training and protocols, forcing professionals to rely on informal knowledge. Yet, the challenge of keeping pace with technological advancements stems not only from the swift progress in technology but also from inadequate resources. The limited allocation of resources to address this issue can be attributed, in part, to the undervaluation of events occurring in digital realms, even if they involve violence. Generally, online violence tends to be relegated to a lower position in the hierarchy of violence within the prevalent perspective that separates online

and offline domains. Despite recognition by professionals of the potential impact of online violence on women, there is a noticeable gap between awareness and practical efforts to provide support in digital spaces. The perceived dichotomy between the “real” and the “digital” world contributes to the perception of online violence as “not real”, despite its profound impact, magnified by the ubiquity of digital spaces facilitated by mobile phones and the persistent nature of the violence. In this paradigm, gendered online violence is simultaneously seen as inevitable (and normalised), yet it remains largely invisible within the hierarchy of violence when compared to physical violence. Practically, this perception of online spaces and online violence significantly influences the available support for those targeted by gendered online violence. Acknowledging and addressing these perceptions is crucial for enhancing the support mechanisms for victims in the realm of gender-based online violence.

Public services aiding victims of crimes and discriminatory incidents show a notable divide, as they focus on either hate crimes (racist and anti-LGBT incidents) or gender-based violence (partner or ex-partner violence). Women facing other forms of gendered violence, like online hate speech, fall through these gaps, as these incidents are not categorised as “gender-based violence” nor as “hate crimes”. This classification results in a lack of comprehensive analysis, understanding, statistical data, as well as specialised support services. Failure to recognise the gender motivation in digital crimes outside intimate relationships creates a deficiency in expertise, impacting victim support and technical aspects. This knowledge gap may affect the entire justice process, from reporting to outcomes, significantly impacting women targeted by such violence.

10.7. Final reflections and future research paths

This thesis has underscored the role of hate speech in causing a type of linguistic injury, aligning with Judith Butler's (1997) conceptualisation of this phenomenon. The absence of precise language for describing linguistic injury makes it challenging to discern the distinct nature of linguistic vulnerability compared to physical vulnerability (Ibid.). At the same time, the use of physical metaphors to

illustrate linguistic injury underscores the crucial role of the bodily dimension in comprehending linguistic distress. Thus, the metaphorical association between physical and linguistic vulnerability seems almost indispensable in describing linguistic vulnerability itself. Experiencing injury through speech often entails a loss of context, resulting in a feeling of disorientation—a sense of being displaced or losing control. Essentially, harmful speech can instil a sense of insecurity in its targets. Through this process, a crucial function of hate messages is to create feelings insecurity and non-belonging among targeted individuals and groups.

The research and findings of this thesis bring evidence to how contemporary anti-feminist or anti-gender campaigns utilise individualised approaches. Consequently, distinguishing attacks on feminism from those specifically aimed at individual women becomes challenging. This phenomenon highlights the dual nature of gender-based hate speech, occurring both against individual women and women collectively as a group (Igareda, 2022: 102). Despite the profound impact on individual targets, these are attacked as symbolic representations of feminism, available as easy targets in digital spaces.

Deprived of institutional or platform support, individuals targeted may resort to seeking justice through informal means, engaging in digilantism, which includes speaking out or naming and shaming on social media. These approaches are closely tied to discussions on “viral justice” (Wood et al, 2019), which is a concept used to describe instances where victims share their own evidence of abuse or “survivor selfie”, leading to it going viral. As noted by Anastasia Powell and Nicola Henry (2017: 278), this often occurs in the context of the criminal justice system’s failure to hold perpetrators accountable. While “justice” is often narrowly associated with formal criminal justice responses by state institutions, some response strategies observed in the ethnography can be seen as informal attempts to seek “justice” through online spaces and activist communities. To some extent, the justice system is no longer seen as an essential tool for feeling that “justice has been done” but is viewed as inefficient, particularly in addressing online harm. To improve the response by institutions in general and the justice system in particular, a more victim-centred approach is needed,

including forms of community-led or informal justice operating outside of the state (e.g., Daly, 2014; McGlynn, 2011).

Feminists, as well as LGBT+ and anti-racist activists, need closed safe spaces for purposes of empowerment, to discuss strategies and to express support for each other; however, they also need to be able to express their opinions in public spaces of debate without the constant risk of suffering hateful attacks. Constructing alternative platforms that facilitate active participation can be a crucial step, but the process requires time and resources for building a new online environment.

Looking ahead, there is a pressing need for future research endeavours to delve into the realm of perpetrators, with the aim to deepen our understanding of the underlying mechanisms that propel parasitic behaviour and the monetisation of hate against feminists and other particularly targeted groups of women and/or LGBT+ persons. Investigating the motivations and dynamics behind individuals engaged in online hate speech is crucial for developing effective prevention and intervention strategies, and for exploring the responsibilities of platforms in this regard. Further, more research would be needed to throw light on the development of algorithms and their role in obscuring or shaping content related to gender equality and feminism on social media platforms. This would focus, for instance, on the invisibilisation of content through what is often informally referred to as “shadowbanning”.¹ Shadowbanning takes place when posts or activities remain hidden from other users, even though the affected user has not received an official ban or notification by the platform. Initially designed to allow platforms to let spammers continue their activities without being visible to the community, there is evidence from feminists suggesting that this platform mechanism also targets specific activist content. Consequently, more comprehensive research is needed to elucidate how these platform functions and algorithms play a role in shaping feminism in digital spaces. A further aspect that would need more in-depth qualitative research is the role of “survivor selfies” and “virtual justice” as paths for creating restoration and support networks for hate crime victims, linked to a discussion on the impact of these informal justice-

making processes on formal processes of justice. Finally, there is a need to expand research efforts beyond the current focus on gender-based online hate speech to address other forms of under-researched digital violence. For instance, digital violence targeting trafficking victims emerged as a noteworthy concern in interviews with representatives of support services for victims of gender-based violence, indicating a need for dedicated studies on how trafficking victims are digitally controlled by perpetrators. While acknowledging the limitations of the present thesis in comprehensively covering such areas, future research endeavours should bridge these gaps to provide a holistic understanding of the diverse facets of gender-based digital violence. These emerging fields would benefit from an ethnographic approach that combines digital ethnography with in-depth interviewing.

10.7.1. On a positive note...

When writing this chapter in February 2024, I'm suddenly reached by the news that EU legislators have attained a provisional deal on rules to combat gender-based violence and protect its victims.⁹¹ This is an important step towards a European Directive on combating violence against women. This Directive is in lines with the Istanbul Convention, and as such, includes a broad definition of gender-based violence. It recognises the impact of gender-based digital violence and criminalises five offences of cyberviolence. These include non-consensual sharing of intimate and manipulated material, cyberstalking, cyberharassment, cyberflashing and cyber incitement to hatred or violence. When transposed, this Directive will oblige Member States to adapt their laws to include these provisions.

While enhancing laws should not be the sole method for increasing societal responses, the symbolic importance of enacting legislation against gendered online violence cannot be understated in terms of delineating the boundaries

⁹¹ European Parliament News: "First ever EU rules on combating violence against women: deal reached" (February 6, 2024), available at <https://www.europarl.europa.eu/news/en/press-room/20240205IPR17412/first-ever-eu-rules-on-combating-violence-against-women-deal-reached>, last accessed on April 19, 2024.

between socially and culturally acceptable and unacceptable behaviours (e.g., Dixon and Gadd, 2006; Mason-Bish, 2011; Smith, 2019). If hate crimes and gender-based violence function as "message crimes" (Chakraborti, 2012) intended to convey a message to all members of the targeted group, then the framework for identifying and dealing with such behaviours holds a potent symbolic role in the shaping of normative standards. Including online misogyny in this framework, allowing its reporting and visibility in a novel manner, achieves the same symbolic function: explicitly declaring these once-normalised behaviours as unacceptable.

References

- Adkins, L. (1999). Community and economy: A retraditionalization of gender?, *Theory, Culture & Society*, 16(1), p.117-137.
- Adorno, T. (1975). Culture industry reconsidered, *New German critique*, 6, 12–19. (Original work published 1967).
- Alabao, N. (2021). Las guerras de género: La extrema derecha contra el feminismo. In Ramos, M. (Ed.). *De los neocon a los neonazis: La derecha radical en el estado español*. Madrid: Fundación Rosa Luxemburgo, pp. 397-423. <https://www.rosalux.eu/es/article/1954.las-guerras-de-g%C3%A9nero.html>
- Aguilar Fernandez, S. (2013). La jerarquía católica española en perspectiva comparada: La confrontación política entre la Iglesia y el Gobierno socialista a comienzos del siglo XXI, *Revista internacional de Sociología*, 71(2), 309–334.
- Allen, C. (2015). ‘People Hate You Because of the Way You Dress’: Understanding the Invisible Experiences of Veiled British Muslim Women Victims of Islamophobia, *International Review of Victimology* 21: 287–301
- Althusser, L. (2001 [1971]). *Lenin and Philosophy and Other Essays*. New York and London: Monthly Review Press, pp. 170-86.
- Álvarez Rodríguez, I. (2019). El discurso de odio sexista (en Construcción), *Revista Jurídica de Castilla y León*, 48, 43.
- Amnesty International (2018). *Toxic Twitter*. Available at <https://www.amnesty.org/en/latest/research/2018/03/online-violence-against-women-chapter-1-1/>
- Andrejevic, M. (2015). Foreword, in R. E. Dubrofsky and S.A. Magnet (Eds.) *Feminist Surveillance Studies*, ix–xviii. Durham: Duke University Press.
- Anthias, F. (2013). Intersectional what? Social divisions, intersectionality and levels of analysis, *Ethnicities*, 13: 3.
- Antoniades, A. (2012). Viral justice: Domestic abuse victim calls out attacker on Facebook, *Take Apart*. Available at <http://www.takepart.com/article/2012/14/viral-justice-domestic-abuse-victim-calls-out-attacker-facebook>
- Arce-García, S. and Menéndez-Menéndez, M.I. (2023). Inflaming public debate: a methodology to determine origin and characteristics of hate speech about sexual and gender diversity on Twitter, *Profesional de la información*, v. 32, n. 1, e320106. <https://doi.org/10.3145/epi.2023.ene.06>
- Arcila-Calderón C., Amores J.J., Sánchez-Holgado P., Blanco-Herrero D. (2021). Using Shallow and Deep Learning to Automatically Detect Hate Motivated by Gender and Sexual Orientation on Twitter in Spanish, *Multimodal Technologies and Interaction*; 5(10):63. <https://doi.org/10.3390/mti5100063>
- Arcila-Calderón, C., Blanco-Herrero, D., Valdez-Apolo, M.B. (2020). Rechazo y discurso de odio en Twitter: análisis de contenido de los tuits sobre migrantes y

refugiados en español, *Reis*, v. 172, pp. 21-40.
<https://doi.org/10.5477/cis/reis.172.21>

Arthur, B.W. (2009). *The Nature of Technology: What It Is and How It Evolves*. Free Press.

Association of Internet Researchers (2012). *Ethical Decision-Making and Internet Research*, available at <http://aoir.org/reports/ethics2.pdf>

Association of Social Anthropologists of the UK and the Commonwealth (ASA) (2011). *Ethical Guidelines for good research practice*.

Aull Davies, C. (1999). *Reflexive Ethnography. A Guide to Researching Selves and Others*. London: Routledge.

Austin, J.L. (1962). *How To Do Things With Words*. Oxford: Oxford University Press.

Awan, I. and Zempi, I. (2015). *We Fear for our Lives: Offline and Online Experiences of Anti-Muslim Hostility*. Tell Mama.

Balka, E. (1993). Women's access to on-line discussions about feminism, *Electronic Journal of Communication*, 3(1).

Banet-Weiser, S. and Miltner, K. (2016). #MasculinitySoFragile: culture, structure and networked misogyny, *Feminist Media Studies*, 16(1), pp. 171–174. DOI: 10.1080/14680777.2016.1120490

Bangstad, S. (2017). Can there be an anthropology of hate speech?, *Anthropology News*, May 11, 2017.

Barak, A. (2005). Sexual harassment on the internet, *Social Science Computer Review*, 23(1), 77–92.

Barcellona, M. (2022). Incel violence as a new terrorism threat: A brief investigation between Alt-Right and Manosphere dimensions, *Sortuz: Oñati Journal of Emergent Socio-Legal Studies*, 11(2), pp. 170–186.

Bareither, C. (2019). Doing Emotion through Digital Media: An Ethnographic Perspective on Media Practices and Emotional Affordances, *Ethnologia Europea* 49(1)

Barendt, E. (2016). *Anonymous Speech*. Oxford: Hart Publishing.

Barker, K. and Jurasz, O. (2022). Online misogyny as a hate crime: #TimesUp. In Zempi, I. and Smith, J. (Eds.) *Misogyny as Hate Crime*. Abingdon/New York: Routledge.

Barth, F. (1969). *Ethnic Groups and Boundaries: The Social Organisation of Culture Difference*. Oslo: Universitetsforlaget.

Bauman, Z. (1991). *Modernity and Ambivalence*. Cambridge: Polity Press

Bauman, Z. (2000). *Liquid modernity*. Cambridge: Polity Press

Bauman, Z. (2003). *Liquid Love*. Cambridge: Polity Press

- Beard, M. (2013). Internet Fury: Or Having Your Anatomy Dissected Online, *The Times Literary Supplement: A Don's Life*, available online at http://timesonline.typepad.com/dons_life/2013/01/internet-fury.html.
- Behar, R. (1997). *The Vulnerable Observer: Anthropology That Breaks your Heart*. Boston: Beacon Press.
- Beninger, K., Fry, A., Jago, N., Lepps, H., Nass, L., and Silvester, H. (2014). *Research using Social Media Users' Views*. London: National Centre for Research Methods.
- Biglia, B. (2007). Desde la investigación-acción hacia la investigación activista feminista. In Romay Martínez, J. (Coord.) *Perspectivas y retrospectivas de la psicología social en los albores del siglo XXI* (pp. 415-422). Madrid: Biblioteca Nueva.
- de Block, L. and Buckingham, D. (2007). *Global Children, Global Media*. London: Palgrave.
- Boellstorff, T., Nardi, B., Pearce, C. and Taylor, T.L. (2012). *Ethnography and Virtual Worlds: A Handbook of Method*. Princeton, NJ; Oxford, UK: Princeton University Press.
- Bolter, J.D., & Grusin, R.A. (1996). Remediation, *Configurations* 4(3), 311-358. <https://doi.org/10.1353/con.1996.0018>.
- Boromisza-Habashi, D. (2021). There's no such thing as hate speech and it's a good thing too. In Udupa, S., Gagliardone, I. and P. Hervik (Eds.) *Digital hate: the global conjuncture of extreme speech*. Bloomington, Indiana: Indiana University Press.
- Bourdieu, P. (1991). *Language and symbolic power*. Cambridge: Polity Press.
- Bourdieu, P. and Wacquant, L.J.D. (1992). *An Invitation to Reflexive Sociology*. Chicago: The University of Chicago Press.
- Brah, A. (2004). Ain't I A Woman? Revisiting Intersectionality, *Journal of International Women's Studies*, Vol 5, nº 3.
- Broadband Commission (2015). *Cyber Violence against Women and Girls: A world-wide wake-up call*.
- Brown, A. (2018). What is so special about online (as compared to offline) hate speech?, *Ethnicities*, Vol. 18(3) 297–326
- Burawoy M. (2003). Revisits: An outline of a theory of reflexive ethnography, *American Sociological Review* 68(5): 645-679.
- Burgess, M., Byars, F., Sadeghi-Azar, L. and K.E. Dill-Shackleford (2017). Online Misogyny Targeting Feminist Activism: Anita Sarkeesian and Gamergate. In Sturme, P. (Ed.) *The Wiley Handbook of Violence and Aggression*. Hoboken, NJ: Wiley-Blackwell
- Burnap, P. and Williams, M. (2016). Us and them: identifying cyber hate on Twitter across multiple protected characteristics, *EPJ Data Science* 2016 5:11:

- Burruss, G., Howell, C.J., Bossler, A. and Holt, T.J. (2020). Self-perceptions of English and Welsh constables and sergeants preparedness for online crime: A latent class analysis, *Policing: An International Journal*, Vol. 43 No. 1, pp. 105-119. <https://doi.org/10.1108/PIJPSM-08-2019-0142>
- Butler, J. (1988). Performative acts and gender constitution: An essay in phenomenology and feminist theory, *Theatre Journal*, 40(4), p. 519-531.
- Butler, J. (1990). *Gender Trouble: Feminism and the Subversion of Identity*. New York: Routledge.
- Butler, J. (1997). *Excitable Speech. A Politics of the Performative*. New York: Routledge.
- Caballero-Gálvez, A.; Zaera, A.; Tortajada, I. and Willem, C. (2022). Fe-MI-nismo en las redes sociales: riesgos y oportunidades para el feminismo contemporáneo Atlánticas, *Revista Internacional de Estudios Feministas*, 7 (1), 62-87. doi: <https://dx.doi.org/10.17979/arief.2022.7.1.7049>
- Calala Fondo de Mujeres (Coord.) (2020). *Las violencias machistas en línea hacia activistas. Datos para entender el fenómeno*.
- Caldevilla-Domínguez, D.; Barrientos-Báez, A.; García-Manso, A.; Matarín-Rodríguez-Peral, E. (2022). Neurocomunicación y Manosferas: estudio de caso Forocoches, *Historia y comunicación social*, 27(2), pp. 509-519
- Canyelles, C. (2018). *Masclisme i cultura jurídica. Una etnografia del procés judicial de la violència de gènere*. Tesis doctoral, Departament d'Antropologia Social, Universitat de Barcelona.
- Caplan, R. (2018). *Content or context moderation? Artisanal, Community-Reliant, and Industrial Approaches*. Data & Society. Available at https://datasociety.net/wp-content/uploads/2018/11/DS_Content_or_Context_Moderation.pdf
- Caputi, J. (1993). The Sexual Politics of Murder. In Bart, P. and Moran, E. (Eds.) *Violence against Women*. Newbury Park, CA: Sage, 5-25.
- Casemajor, N., Couture, S., Delfin, M., Goerzen, M. and Delfanti, A. (2015). Non-Participation in Digital Media: Toward a Framework of Mediated Political Action, *Media, Culture & Society* 37 (6): 850-866. doi:10.1177/0163443715584098.
- Castells, M. (2001). *Internet Galaxy. Reflections on the Internet, Business and Society*. Oxford: Oxford University Press.
- Çelik, S., & Paulos, B. (2021). The challenges of regulating hate speech on social media in light of the theory of freedom of expression, *Süleyman Demirel Üniversitesi Hukuk Fakültesi Dergisi*, 11(1), 97-134. <https://doi.org/10.52273/sduhfd..922588>
- Chadwick, R. (2021). Theorizing voice: toward working otherwise with voices, *Qualitative Research*, Vol. 21(1) 76 -101.
- Chakraborti, N., and Garland, J. (2015). *Hate crime: Impact, causes and responses* (2nd ed.). London: Sage.

- Chakraborti, N. and Hardy, S.J. (2015). *LGB&T Hate Crime Reporting. Identifying Barriers and Solutions*. Equality and Human Rights Commission.
- Citron, D.K. (2009a). Law's Expressive Value in Combating Cyber Gender Harassment, *108 Michigan Law Review*, 373-415
- Citron, D.K. (2009b). Cyber civil rights, *Boston University Law Review*, 89. 61–125.
- Citron, D.K. and Norton, H.L. (2011). Intermediaries and hate speech: Fostering digital citizenship for our information age, *Boston University Law Review*, 91(4).
- Citron, D.K. (2014). *Hate crimes in cyberspace*. Cambridge, MA: Harvard University Press.
- Clifford, J. (1986) Introduction: Partial Truths. In Clifford, J. and Marcus, G.E. (Eds.) *Writing Culture: The Poetics and Politics of Ethnography*. Berkeley: University of California Press.
- Coburn P., Connolly, D. and Roesch, R. (2015). Cyberbullying: Is federal criminal legislation the solution?, *Canadian Journal of Criminology and Criminal Justice* 57(4): 566–579
- Coleman, G.E. (2010). Ethnographic Approaches to Digital Media, *Annual Review of Anthropology*, 39: 487-505.
- Coleman, G. and Jandrić, P. (2019). Postdigital Anthropology: Hacks, Hackers, and the Human Condition, *Postdigital Science and Education* 1(2): 525–50. <https://doi.org/10.1007/s42438-019-00065-8>.
- Collier, B. and Bear, J. (2012). Conflict, criticism, or confidence: an empirical examination of the gender gap in Wikipedia contributions, *Proceedings of the ACM 2012 conference on computer supported cooperative work*, 11–15 February 2012, Seattle, WA, pp. 383–392. New York: ACM.
- Colom, A. (2022). Using WhatsApp for focus group discussions: ecological validity, inclusion and deliberation, *Qualitative Research*, 22(3), 452-467. <https://doi.org/10.1177/1468794120986074>
- Comas, D. (2011). “La perspectiva de gènere en la producció científica i en la societat. L'exemple de la violència contra les dones”. In Pastor, I. (Ed.) *Cent anys de dones a la Universitat*. Tarragona: Publicacions URV
- Cornejo-Valle, M. and Pichardo, J.L. (2017). La “ideología de género” frente a los derechos sexuales y reproducción. El escenario español, *Cadernos Pagu* (50), <https://doi.org/10.1590/18094449201700500009>
- Council of Europe (2018). *Mapping study on cyberviolence*. Cybercrime Convention Committee (T-CY). Working Group on cyberbullying and other forms of violence, especially against women and children.
- Crenshaw, K. (1989). Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics, *University of Chicago Legal Forum*, 1989: 139-67.
- Crosas Remón, I. and Medina-Bravo, P. (2018). Ciberviolencia en la red. Nuevas formas de retórica disciplinaria en contra del feminismo, *Papers* 2019, 104/1

- Csordas, T.J. (1990). Embodiment as a Paradigm for Anthropology, *Ethos*, 18 (1).
- Daly, K. (2014). Reconceptualizing sexual victimization and justice. In I. Vanfraechem, I., Pemberton, A. and Ndahinda, F.M. (Eds.) *Justice for victims: Perspectives on rights, transition and reconciliation*, pp.378-395. London: Routledge.
- Daniels, J. (2008). Race, Civil Rights, and Hate Speech in the Digital Era. In Everett, A., The, J. D. and MacArthur, C.T. (Eds.) *Learning Race and Ethnicity: Youth and Digital Media*. Foundation Series on Digital Media and Learning. Cambridge, MA: The MIT Press
- Daniels, J., Gregory, K. and McMillan Cottom, T. (Eds.) (2017) *Digital Sociologies*. Bristol: Policy Press.
- Davis, K. (2008) Intersectionality as buzzword: A sociology of science perspective on what makes a feminist theory successful, *Feminist Theory*, 2008; 9: 67
- De Beauvoir, S. (1976[1949]). *The Second Sex*. Paris: Gallimard.
- Delgado, R. and Stefancic, J. (2014). Hate speech in cyberspace, *Wake Forest Law Review* 49: 319–343
- Derrida, J. (1976). *Of Grammatology*. Baltimore: Johns Hopkins University Press.
- Deuze, M. (2007). *Media Work*. Cambridge: Polity Press, 2007.
- Deuze, M., Blank, P. and L. Speers (2012). A Life Lived in Media, *Digital Humanities Quarterly*, 6(1).
- Díaz-Fernández, S. and García-Mingo, E. (2022). The bar of Forocoches as a masculine online place: Affordances, masculinist digital practices and trolling, *New Media & Society*, 0(0). <https://doi.org/10.1177/14614448221135631>
- Directorate General Research and Innovation of the European Commission (2021) *Ethics and Data Protection*, available at https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/guidance/ethics-and-data-protection_he_en.pdf
- Dixon, B. & Gadd, D. (2006). Getting the message? New labour and the criminalization of hate, *Criminology & Criminal Justice*, 6(3), 309-328.
- Dodge, A. and Spencer, D.C. (2017). Online Sexual Violence, Child Pornography or Something Else Entirely? Police Responses to Non-Consensual Intimate Image Sharing among Youth, *Social & Legal Studies* 1–22.
- Domínguez Ruiz, I.E., Roiha, M. and Jubany, O. (2023). Restorative solutions for anti-LGBT victimisation experiences: potential pathways for victims' wellbeing and key challenges and needs, *Culture, Health & Sexuality*, 25:12, 1626-1639, DOI: 10.1080/13691058.2022.2105401
- Donath, J.S. (1999) "Identity and deception in the virtual community," In Smith, M.A. and Kollock, P. (Eds.) *Communities in cyberspace*. London: Routledge.
- Easter, B. (2018). Feminist_Brevity_in_light_of_masculine_long-windedness: Code, space, and online misogyny, *Feminist Media Studies*, 18 (4): 675-685.

- Eckert, S. (2018). Fighting for recognition: Online abuse of women bloggers in Germany, Switzerland, the United Kingdom, and the United States, *New Media & Society* 20(1), DOI:10.1177/1461444816688457
- Eichhorb, K. (2001). Re-in/citing linguistic injuries: speech acts, cyberhate, and the spatial and temporal character of networked environments, *Computers and Composition*, 18 (2001) 293–304.
- Ellis-Sloan, K. (2014). Understanding teenage motherhood through feminist research: A reflection on the challenges and advantages, *Athenea Digital*, n.14(4), pp. 129-152
- European Commission against Racism and Intolerance (ECRI) (2016). *General Policy Recommendation N° 15*. Brussels: ECRI.
- Evans, E. (2015). What Makes a (Third) Wave?, *International Feminist Journal of Politics*, 18 (3): 409–428.
- Everett, A. (2009). *Digital Diaspora*. Albany: SUNY Press.
- Evolvi, G. (2018) Hate in a Tweet: Exploring Internet-Based Islamophobic Discourses, *Religions*, 9(10):307
- Feldman, M. and Littler, M. (2014). *Anti-Muslim Overview, Analysis and 'Cumulative Extremism'*. Tell MAMA Reporting, July 2014.
- Foucault, M. (1990). *The history of sexuality: An introduction, volume I*. New York: Vintage.
- Fox, G. and Nolte, G. (1995). Intolerant Democracies, *Harvard International Law Journal*, n° 36.
- Franzke, A.S., Bechmann, A., Zimmer, M., Ess, C, and the Association of Internet Researchers (2019). *Internet Research: Ethical Guidelines 3.0*, available at <https://aoir.org/reports/ethics3.pdf>;
- Franzke, A.S. (2020). *Feminist Research Ethics*, IRE 3.0 Companion 6.3.
- Fuchs, C. (2009) Information and communication technologies and society: A contribution to the critique of the political economy of the internet, *European Journal of Communication*, 24 (1): 69-87.
- Gagliardone, I. (2016). *The politics of technology in Africa: Communication, development, and nation-building in Ethiopia*. Cambridge, UK: Cambridge University Press.
- Gagliardone, I. (2019). Defining Online Hate and Its 'Public Lives': What Is the Place for 'Extreme Speech'?, *International Journal of Communication*, 13:3068–3087.
- Gagliardone, I. and Pohjonen, M. (2016). Engaging in polarized society: Social media and political discourse in Ethiopia. In Mutsavairo, B. (Ed.) *Digital activism in the social media era*, pp. 22–44. London, UK: Springer International Publishing.
- Gelber, K. (2002). *Speaking Back: The free speech versus hate speech debate*. Amsterdam: John Benjamins Publishing Company.

- Gelber, K. (2012). "Speaking back" – the Likely Fate of Hate Speech Policy in the United States and Australia. In Maitra, I. and McGowan, M.K (Eds.) *Speech and Harm: Controversies over Free Speech*. Oxford: Oxford University Press.
- Gelber, K. and McNamara, L. (2015). Evidencing the harms of hate, *Social Identities*, 324-341, 22(3)
- Gerard, J.F. and Whitfield, K.C. (2016). The experiences of victims of online Islamophobia. In Awan, I. (Ed.) *Islamophobia in Cyberspace*. London: Routledge.
- Gergen, K. (1991). *The saturated self. Dilemmas of identity in contemporary life*. New York: Basic Books.
- Gill, A. K., & Mason-Bish, H. (2013). Addressing Violence against Women as a Form of Hate Crime: Limitations and Possibilities, *Feminist Review*, 105(1), 1-20. <https://doi.org/10.1057/fr.2013.17>
- Gillespie, T. (2010). The politics of 'platforms', *New Media & Society*, 12(3): 347–364.
- Gillespie, T. (2015). Platforms Intervene, *Social Media + Society*, 1 (1): 2056305115580479.
- Gillespie, T. (2018). *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media*. New Haven, CT: Yale University Press.
- Glitch (2023). *The Digital Misogynoir Report: Ending the dehumanising of Black women on social media*. Available at https://glitchcharity.co.uk/wp-content/uploads/2023/07/Glitch-Misogynoir-Report_Final_18Jul_v5_Single-Pages.pdf
- Goffman, E. (1959). *The presentation of self in everyday life*. New York: Anchor Books.
- Gómez, V. (2016). Incitación al odio y género. Algunas reflexiones sobre el nuevo art. 510 CP y su aplicabilidad al discurso sexista, *Revista Electrónica de Ciencia Penal y Criminología*, Vol. 18, 2016
- Gómez, V. (2020). Odio en la Red. Una revisión crítica de la reciente jurisprudencia sobre Ciberterrorismo y Ciberodio, *Revista de Derecho Penal y Criminología*, 20(20): 411. <https://doi.org/10.5944/rdpc.20.2018.26458>
- Gomez Garcia, R. and Treré, E. (2014). The #YoSoy132 Movement and the Struggle for Media Democratization in Mexico, *Convergence: The Journal of Research into New Media Technologies*, 20 (4):496–510. doi: 10.1177/1354856514541744.
- Gray, K.L. (2012). Intersecting Oppressions and Online Communities: Examining the Experiences of Women of Color in Xbox Live, *Information, Communication & Society*, 15(3), 411
- Gray, K.L. (2013). Collective Organizing, Individual Resistance, or Asshole Grieffers? An Ethnographic Analysis of Women of Color in Xbox Live, *Ada: A*

Journal of Gender, New Media, and Technology, 2: Feminist Game Studies (June 2013)

Gurumurthy, A. (2011). Feminist Visions of the Network Society, *Development*, 54 (4): 464–469. doi:10.1057/dev.2011.82.

Gusterson, H. (1996). *Nuclear Rites: A Weapons Laboratory at the End of the Cold War*. Berkeley: University of California Press.

Graeber, D. (2004). *Fragments of an Anarchist Anthropology*. Chicago: Prickly Paradigm Press.

Graddol, D. and Swann, J. (1989). *Gender Voices*. London: Basil Blackwell

Gregorio Gil, C. (2014). Traspasando las fronteras dentro-fuera. Reflexiones desde una etnografía feminista, *Revista de Antropología Iberoamericana*, 9(3) Septiembre – Diciembre, pp. 297 – 322.

Guyan, K. (2022). *Queer Data. Using Gender, Sex and Sexuality Data for Action*. London: Bloomsbury.

Habermas, J. (1989). *The Structural Transformation of the Public Sphere*. New York: Polity Press.

Halford, S. and M. Savage (2010). Reconceptualizing digital social inequality, *Information, Communication & Society*, 13(7), p.937-955.

Haraway, D. J. (1991). *Simians, cyborgs, and women: the reinvention of nature*. New York: Routledge.

Harmer, E. and Lumsden, K. (2019). Introduction: Online Othering. In Lumsden, K. and Harmer, E. (Eds.) *Online Othering: Exploring Violence and Discrimination on the Web*. Basingstoke: Palgrave Macmillan.

Harris, B. and Vitis, L. (2020). Digital intrusions: technology, spatiality and violence against women, *Journal of Gender-Based Violence*, vol 4, no 3, 325–341, DOI: 10.1332/239868020X15986402363663

Hasinoff, A. (2014). Contradictions of Participation: Critical Feminist Interventions in New Media Studies, *Communication and Critical/Cultural Studies*, 11 (3): 270–272. doi:10.1080/147914 20.2014.926242.

Hasinoff, A. (2015). *Sexting Panic: Rethinking Criminalization, Privacy, and Consent*. Illinois: University of Illinois Press.

Hecht, G. (2001). Technology, Politics, and National Identity in France. In Allen, M. and Hecht, G. (Eds.) *Technologies of Power*, 253-294. Boston: MIT Press.

Heinze, E. (2006). Viewpoint Absolutism and Hate Speech, *The Modern Law Review*, 69(4), 543–582.

Henry, N. and Powell, A. (2014). The dark side of the virtual world. In Henry, N. and Powell, A. (Eds.) *Preventing Sexual Violence: Interdisciplinary Approaches to Overcoming a Rape Culture*. Basingstoke, UK and New York, US: Palgrave Macmillan.

- Herring, S. (1993). Gender and democracy in computer mediated communication, *Electronic Journal of Communication*, 3(2), 476–489.
- Herring, S. (1996). Bringing familiar baggage to the new frontier: Gender differences in computer-mediated communication. In Selzer, J. (Ed.) *Conversations* (pp. 1069-1082). Boston: Allyn & Bacon.
- Herring, S. (2001). Gender and power in online communication, *Center for Social Informatics Working Papers*, no. WP-01-05.
- Herring, S. (2002). Computer-mediated communication on the internet, *Ann. Rev. Info. Sci. Tech.*, 36: 109-168. <https://doi.org/10.1002/aris.1440360104>
- Herring, S. (2004). Computer-mediated communication and woman's place". In Bucholtz, M. (Ed.) *Robin Tolmach Lakoff Language and Woman's Place: Text and Commentaries*. NY: Oxford University Press.
- Herring, S. (2010). Who's got the floor in computer-mediated conversation? Edelsky's gender patterns revisited, *Language@Internet*, 7, article 8.
- Hill Collins, P. (1989). The Social Construction of Black Feminist Thought, *Signs*, Vol. 14, No. 4, Common Grounds and Crossroads: Race, Ethnicity, and Class in Women's Lives, pp. 745-773
- Hill Collins, P. (2017). On Violence, Intersectionality and Transversal Politics, *Ethnic & Racial Studies*, 40:9, 1460.
- Hill Collins, P. and Bilge, S. (2016). *Intersectionality. Key Concepts Series*. Cambridge, UK; Malden, MA: Polity Press.
- Hine, C. (2015). *Ethnography for the Internet: Embedded, Embodied and Everyday*. London, UK; New York, NY: Bloomsbury Publishing.
- Hylland Eriksen, T. (2007). *Globalization: The key concepts*. Oxford: Berg
- Internet Governance Forum (IGF) (2015). *Best Practice Forum (BPF) on Online Abuse and Gender-Based Violence Against Women*.
- Igareda, N., Pascale, A., Cruells, M. and Paz, O. (2019). *Les Ciberviolències masclistes*. Barcelona: Institut Català de les Dones.
- Igareda, N. (2022). El discurso de odio anti-género en las redes sociales como violencia contra las mujeres y como discurso de odio, *Derechos y libertades*, N° 47, Época II, junio 2022, pp. 97-122
- Instituto Andaluz de la Mujer (2015). *Protocolo para la detección e intervención a víctimas de ciberviolencias*. Available online at <https://www.juntadeandalucia.es/iam/catalogo/doc/iam/2015/143528391.pdf>
- Iphofen, R. (2021). *Research Ethics in Ethnography/Anthropology*. Brussels: European Commission, DG Research and Innovation.
- Ito, M. and Okabe, D (2005). Technosocial situations: Emergent structurings of mobile email use. In Ito, M., Okabe, D. and Matsude, M. (Eds.) *Personal, portable, pedestrian; Mobile phones in Japanese life*. Cambridge, MA: MIT Press.
- Jane, E. (2014a). You're a Ugly, Whorish, Slut, *Feminist Media Studies*, 14(4)

- Jane, E. (2014b). Back to the kitchen, cunt”: Speaking the unspeakable about online misogyny, *Continuum: Journal of Media and Cultural Studies*, 28: 558–570, DOI:10.1080/10304312.2014.924479
- Jane, E. (2015). Flaming? What Flaming? The Pitfalls and Potentials of Researching Online Hostility, *Ethics of Information Technology*, 17: 65–87.
- Jane, E. (2016). Online misogyny and feminist digilantism, *Continuum: Journal of Media and Cultural Studies*, DOI: 10.1080/10304312.2016.1166560
- Jane, E. (2017a). *Misogyny Online A Short (and Brutish) History*. Sage Swifts.
- Jane, E. (2017b). Feminist flight and fight responses to gendered cyberhate. In Segrave, M. and Vitis, L. (Eds.) *Gender, Technology and Violence*. London: Routledge.
- Jane, E. (2018). Systemic misogyny exposed: translating rape-gish from the manosphere with a random rape threat generator, *International Journal of Cultural Studies*, 21(6): 661–80
- Jay, P. (2023). *Put a gorda feminazi*. Madrid: Montena.
- Jeong, S. (2018). *The Internet of Garbage*. Forbes Media.
- Johnson, A. (2017). The multiple harms of sealions, *Perspectives of Harmful Speech Online*. Available at https://dangerousspeech.org/wp-content/uploads/2017/08/2017-08_harmfulspeech.pdf#page=13
- Jubany, O. and Roiha, M. (2018). *Las palabras son armas*. Barcelona: Edicions UB.
- Jubany, O. and Roiha, M. (2018). *REACT WP2 Comparative Report Part II: Counter-narratives to online hate speech: Conceptualizations, constructions and lessons learned*. REACT Project, available at http://www.reactnohate.eu/wp-content/uploads/2019/09/D2.7.-REACT_WP2-Comparative-report..pdf
- Jubany, O. and Guasch, O. (2020). Introducción. In Jubany, O. and Guasch, O. (Eds.) *Intersecciones encarnadas: [Con]textos críticos en género, identidad y diversidad*. Barcelona: Bellaterra.
- Juliano, D. (2004). *Excluidas y marginales*. Colección Feminismos. Madrid: Ediciones Cátedra
- Jägervi, L. (2014). Who wants to be an ideal victim? A narrative analysis of crime victims’ self- presentation, *Journal of Scandinavian Studies in Criminology and Crime Prevention*, 15 (2014), 73–88.
- Katiambo, D. (2021). It’s incivility, not hate speech. In Udupa, S., Gagliardone, I. and P. Hervik (Eds.) *Digital hate: the global conjuncture of extreme speech*. Bloomington, Indiana: Indiana University Press,
- Kayany, J. M. (1998). Contexts of uninhibited online behavior: Flaming in social newsgroups on Usenet, *Journal of the American Society for Information Science*, 49(12), 1135-1141.

- Kelly L. (1987). The Continuum of Sexual Violence. In Hanmer, J., Maynard, M. (Eds.) *Women, Violence and Social Control. Explorations in Sociology* (British Sociological Association Conference Volume series). London: Palgrave Macmillan.
- Kelly, L. (1987). *Surviving sexual violence*. Cambridge: Polity Press.
- Kendall, L. (2002). *Hanging Out in the Virtual Pub: Masculinities and Relationships Online*. Berkeley: University of California Press.
- KhosraviNik, M. and Esposito, E. (2018) Online hate, digital discourse and critique: Exploring digitally-mediated discursive practices of gender-based hostility, *Lodz Papers in Pragmatics*, 14.1, 45–68 Special issue on Narrating hostility, challenging hostile narratives: <https://doi.org/10.1515/lpp-2018-0003>
- Kingston Mann, L. (2014). What Can Feminism Learn from New Media?, *Communication and Critical/Cultural Studies*, 11:3, 293-297, DOI: 10.1080/14791420.2014.926244
- Kirkpatrick, D. (2011) *The Facebook Effect: The Inside Story of the Company That Is Connecting the World*. New York: Simon & Schuster.
- Kramarae, C. and Taylor, H.J. (1993). Women and Men on Electronic Networks: A Conversation or a Monologue?. In Taylor, H.J, Kramarae, C. and Ebben, M. (Eds.) *Women, Information Technology, and Scholarship*. Urbana, IL: Center for Advanced Study
- Kramer, M. (2021). The moral economy of extreme speech. Resentment and Anger in Indian Minority Politics. In Udupa, S., Gagliardone, I. and P. Hervik (Eds.) *Digital hate: the global conjuncture of extreme speech*. Bloomington, Indiana: Indiana University Press.
- Krasnova, H., Günther, O., Spiekermann, S. and Koroleva, K. (2009). Privacy Concerns and Identity in Online Social Networks, *Identity in the Information Society* 2 (1): 39–63.
- Krendel, A. (2022). From sexism to misogyny: can online echo chambers stay quarantined?. In Zempi, I. and Smith, J. (Eds.) *Misogyny as Hate Crime*. Abingdon/New York: Routledge.
- Krotoszynski, R. (2004). A comparative perspective on the First Amendment: free Speech, militant democracy, and the primacy of dignity as a preferred constitutional value in Germany, *Tulane Law Review*, nº78 (2004), pp. 154-
- Kunreuther, L. (2014). *Voicing Subjects: Public Intimacy and Mediation in Kathmandu*. Oakland, CA: University of California Press.
- Langarita, J.A., Sadurní, N., Dorado Caballero, A., and Albertín Carbó, P. (2018). From Zero to Hero: A Transnational Perspective on Countering Anti-LGBT Hate Crime in Europe. In Godzidz, P. and Viaggiani, G. (Eds.) *Running through Hurdles: Obstacles in the Access to Justice for Victims of Anti-LGBTI Hate Crimes*. Lambda Warsaw Association.
- Lange, P.G (2006). What is your claim to flame?, *First Monday*, Vol 11, nº 9.

- Langlois, G. (2015). What Are the Stakes in Doing Critical Research on Social Media Platforms?, *Social Media +Society*, 1 (1): 1–2. doi:10.1177/2056305115591178
- Lather, P. (1986). Research as Praxis, *Harvard Educational Review*, 56(3), pp. 257–278
- Latour, B. (2007). *Reassembling the Social: An Introduction to Actor-Network-Theory*. Oxford: Oxford University Press.
- Lawrence III, C.R. (1993). If he hollers let him go: Regulating racist speech on campus. In Matsuda, M.J., Lawrence III, C.R., Delgado, R. and K. Crenshaw (Eds). *Words that Wound: Critical race Theory, Assaultive Speech and the First Amendment*. Boulder: Westview Press.
- Lea, M., O'Shea, T., Fung, P., & Spears, R. (1992). *'Flaming' in computer-mediated communication: Observations, explanations, implications*. Harvester Wheatsheaf.
- Lévi-Strauss, C. (1963). *Structural Anthropology*. New York: Basic Books.
- Lewis, R., Rowe, M. and Wiper, C. (2016). Online abuse of feminists as an emerging form of violence against women and girls, *British Journal of Criminology*, 57(6), 1462–1481.
- Lewis, R., Rowe, M. and Wiper, C. (2019). Online/Offline Continuities: Exploring Misogyny and Hate in Online Abuse of Feminists. In Lumsden, K. and Harmer, E. (Eds.) *Online othering: Exploring digital violence and discrimination on the web*. Palgrave Macmillan.
- Linabary, J.R. and Batti, B. (2019). 'Should I even be writing this?': Public narratives and resistance to online harassment. In Ging, D. and Siapera, E. (Eds.) *Gender hate online: Understanding the new anti-feminism*. Palgrave Macmillan.
- Lister, R. (2004). *Poverty*. Cambridge: Polity Press.
- Lizardo-González, E. (2021). *Génesis, usos jurídicos y consecuencias (sociales y biográficas) de los Delitos de Odio y Discriminación por Orientación sexual e Identidad de género en España* [Tesis doctoral, UNIR]. Re-UNIR. <https://reunir.unir.net/handle/123456789/12828>
- Lizardo, E. (2023). *Delitos de odio y discriminación. El caso LGTB*. Barcelona: J.M. Bosch Editor.
- Long, J. (2015). The Oppression That Dare Not Speak Its Name? Silences around Heterosexuality in Contemporary Feminism. In Kiraly, M. and Tyler, M. (Eds.) *The Freedom Fallacy: The Limits of Liberal Feminism*, 145–152. Ballarat: Connor Court.
- MacKinnon, C (1991). *Toward a Feminist Theory of the State*. Harvard University Press.
- MacKinnon, C. (1993). *Only Words*. Cambridge, Mass: Harvard University Press
- MacKinnon, C. (2013). Intersectionality as Method: A Note, *Signs*, Vol.38, No 4, Intersectionality: Theorizing Power, Empowering Theory.

- Madianou, M. and Miller, D. (2012). Polymedia. Towards a new theory of digital media in interpersonal communication, *International Journal of Cultural Studies* 16 (2), 169–187.
- Mancinelli, F., Roca, J. And Jubany, O. (2023). Introducción. Categorías desbordadas: Cuestionando etiquetas, jerarquías y alteridades. In Jubany, O., Roca, J. and Mancinelli, F. (Eds.) *Categorías desbordadas: Cuestionando etiquetas, jerarquías y alteridades*. Barcelona: Bellaterra.
- Mantilla, K. (2015). *Gender trolling. How misogyny went viral*. Santa Barbara: Praeger.
- Marin de Espinosa, E. (2018). La agravante genérica de discriminación por razones de género (art. 22.4 CP), *Revista electrónica de ciencia penal y criminología*, N.º. 20, 2018, p. 27.
- Marques Gonçalves, G.; Willem, C. (2021). Lucha feminista gitana en España, lucha interseccional: el combate contra el Antigitanismo en las redes sociales en España, *Revista de Investigaciones Feministas*, 12(1), 127-143.
- Marwick, A.E. and Caplan, R. (2018). Drinking male tears: language, the manosphere, and networked harassment, *Feminist Media Studies*, DOI: 10.1080/14680777.2018.1450568
- Marwick, A.E and Lewis, R. (2017). *Media Manipulation and Disinformation Online*. New York: Data & Society Research Institute.
- Martin, J. S., Vaccaro, C. A., Heckert, D. A., & Heasley, R. (2015). Epic Glory and manhood acts in fantasy role-playing Dagorhir as a case study, *The Journal of Men's Studies*, 23(3), 293– 314
- Martínez-Valerio, L. (2021). Mensajes de odio hacia la comunidad LGTBIQ+: análisis de los perfiles de Instagram de la prensa española durante la “Semana del Orgullo”, *Revista Latina de Comunicación Social*, 80, 363-388. <https://doi.org/10.4185/RLCS-2022-1749>
- Mason-Bish, H. (2011). Examining the boundaries of hate crime policy: Considering age and gender, *Criminal Justice Policy Review*, 24(3), 297-316.
- Massanari, A. (2017). #Gamergate and the Fapping: How Reddit’s algorithm, governance, and culture support toxic technocultures, *New Media and Society*, 19 (3): 329-346
- Matsuda, M. J. (1993). Public Response to Racist Speech: Considering the Victim’s Story”. In Matsuda, M. J, Lawrence III, C.R., Delgado, R. and Crenshaw, K. (Eds.) *Words that Wound: Critical race Theory, Assaulted Speech and the First Amendment*. Boulder: Westview Press.
- Mauri, A. (2022). *Las locas del coño*. Letras libres. Available at <https://letraslibres.com/cultura/las-locas-del-cono/>
- McGlynn, C. (2011). Feminism, rape and the search for justice, *Oxford Journal of Legal Studies*, 31(4), 825-842.

- McGranahan, C. (2021) A presidential archive of lies. Racism, Twitter, and a History of the Present. In Udupa, S., Gagliardone, I. and P. Hervik (Eds.) *Digital hate: the global conjuncture of extreme speech*. Bloomington, Indiana: Indiana University Press
- Mead, M. (1935). *Sex and Temperament in Three Primitive Societies*. New York: William Morrow and Company.
- Megarry, J. (2014). Online incivility or sexual harassment? Conceptualising women's experiences in the digital age, *Women's Studies International Forum*, Vol 47, issue PA.
- Megarry, J. (2017). Under the watchful eyes of men: theorising the implications of male surveillance practices for feminist activism on social media, *Feminist Media Studies*, 18(6), p.1070-1085.
- Megías, I., Amezága, A., García, M.C., Kuric, S., Morado, R. and Orgaz, C. (2021). *Romper cadenas de odio, tejer redes de apoyo: los y las jóvenes ante los discursos de odio en la red*. Madrid: Centro Reina Sofía sobre Adolescencia y Juventud, Fad. DOI: 10.5281/zenodo.4288486
- Menéndez-Menéndez, M.I., Amigot-Leache, P. and Iturbide-Rodrigo, R. (2021). Narrativas sexistas y hostilidad en foros de prensa digital: Análisis en diarios de ámbito local, *Investigaciones feministas*, Vol. 12, Nº. 1, 2021
- Merleau-Ponty, M. (1962). *Phenomenology of Perception*. London: Routledge and Kegan Paul.
- Merleau-Ponty, M. (1968). *The Visible and the Invisible*. Northwestern University Press.
- Meyer, D. (2010). Evaluating the Severity of Hate-motivated Violence: Intersectional Differences among LGBT Hate Crime Victims, *Sociology*, 44(5), 980-995. <https://doi.org/10.1177/0038038510375737>
- Miles, R. and Brown, M. (2003) *Racism*. 2nd Edition. London: Routledge.
- Miller, D. and Slater, D. (2000). *The Internet: An Ethnographic Approach*. Oxford: Berg.
- Miller, D. and Horst H.A. (2012). The Digital and the Human: A Prospectus for Digital Anthropology. In Horst, H.A. and Miller, D. (Eds.) *Digital Anthropology*. London: Berg.
- Ministerio del Interior (2021). *Informe de la encuesta sobre los delitos de odio*. Available at https://www.interior.gob.es/opencms/pdf/servicios-al-ciudadano/Delitos-de-odio/descargas/Informe-de-la-encuesta-sobre-delitos-de-odio_2021.pdf
- Ministerio del Interior (2023). *Informe sobre la evolución de los delitos de odio en España*. Available at https://www.interior.gob.es/opencms/export/sites/default/galleries/galeria-de-prensa/documentos-y-multimedia/balances-e-informes/2022/Informe_Evolucion_delitos_odio_2022.pdf

- Miró Llinares, F. (2016). Taxonomía de la comunicación violenta y el discurso del odio en Internet, *Revista de Internet, Derecho y Política*, Nº 22, págs. 82-107. UOC.
- Mitchell, M., Wood, J., O'Neill, T., Wood, M., Pervan, F., Anderson, B., & Arpke-Wales, W. (2022). Technology-facilitated violence: A conceptual review. *Criminology & Criminal Justice*, 0(0), doi.org/10.1177/17488958221140549
- Moor, P., Heuvelman, A. and Verleur, R. (2010). Flaming on YouTube, *Computers in Human Behavior*, 26(6), pp. 1536-1546, <https://doi.org/10.1016/j.chb.2010.05.023>.
- Morena-Balaguer, D., García-Romeral, G. and Binimellis-Adell, M. (2021). *Diagnóstico sobre las violencias de genero contra activistas feministas en el ámbito digital*. Calala Fondo de Mujeres and Servei de Publicacions de UVic.
- Moya, B. and Trudy (2018) On Misogynoir: Citation, Erasure, and Plagiarism, *Feminist Media Studies*, 18, no. 4 (March 13, 2018): 762–68, <https://doi.org/10.1080/14680777.2018.1447395>
- Mullaney, L. and Trickett, L. (2018). *Misogyny Hate Crime Evaluation Report*. Available at <https://www.nottingham.ac.uk/lipp/documents/misogyny-evaluation-report.pdf>
- Murji, K. and Solomos, J. (2005). *Racialization: Studies in Theory and Practice*. Oxford: Oxford University Press.
- Nader, L. (2002a). *The Life of the Law. Anthropological Projects*. Berkeley: University of California Press.
- Nader, L. (2002b). A user theory of the law, *Southwestern Law Review*, 38(4), pp.951-963.
- Nielsen, L.B. (2002). Subtle, pervasive, harmful: Racist and sexist remarks in public as hate speech, *Journal of Social Issues*, (2002), 265-280, 58(2).
- Noble, S.U. (2018). *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York: New York University Press.
- Observatorio Nacional de Tecnología y Sociedad (Ontsi) (2022). *Violencia digital de género: una realidad invisible*. Available at <https://www.ontsi.es/es/publicaciones/violencia-digital-de-genero-una-realidad-invisible-2022>
- Oksanen, A., Hawdon, J., Holkeri, E., Näsi, M and Räsänen, P. (2014). Exposure to online hate among young social media users. In Warehime, M. (Ed.) *Soul of Society: A focus on the lives of children and youth*. Bingley: Emerald Group
- O'Neil, M. (2009). *Cyberchiefs: Autonomy and Authority in Online Tribes*. London: Pluto Press.
- Ortega, E. (2019). Las negras siempre fuimos queer. In Vila, F. and Sáez, J. (Eds.). *El libro de Buen Vmor. Sexualidades raras y políticas extrañas*. Madrid: Ayuntamiento de Madrid, 222-229.

- O'Sullivan, P.B. and Flanagan, A.J. (2003). Reconceptualizing 'Flaming' and Other Problematic Messages, *New Media & Society*, vol 5, nº 1, pp. 69–94. <http://dx.doi.org/10.1177/1461444803005001908>
- Palese, E. (2013). Zygmunt Bauman. Individual and society in the liquid modernity, *SpringerPlus*, 2013, 2:191
- Paternotte, D. and Kuhar, R. (2018). Disentangling and Locating the “Global Right”: Anti-Gender Campaigns in Europe, *Politics and Governance*, 6(3), 6-19. doi:<https://doi.org/10.17645/pag.v6i3.1557>
- Paterson, J., Walters, M.A., Brown, R. and Fearn, H. (2018). *The Sussex Hate Crime Project. Final Report*. Leverhulme Trust.
- Patterson, A. N. (2018). YouTube Generated Video Clips as Qualitative Research Data: One Researcher's Reflections on the Process, *Qualitative Inquiry*, 24(10), 759–767. <https://doi.org/10.1177/1077800418788107>.
- Pease, B. (2020). The rise of angry white men: Resisting populist masculinity and the backlash against gender equality. In Noble, C. and Ottmann, G. (Eds.) *The Challenge of Right-wing Nationalist Populism for Social Work: A Human Rights Approach*, 55-68.
- Perry, B. (2001). *In the name of hate: Understanding hate crimes*. London: Routledge.
- Pfeifer, S. and Jovovic, S. (2023). Co-Teaching Postdigital Ethnography, *Working Paper Series Collaborative Research Center 1187 Media of Cooperation*, Vol. 34. Universitätsbibliothek Siegen. <https://doi.org/10.25819/UBSI/10427>.
- Phillips, W. (2015). *This Is Why We Can't Have Nice Things: Mapping the Relationship between Online Trolling and Mainstream Culture*. MIT Press.
- Phillipson, G. (2015). Hate Speech Laws: What they should and shouldn't try to do, *Revue générale du droit*, Etudes et réflexions 2015, nº 13.
- Pina-Cabral, J. (2020). On embracing the vague, *HAU: Journal of Ethnographic Theory*, Vol 10, Nº 3, Winter 2020.
- Piñeiro-Otero, T. and Martínez-Rolán, X. (2021). Say it to my face: analysing hate speech against women on Twitter, *Profesional de la información*, v. 30, n. 5, e300502. <https://doi.org/10.3145/epi.2021.sep.02>
- Plant, S. (1995). The Future Looms: Weaving Women and Cybernetics, *Body & Society*, 1(3-4), 45-64. <https://doi.org/10.1177/1357034X95001003003>
- Platero, R. L., and Ortega Arjonilla, E. (2016). Building coalitions: The interconnections between feminism and trans* activism in Spain, *Journal of Lesbian Studies*, 20(1), 46– 64.
- Platero, R. L. (2020). Recognition and distribution of trans rights in Spain, *Politics and Governance*, 8(3), 253– 265
- Pohjonen, M. and Udupa, S. (2017) Extreme Speech Online: An Anthropological Critique of Hate Speech Debates, *International Journal of Communication*, 11(2017), 1173–1191

- Portwood-Stacer, L. (2014). Feminism and Participation: A Complicated Relationship, *Communication and Critical/Cultural Studies*, 11:3, 298-300, DOI: 10.1080/14791420.2014.926246
- Powell, A., & Henry, N. (2016). Policing technology-facilitated sexual violence against adult victims: police and service sector perspectives, *Policing and Society*, 28(3), 291–307. doi:10.1080/10439463.2016.1154964
- Powell, A. and Henry, N. (2017). *Sexual violence in a digital age*. London: Palgrave MacMillan.
- Pratt, M.L. (1986). Fieldwork in common places. In Clifford, J. and Marcus, G.E. (Eds.) *Writing Culture: The Poetics and Politics of Ethnography*. Berkeley: University of California Press.
- Ragnedda, M. and G. Muschert, G.W. (2013). *The digital divide: The Internet and social inequality in international perspective*. London: Routledge.
- Rajani, N. (2022). “I Bet You Don’t Get What We Get”: An Intersectional Analysis of Technology-Facilitated Violence Experienced by Racialized Women Anti-Violence Online Activists in Canada, *Canadian Journal of Law and Technology*, 19:2, 217.
- Reagle, J. (2012). Free as in sexist? Free culture and the gender gap, *First Monday*, 18(1), doi:10. 5210/fm.v18i1.4291
- Reguero, P. (2020). Medio siglo de feminismo y transfobia. In *Transfeminismo o barbarie*. Málaga: Kaótica Libros, pp. 227-240. ISBN: 978 84 12212945
- Reinig, B. A., Briggs, R. O., & Nunamaker Jr, J. F. (1997). Flaming in the electronic classroom, *Journal of Management Information Systems*, 14(3), 45-59.
- Rivera-Martín B. (2022). Hate speech towards LGBT people in Spain: Media and social audience, *Prisma Social*, 39, 213-233.
- Roca, J. and Folguera, L. (2021). Amores en movimiento: mujeres, novias y migrantes. Fuentes y cuestiones metodológicas, *Revista de demografía histórica - Journal of Iberoamerican Population Studies*, XXXIX, I, 2021, PP. 77-98
- Rodotà, S. (1997). *Tecnopolitica. La Democrazia e Le Nuove, Tecnologie Della Comunicazione*. Bari: Laterza.
- Romero, C. (2020). ¿Quién teme al transfeminismo?. In *Transfeminismo o barbarie*, pp. 17-38. Málaga: Kaótica Libros.
- Rubin, G. (1975). The Traffic in Women: Notes on the 'Political Economy' of Sex. In Reiter, R. (Ed.) *Toward an Anthropology of Women*. New York: Monthly Review Press
- Sadurní Balcells, N. and Langarita Adiego, J. (2020). Estado penal y disidencias sexo-genéricas. Apuntes para una comprensión crítica. In O. Jubany, and O. Guasch (Eds.) *Intersecciones encarnadas. [Con]textos críticos en género, identidad y diversidad*. Barcelona: Edicions Bellaterra.

- Sánchez-Sánchez, A.M., Ruiz-Muñoz, D., Sánchez-Sánchez, F.J. (2023). Mapping Homophobia and Transphobia on Social Media, *Sexuality Research and Social Policy*, [10.1007/s13178-023-00879-z](https://doi.org/10.1007/s13178-023-00879-z)
- Sandvig, C. (2015). The Social Industry, *Social Media +Society*, 1 (1): 1–4. doi:10.1177/2056305115582047
- Sassen, S. (2002). Towards a Sociology of Information Technology, *Current Sociology*, 50(3).
- Schweppe, J. and Perry, B. (2022). A continuum of hate: delimiting the field of hate studies, *Crime, Law and Social Change*, 77, 503–528 (2022). <https://doi.org/10.1007/s10611-021-09978-7>
- Segato, R. (2003). *Las estructuras elementales de la violencia*. Universidad Nacional de Quilmes.
- Selfe, C. and Meyer, P. (1991). Testing Claims for On-Line Conferences, *Written Communication*, 8(2), 163-192. <http://doi.org/10.1177/0741088391008002002>
- Serra, L. (2018a). *Las violencias de género en línea*. Pikara Magazine.
- Serra, L. (2018b) *Discurso de incitación al odio: Análisis desde los derechos humanos y pautas interpretativas*. Barcelona: IDHC - Institut de Drets Humans de Catalunya
- Shariff, S. and DeMartini, A. (2015). Defining the legal lines: egirls and intimate images. In Bailey J and Steeves V (Eds.) *eGirls, eCitizens*. Ottawa: University of Ottawa Press, pp. 282–307
- Sheffield, C. (1995). Hate Violence, in Rothenberg, P. (Ed.) *Race, Class and Gender in the United States* (3rd ed.). New York: St. Martin's Press, 432–441.
- Siapera, E. (2019). Online Misogyny as Witch Hunt: Primitive Accumulation in the Age of Techno-capitalism. In Ging, D., Siapera, E. (Eds.) *Gender Hate Online*. Palgrave Macmillan. https://doi.org/10.1007/978-3-319-96226-9_2
- Sills, S., Pickens, C., Beach, K., Jones, L., Calder-Dawe, O., Benton-Greig, P. and Gavey, N. (2016). Rape culture and social media: young critics and a feminist counterpublic, *Feminist Media Studies*, 16:6, 935-951, DOI: [10.1080/14680777.2015.1137962](https://doi.org/10.1080/14680777.2015.1137962)
- Smith, J. (2019). 'When I saw women being attacked... It made me want to stand up and fight': Reporting, responding to, and resisting online misogyny. In Lumsden, K. and Harmer, E. (Eds.) *Online othering: Exploring digital violence and discrimination on the web*. Palgrave Macmillan.
- Sobieraj, S. (2018). Bitch, slut, skank, cunt: patterned resistance to women's visibility in digital publics, *Information Communication and Society*, 21(11), 1700–1714. <https://doi.org/10.1080/1369118X.2017.1348535>
- Sobieraj, S. (2021). *Credible threat: Attacks against women online and the future of democracy*. Oxford: Oxford University Press.
- Souto Galván, B. (2015). Discurso del odio: género y libertad religiosa, *Revista General de Derecho Penal*, 23, 14.

- Spivak, G.C. (1985). The Rani of Sirmur: An essay in reading the archives, *History and Theory*, 24 (3): 247–272.
- Stanko, E. (1985). *Intimate intrusions: Women's experience of male violence*. London: Routledge.
- Stacey, J. (1988). Can there be a feminist ethnography?, *Women's Studies International Forum*, v. 11, Issue 1,
- Stanford Law School (2021). *Exploring Alternative Approaches to Hate Crimes*. Brennan Centre for Justice.
- Stoycheff, E. (2016). Under Surveillance: Examining Facebook's Spiral of Silence Effects in the Wake of NSA Internet Monitoring, *Journalism & Mass Communication Quarterly*, 93 (2): 296–311.
- Sunstein, C. (2007). *Republic.com 2.0*. Princeton. New Jersey: Princeton University Press.
- Tomlinson, J. (1999). *Globalization and Culture*. Cambridge: Polity Press.
- Treré, E. and Barranquero Carretero, A. (2018). Tracing the Roots of Technopolitics: Towards a North-South Dialogue. In Sierra Caballero, S. and Gravante, T. (Eds.) *Networks, Movements and Technopolitics in Latin America: Critical Analysis and Current Challenges*, 43–63. Cham: Springer.
- Turner, B.S. (2000). An outline of a general sociology of the body. In B.S. Turner (Ed.) *The Blackwell Companion to Social Theory*. Massachusetts and Oxford: Blackwell Publishers.
- Udupa, S., Budka, P. and E. Costa (2018). The Digital Turn: New Directions in Media Anthropology, *Follow-Up E-Seminar on the EASA Media Anthropology Network Panel "The Digital Turn"* at the 15th European Association of Social Anthropologists (EASA) Biennial Conference, Stockholm, Sweden, 14-17 August 2018
- Udupa, S. and Pohjonen, M. (2019). Extreme Speech and Global Digital Cultures, International, *Journal of Communication*, 13(2019), 3049–3067.
- Udupa, S., Gagliardone, I. and P. Hervik (2021). Introduction. In Udupa, S., Gagliardone, I. and P. Hervik (Eds.). *Digital hate: the global conjuncture of extreme speech*. Bloomington, Indiana: Indiana University Press.
- UN Broadband Commission for Digital Development Working Group on Broadband and Gender (2015). *Cyberviolence against women and girls. A world-wide wake-up call*. The Broadband Commission for Digital Development.
- Van Doorn, N. (2011). Digital spaces, material traces: How matter comes to matter in online performances of gender, sexuality and embodiment, *Media, Culture & Society*, 33(4) 531 –547.
- Van Meijl, T. (2008). Culture and identity in anthropology: reflections on 'unity' and 'uncertainty' in The Dialogical Self, *International Journal for Dialogical Science*, Vol. 3, No. 1, 165-190.

- Van Zoonen, L. (2002). Gendering the Internet: Claims, controversies and cultures, *European Journal of Communication*, 17(1), 5-23. <https://doi.org/10.1177/0267323102017001605>
- Vergès, N. (2017). *Redes sociales en perspectiva de género: Guía para conocer y contrarrestar las violencias de género online*. Instituto Andaluz de Administración Pública.
- Vergés Bosch, N. and Gil-Suarez, A. (2021). Un acercamiento situado a las violencias machistas online y a las formas de contrarrestarlas, *Revista Estudios Feministas*, 29 (3)
- Villar-Aguilés, A.; Pecourt Gracia, J. (2021). Antifeminismo y troleo de género en Twitter. Estudio de la subcultura trol a través de #STOPfeminazis. *Teknokultura. Revista de Cultura Digital y Movimientos Sociales*, 18(1), 33-44
- Wajcman, J. (2002). Addressing technological change: The challenge to social theory, *Current Sociologies*, 50(3), p.347-363.
- Wajcman, J. (2010). Feminist theories of technology, *Cambridge Journal of Economics*, 34(1), p.143-152.
- Waldron, J. (2012). *The Harm in Hate Speech*. Cambridge MA: Harvard University Press.
- Walters, M. (2019). Repairing the harms of hate crime: towards a restorative justice approach?, *University of Sussex. Journal contribution*. <https://hdl.handle.net/10779/uos.23308100.v1>
- Waltman, M. S., and Davis, M. J. (2005). How to be a proper racist: Mythic representations in The Turner Diaries, *Journal of Intergroup Relations*, 32, 19–39.
- Waltman, M. S., and Haas, J. W. (2011). *The communication of hate*. New York: Peter Lang Publishers.
- Wang, L. (1999). The complexities of 'hate', *Ohio State Law Journal*, 60,799–898.
- West, C. and Zimmerman, D. (1987). Doing Gender, *Gender and Society*, 1(2):125–151.
- Willem, C., and Tortajada, I. (2021). Gender, Voice and Online Space: Expressions of Feminism on Social Media in Spain, *Media and Communication*, 9(2), 62-71. doi:<https://doi.org/10.17645/mac.v9i2.3851>
- Willem, C., Platero, R.L., and Tortajada, I. (2022). Trans-exclusionary Discourses on Social Media in Spain. In Krijnen, T., Nixon, P.G., Ravenscroft, M.D., Scarcelli, C.M. (Eds.) *Identities and Intimacies on Social Media*. London: Routledge.
- Wilson, R. A. (2019). The digital ethnography of law, *Journal of Legal Anthropology*, 3(1), 1-20.
- Wolf, M. (1992). *A Thrice Told Tale: Feminism, Postmodernism, and Ethnographic Responsibility*. Stanford University Press.

Wood, M., Rose, E., & Thompson, C. (2019). Viral justice? Online justice-seeking, intimate partner violence and affective contagion, *Theoretical Criminology*, 23(3), 375-393. <https://doi.org/10.1177/1362480617750507>

Wren, K. (2001). Cultural racism: Something rotten in the state of Denmark?, *Social & Cultural Geography* 2(2): 141-163.

Zurita, E. (2011). Violencia de género y sociedad de consumo, *Actas de I Congreso Internacional Virtual Violencia de Género en el Contexto de la Globalización*, UNIA (Universidad Internacional de Andalucía), p. 607-625

