



#TOXICTWITTER

VIOLENCE AND ABUSE AGAINST WOMEN ONLINE

AMNESTY
INTERNATIONAL



Amnesty International is a global movement of more than 7 million people who campaign for a world where human rights are enjoyed by all.

Our vision is for every person to enjoy all the rights enshrined in the Universal Declaration of Human Rights and other international human rights standards.

We are independent of any government, political ideology, economic interest or religion and are funded mainly by our membership and public donations.

© Amnesty International 2018
Except where otherwise noted, content in this document is licensed under a Creative Commons (attribution, non-commercial, no derivatives, international 4.0) licence.
<https://creativecommons.org/licenses/by-nc-nd/4.0/legalcode>
For more information please visit the permissions page on our website: www.amnesty.org
Where material is attributed to a copyright owner other than Amnesty International this material is not subject to the Creative Commons licence.
First published in 2017
by Amnesty International Ltd
Peter Benenson House, 1 Easton Street
London WC1X 0DW, UK

Cover photo: Portrait of a young women using her personal device to access her social media.
© lechatnoir/Getty Images

Index: ACT 30/8070/2018
Original language: English
amnesty.org

AMNESTY
INTERNATIONAL 

CONTENTS

SECTION 1: A TOXIC PLACE FOR WOMEN	5
METHODOLOGY	5
A TOXIC PLACE FOR WOMEN	7
SECTION 2: TRIGGERS OF VIOLENCE AND ABUSE AGAINST WOMEN ON TWITTER	14
TRIGGERS OF VIOLENCE AND ABUSE	14
THE INTERSECTIONAL NATURE OF VIOLENCE AND ABUSE AGAINST WOMEN ONLINE	19
SECTION 3: WOMEN'S EXPERIENCES OF VIOLENCE AND ABUSE ON TWITTER	22
WHAT IS VIOLENCE AND ABUSE AGAINST WOMEN ONLINE?	22
THE MANY FORMS OF VIOLENCE AND ABUSE	23
SECTION 4: THE REPORTING PROCESS	33
TWITTER'S DUE DILIGENCE RESPONSIBILITIES	33
WOMEN'S EXPERIENCES: INCONSISTENT ENFORCEMENT OF TWITTER RULES	37
WOMEN'S EXPERIENCES: TWITTER'S INACTION ON REPORTS OF ABUSE	41
WOMEN STOP REPORTING ABUSE	44
TWITTER'S HUMAN RIGHTS FAILURES	45
SECTION 5: THE SILENCING EFFECT	46
WHY VIOLENCE AND ABUSE AGAINST WOMEN ON TWITTER IS A FREEDOM OF EXPRESSION ISSUE	46
LIMITING INTERACTIONS AND CHANGING BEHAVIOUR	48
SELF-CENSORSHIP TO AVOID ABUSE	49
WOMEN LEAVING TWITTER	50
IMPACT OF VIOLENCE AND ABUSE ON WOMEN'S PARTICIPATION IN PUBLIC LIFE	51
SECTION 6: THE PSYCHOLOGICAL HARMS OF VIOLENCE AND ABUSE AGAINST WOMEN ONLINE	54
ONLINE VERSUS OFFLINE REALITIES	54
THE TOLL OF VIOLENCE AND ABUSE AGAINST WOMEN ONLINE	55
SECTION 7: HUMAN RIGHTS RESPONSIBILITIES	61
TWITTER'S FAILURE TO RESPECT WOMEN'S HUMAN RIGHTS	61
THE DUTIES OF STATES UNDER INTERNATIONAL HUMAN RIGHTS LAW	64
SECTION 8: SOLUTIONS	66
ANNEX I: LETTER FROM TWITTER TO AMNESTY INTERNATIONAL (14 FEBRUARY 2018)	69

SECTION 1: A TOXIC PLACE FOR WOMEN

Twitter is a social media platform used by hundreds of millions of people around the world to debate, network and share information with each other. From high-level female politicians to journalists, activists, writers and bloggers, to women who simply want to know what's happening around them - Twitter can be a powerful tool for women to make connections and express themselves. In fact, the company has touted itself as a place where *'every voice has the power to impact the world'*.¹ But for many women, Twitter is a platform where violence and abuse against them flourishes, often with little accountability. As a company, Twitter is failing in its responsibility to respect women's rights online by inadequately investigating and responding to reports of violence and abuse in a transparent manner. The violence and abuse many women experience on Twitter has a detrimental effect on their right to express themselves equally, freely and without fear. Instead of strengthening women's voices, the violence and abuse many women experience on the platform leads women to self-censor what they post, limit their interactions, and even drives women off Twitter completely. At a watershed moment when women around the world are using their collective power to speak out and amplify their voices through social media platforms, Twitter's failure to adequately respect human rights and effectively tackle violence and abuse on the platform means that instead of women using their voices *'to impact the world'*, many women are instead being pushed backwards to a culture of silence.

"Online Abuse is unacceptable for women in politics, just as it's unacceptable for a woman anywhere to suffer that kind of abuse"

Nicola Sturgeon, First Minister of Scotland²



Examples of abusive tweets mentioning Nicola Sturgeon.

METHODOLOGY

Between December 2016 and March 2018, Amnesty International conducted qualitative and quantitative research about women's experiences on social media platforms including the scale, nature and impact of violence and abuse directed towards women on Twitter.

Over the course of this investigation, Amnesty International interviewed 86 women and non-binary individuals in the United Kingdom (UK) and the United States of America (USA) about their experiences on Twitter including in 5 focus group discussions with 37 young women coordinated by the University of Sussex. Given the public nature of the social media platform and the fact that many of the women interviewed are leading

¹ Twitter, *About Us*, available at https://about.twitter.com/en_us/values.html (last accessed 12 March 2018)

² Amnesty International interview with Nicola Sturgeon, 22 June 2017.

debates and opinions in various areas and industries, Amnesty International mainly interviewed female public figures including politicians, journalists, activists, bloggers, writers, comedians and games developers.

In addition to aforementioned interviews, Amnesty International conducted a qualitative survey in early 2017 which received 162 responses about women's experiences of violence and abuse on social media platforms in order to capture the experiences of female users without a large public following. 101 out of the respondents (62.35%) said they had experienced abuse on Twitter. Some of these testimonies were included in the research to help illustrate how women without a large following on Twitter are impacted by violence and abuse on the platform.

Also, in September 2017 Amnesty International commissioned a data scientist to use machine learning tools to detect and analyse online abuse against female Members of Parliament (MPs)³ active on Twitter in the UK. The analysis found that between 1 January and 8 June 2017 25,688 tweets out of 900,223 were abusive.

In addition to qualitative research methods, in November 2017, the organization published an extensive online poll in 8 countries, including the UK and USA,⁴ about women's experiences of abuse or harassment on social media platforms, including Twitter, and its impact on their freedom of expression online as well as the psychological impact of online abuse and harassment. This research was carried out by Ipsos MORI using an online quota survey of 500 women aged 18–55 in each country, via the Ipsos Online Panel system. The survey sample in each country was designed to be nationally representative of women in that country.

The UK and USA were chosen as focus countries given the heightened attention this issue has received in media and political spheres. Both countries also have a significant number of female Twitter users. Despite Amnesty International's focus on these two countries, it is important to highlight that violence and abuse against women online, on Twitter and more generally on social media, is a global issue that affects women in different countries around the world.

The research highlights the particular experiences of violence and abuse on Twitter against women of colour, women from ethnic or religious minorities, lesbian, bisexual or transgender women, non-binary individuals, and women with disabilities, to demonstrate the intersectional nature of abuse on the platform. Amnesty International also spoke with dozens of experts in the United Kingdom and United States working in the field of women's rights, identity-based discrimination, technology, and digital rights about violence and abuse against women on social media platforms. Amnesty International consulted with multiple organizations and individuals, particularly in the UK, when developing our recommendations and solutions for Twitter.

As of 16 March 2018, Amnesty International has met with Twitter in-person on three separate occasions (May 2017 in Washington, D.C, USA; February 2018 in San Francisco, California, USA; and March 2018 in London, United Kingdom). In addition to these meetings with Twitter Legal and Public Policy Experts, Amnesty International had further correspondence with the company via email and telephone and sent a letter to Twitter on 28 January 2018 outlining our concerns about violence and abuse against women on the platform and requested clarity and data related to various aspects of Twitter's operations in relation to the company's human rights responsibilities. Twitter responded to Amnesty International's general concerns (see Annex I) but refused to provide the data requested about the reporting process and content moderation. These requests were further reiterated but also denied during the aforementioned meetings with Twitter in February and March 2018.

³ Azmina Dhrodia, *Unsocial Media: The Real Toll of Online Abuse against Women*, Amnesty International, Medium, 20 November 2017, available at <https://medium.com/amnesty-insights/unsocial-media-the-real-toll-of-online-abuse-against-women-37134ddab3f4> (last accessed 13 March 2018).

⁴ Azmina Dhrodia, *Unsocial Media: The Real Toll of Online Abuse against Women*, Amnesty International, Medium, 20 November 2017, available at <https://medium.com/amnesty-insights/unsocial-media-the-real-toll-of-online-abuse-against-women-37134ddab3f4> (last accessed 13 March 2018).

On 12 March 2018, Amnesty International once again wrote to Twitter outlining our allegations of the company's failure to meet its human rights responsibilities. Twitter responded on 15 March 2018 (see Annex II).

A TOXIC PLACE FOR WOMEN

As the world becomes increasingly connected online, our reliance on social media platforms such as Twitter has also become increasingly important. But the online world, and social media platforms like Twitter are not immune to many of the human rights abuses that women face offline.

Over the last 16 months, Amnesty International has conducted qualitative and quantitative research (see Methodology above) about women's experiences on social media platforms including the scale, nature and impact of violence and abuse directed towards women on Twitter, with a particular focus on the United Kingdom (UK) and the United States of America (USA). Such abuse includes direct or indirect threats of physical or sexual violence, discriminatory abuse targeting one or more aspects of a woman's identity, targeted harassment, and privacy violations such as doxing or sharing sexual or intimate images of a woman without her consent.

Over the course of this research, Amnesty International interviewed 86 (see Methodology above) women both individually and in groups in the UK and USA. We spoke to female politicians, journalists, activists, bloggers, writers, comedians, games developers as well as women who use the platform but do not have a large following. Amnesty International also spoke with dozens of experts in the United Kingdom and United States working in the field of women's rights, identity-based discrimination, technology, and digital rights about violence and abuse against women on social media platforms. Amnesty International consulted with multiple organizations and individuals, particularly in the UK, when developing our recommendations and solutions for Twitter. The research highlights the particular experiences of violence and abuse on Twitter against women of colour, women from ethnic or religious minorities, lesbian, bisexual or transgender women - as well as non-binary individuals - and women with disabilities, to expose the intersectional nature of abuse on the platform. In November 2017, the organization commissioned Ipsos MORI to conduct an online poll in 8 countries, including the UK and US, about women's experiences of abuse or harassment on social media platforms more generally and its impact on women's freedom of expression online as well as the psychological impact of online abuse and harassment.⁵

Overall, our findings paint a worrying picture that Twitter can be a toxic place for its female users. The company's failure to meet its responsibilities regarding violence and abuse means that many women are no longer able to express themselves freely on the platform without fear of violence or abuse.

"I think Twitter is the worst of the social media platforms, just because of the quickened and masked flow [of abuse] that happens. The content feels pretty similar across the platforms but the sheer volume of it on Twitter is what's different. "

Jessica Valenti, US journalist and writer⁶

⁵ Amnesty International, *Amnesty reveals alarming impact of online abuse against women*, 20 November 2017, available at <https://www.amnesty.org/en/latest/news/2017/11/amnesty-reveals-alarming-impact-of-online-abuse-against-women/> (last accessed 12 March 2018).

⁶ Amnesty International interview with Jessica Valenti, 15 May 2017.

WHAT IS TWITTER?

Twitter operates in 43 languages and was created in 2006. On 21 March that year, co-founder and current Chief Executive Officer (CEO) Jack Dorsey⁷ sent the first tweet on the platform that read 'just setting up my twtr'.⁸ Initially conceived as a Short Message Service (SMS)-based platform, Twitter has since turned into one of the world's largest online social media platforms and has over 330 million monthly active users.⁹

- Around 500 million tweets are sent per day,¹⁰ that's 6,000 tweets per second and 200 billion tweets per year
- The approximate number of Twitter users in the UK is 20 million.¹¹ 45% of UK adults online use Twitter, with 37% of those account holders logging in daily.
- Male and female users account for 56 and 44 percent of users in the UK,¹² respectively.
- As of 2017 there are over 67 million Twitter users in US.¹³ 36% of people in the USA aged 18-29 use Twitter, more than any other age group.
- 21 % of men and women in the USA use Twitter.¹⁴

Over the last 12 years, Twitter has become a place that people use in both a professional and personal capacity. From human rights activists coordinating protests on Twitter to people from marginalized communities creating solidarity networks or politicians engaging with constituents outside of traditional office hours - Twitter offers its users around the world an opportunity to make connections across industries and regions at a speed that was almost unthinkable only a decade ago.

WHY TWITTER?

Although violence and abuse against women is certainly not limited to any one social media platform, Twitter is one of the world's largest social media companies and its platform is one of the most influential globally. Twitter is, therefore, critical in any effort to tackle violence and abuse online.

Moreover, the very nature of Twitter encourages users to have public conversations and share their thoughts with others (often strangers) on the platform, meaning that users arguably most benefit from Twitter when they are able to participate in discussions openly. In fact, Twitter's ability to provide 'up-to-the-minute reactions'¹⁵ means that public figures are able to bypass traditional media outlets and engage directly with their audiences. However, the open and public nature of these interactions also means that the platform is vulnerable to being used to send violent and abusive content. The ease and speed with which content can proliferate on Twitter means that women's experiences of violence and abuse on the platform requires an urgent and adequate response from the company.

⁷ Jack Dorsey, Twitter account (@jack), available at <https://twitter.com/jack> (last accessed 12 March 2018).

⁸ Jack Dorsey, Tweet, 21 March 2006, available at <https://twitter.com/jack/status/20> (last accessed 12 March 2018).

⁹ Twitter, *Q3 2017 Letter to Shareholders*, 26 October 2017, available at http://files.shareholder.com/downloads/AMDA-2F526X/5896920797x0x9611213D6E4631-9478-453F-A813-8DAB496307A1/Q3_17_Shareholder_Letter.pdf (last accessed 12 March 2018), p. 2.

¹⁰ Internet Live Stats, *Twitter Usage Statistics*, available at <http://www.internetlivestats.com/twitter-statistics/> (last accessed 12 March 2018).

¹¹ Kate Rose, *UK Social Media Statistics for 2017*, undated, available at <http://www.rosemcgrory.co.uk/2017/01/03/uk-social-media-statistics-for-2017/> (last accessed 12 March 2018).

¹² The Statistics Portal, *Gender breakdown of Twitter users in Great Britain (GB) from May 2014 to May 2017*, undated, available at <https://www.statista.com/statistics/278319/gender-breakdown-of-twitter-users-in-great-britain/> (last accessed 12 March 2018).

¹³ Salman Aslam, *Twitter by the Numbers: Stats, Demographics & Fun Facts*, 01 January 2018, available at <https://www.omnicoreagency.com/twitter-statistics/> (last accessed 12 March 2018).

¹⁴ Christina Newberry, *28 Twitter Statistics All Marketers Need to Know in 2018*, Hootsuite, 17 January 2018, available at <https://blog.hootsuite.com/twitter-statistics/> (last accessed 12 March 2018).

¹⁵ Jake Coyle and Associated Press, *Is Twitter the news outlet for the 21st century?*, ABC News, undated, available at <https://abcnews.go.com/Technology/story?id=7979891&page=1> (last accessed 12 March 2018).

In our initial research on violence and abuse against women on social media platforms, many women we interviewed repeatedly highlighted Twitter as a place where violence and abuse against women is widespread and where inadequate remedies exist for those who experience such abuse. Twitter itself has publicly acknowledged that they have a lot more work to do when it comes to tackling violence and abuse against women on the platform. Multiple executives from the company have stated that the platform has a problem with abuse. For example, in 2017, CEO Jack Dorsey stated,¹⁶

“We see voices being silenced on Twitter every day. We’ve been working to counteract this for the past 2 years... We prioritized this in 2016. We updated our policies and increased the size of our teams. It wasn’t enough.”

WOMEN WANT TO BE ON TWITTER

Many of the women who spoke to Amnesty International about the violence and abuse they experience on Twitter emphasized how important the platform is to them – both professionally and personally. Women rely on social media platforms like Twitter to advocate, communicate, mobilize, access information and gain visibility.¹⁷

Pamela Merritt, a US blogger and reproductive rights activist, told us,¹⁸

“I am a bit of a Twitter addict. I wake up, I check Twitter. I have two cups of coffee, and I check it again. Being online is important for my work. I want to know what’s going on. I want to know what people are saying and I want to weigh in, so I’m on Twitter through the day”.

Seyi Akiwowo, UK Politician and activist,¹⁹ talked about how Twitter makes her feel part of a movement and ‘puts words to her experiences’ of being a woman of colour. She told us,²⁰

“I feel torn. I love Twitter. The platform has connected me to people I will probably never see in my entire life and they are amazing, successful and inspirational people. There is now a massive movement of women of colour online. We express our beauty and confidence and talk about self-care and intersectional inequality... and so some of my life changing moments and development into womanhood has happened because of and via Twitter. I remember finding my first international job through Twitter. The possibilities and opportunities on that platform are endless.”

Imani Gandy, US journalist, also stressed how Twitter has become the epi-centre of communications and solidarity for marginalized communities. She explains,²¹

“I think Twitter has become the new public square. I’ve found Twitter to be a really good platform for people who normally don’t have as much of a say in the political process. I’m talking primarily young people and people of colour. There is a reason that ‘Black Twitter’ is a thing. It has been really powerful that black people have been able to come together to reach out to other black people across the country. Whether it’s dishing about the latest episode of Scandal or organizing around Black Lives Matter – it has really become a powerful tool. And elite politicians are not used to having to answer to these types of people.”

Politicians themselves agree. Scottish Parliamentarian and Leader of the Opposition, Ruth Davidson, emphasized how important Twitter is as a tool to communicate and listen to her constituents and the wider public. She notes,²²

¹⁶ Jack Dorsey, Tweet, 13 October 2017, available at <https://twitter.com/jack/status/919028949434241024> (last accessed 14 March 2018).

¹⁷ UN Human Rights Council, *Promotion, protection and enjoyment of human rights on the Internet: ways to bridge the gender digital divide from a human rights perspective*, 05 May 2017, A/HRC/35/9, available at <https://documents-dds-ny.un.org/doc/UNDOC/GEN/G17/111/81/pdf/G1711181.pdf?OpenElement> (last accessed 12 March 2018) p. 7, para. 23.

¹⁸ Amnesty International interview with Pamela Merritt, 07 July 2017.

¹⁹ Glitch! UK, website, available at <https://seyiakiwowo.com/glitchuk/> (last accessed 15 March 2018).

²⁰ Amnesty International interview with Seyi Akiwowo, 30 March 2017.

²¹ Amnesty International interview with Imani Gandy, 15 August 2017.

²² Amnesty International interview with Ruth Davidson, 21 June 2017.

“Social media platforms are where a lot of political debate now happens, particularly as we see the divide in age range of voters. For a lot of younger voters, actually – they don’t want to be told things – they want to be able to discuss them.”

Twitter remains an important tool for women across industries. US writer Chelsea Cain told Amnesty International how she relied on Twitter to talk to other writers when she moved from writing novels to writing comics. She explains,²³

“I have a lot of friends who work in comics and we kept in touch via Twitter. The comic book industry uses Twitter like a bulletin board -- it's the place everyone comes when they have an announcement or a complaint. Readers. Editors. Artists. Writers. Colourists. Letterers. Everyone I know in comics is on Twitter. The platform really fuels a connection between creators and fans. Facebook is for novels. Twitter is for comics. Don't ask me why.”

In addition to movement-building and networking, some women simply use Twitter to communicate what they are up to. Comics writer Kelly Sue DeConnick told us,²⁴

“Sometimes I use social media as a public figure to promote books and appearances. Often I use it just to be social with people, just to chat.”

But despite all the possibilities and the positive ways in which the platform is used by women on a daily basis, Twitter remains fertile ground for reinforcing existing gender inequalities and discrimination against women online. Harmful and negative gender stereotypes of women offline, as well as widespread discrimination against women rooted in patriarchal structures,²⁵ manifest as violent and abusive tweets against some women on Twitter. All forms of violence and abuse against women, both in the physical and digital world, must be seen through the lens of the systematic marginalization of women throughout society.²⁶ Violence and abuse against women on Twitter is not a new phenomenon, it is simply an extension of existing and systematic discrimination against women that has found its way into the digital sphere.²⁷

As First Minister of Scotland, Nicola Sturgeon, explains,²⁸

“There’s a link between the misogyny and abuse that women experiences offline and online... Social media is just a different way of committing these acts. Ultimately it’s the misogyny lying behind it that is the problem. So, it’s that we have to tackle, not just the means by which people are able to spread their hate and misogyny and abuse.”

UK journalist Hadley Freeman agrees. She told Amnesty,²⁹

“I think there is a connection between the misogyny women face offline and the abuse they face online. I mean, a nice guy isn't going to be calling women "infected c*nts" online, is he?”

²³ Amnesty International interview with Chelsea Cain, 09 June 2017.

²⁴ Amnesty International interview with Kelly Sue DeConnick, 28 June 2017.

²⁵ UN Committee on the Elimination of Discrimination against Women, *General recommendation No. 35 on gender-based violence against women, updating general recommendation No. 19*, 14 July 2017, CEDAW/C/GC/35, available at http://tbinternet.ohchr.org/Treaties/CEDAW/Shared%20Documents/1_Global/CEDAW_C_GC_35_8267_E.pdf (last accessed 12 March 2018) p. 6, para. 19.

²⁶ Zarizana Abdul Aziz, *Due Diligence and Accountability for Online Violence against Women*, Due Diligence Project, 2017 available at <https://www.apc.org/sites/default/files/DueDiligenceAndAccountabilityForOnlineVAW.pdf> (last accessed 12 March 2018) p. 13.

²⁷ UN Committee on the Elimination of Discrimination against Women, *General recommendation No. 35 on gender-based violence against women, updating general recommendation No. 19*, 14 July 2017, CEDAW/C/GC/35, available at http://tbinternet.ohchr.org/Treaties/CEDAW/Shared%20Documents/1_Global/CEDAW_C_GC_35_8267_E.pdf (last accessed 12 March 2018) p. 6, para. 19.

²⁸ Amnesty International interview with Nicola Sturgeon, 22 June 2017.

²⁹ Amnesty International interview with Hadley Freeman, 29 January 2018.

Women of colour, in particular, emphasized to Amnesty International that in addition to Twitter manifesting existing patriarchal structures in society, ideologies of white supremacy are also permeating into the platform.³⁰ Renee Bracey Sherman, a US abortion rights activist,³¹ explained how the violence and abuse she experiences on social media platforms went to a deeper level when she started talking about reproductive rights issues or the Black Lives Matter (BLM) movement.³² She told Amnesty,³³

“When I started talking about abortion, the abuse on Twitter went to a deeper level. When I talked about BLM, it went to a deeper level. That’s the deep issue of how much white supremacy and misogyny is really embedded in our culture. And people are upset when marginalized folks, black folks, women of colour, trans women, trans folks – when they reclaim their narrative and are unapologetic – it makes people uncomfortable. And this primal level of attack comes out and they just say such disgusting things.”



Examples of abusive tweets sent to US abortion rights activist @RBraceySherman.

Many women interviewed by Amnesty International spoke about the fact that the online world, more generally, has created spaces where people feel emboldened to say things to people that they would not necessarily say offline.

Scottish Women’s Rights Activist Talat Yaqoob explains how the sexist, racist and Islamophobic abuse she receives on Twitter is far worse than anything she has experienced offline.³⁴

“I, 100%, don’t experience the level of abuse offline that I do online. I experience it – but the frequency of it and the toxic nature of it is more online than what I experience in real life because people know they get away with it more. If people do it, it’s behind an anonymous Twitter profile. Offline, you have to be physically in front of me. I don’t think that people understand the consequences of what they say online.”

³⁰ Noah Kulwin, *Twitter is still trying and failing to solve its white supremacy problem*, Vice News, 16 November 2017, available at https://news.vice.com/en_us/article/gvdda7/twitter-is-still-trying-and-failing-to-solve-its-white-supremacy-problem (last accessed 14 March 2018).

³¹ Reproaction, Website, available at <https://reproaction.org> (last accessed 14 March 2018).

³² Black Lives Matter, Website, available at <https://blacklivesmatter.com> (last accessed 12 March 2018).

³³ Amnesty International interview with Renee Bracey Sherman, July 2017.

³⁴ Amnesty International interview with Talat Yaqoob, 23 June 2017.

TWITTER'S HUMAN RIGHTS RESPONSIBILITIES

Under the United Nations (UN) Guiding Principles on Business and Human Rights,³⁵ Twitter, as a company, has a specific responsibility to respect all human rights – including the rights to non-discrimination and freedom of expression and opinion – and to take concrete steps to avoid causing or contributing to abuses of those rights. This includes taking action to identify, prevent, address and account for human rights abuses that are linked to its operations. Specifically, as part of its human rights due diligence,³⁶ Twitter should be assessing – on an ongoing and proactive basis – how its policies and practices impact on users' right to freedom of expression and opinion as well other rights, and take steps to mitigate or prevent any possible negative impacts. It is also critical that Twitter is transparent about its policies and practices and the steps it is taking to identify and address human rights abuses.

Amnesty International contacted Twitter in January 2018 requesting that the company share information and data on how it responds to reports of violence and abuse on Twitter, including disaggregated information on the number of reports of abuse it receives, the number of reports found in violation of the Twitter rules and response times to reports of abuse. Greater transparency around the number of content moderators it employs and details on how they are trained on gender and other identity-based forms of abuse and international human rights standards was also requested. Twitter stated in a letter dated 14 February 2018 (see Annex I) that releasing such data would be uninformative and potentially misleading. Twitter stated,

“Meaningful transparency is a question we share your passion for. The question of absolute numbers of reports and the proportion of accounts that are actioned is one on which we continue to reflect. However, it is important to note that this type of information can be both uninformative and potentially misleading. One of the challenges, as we have discussed with your Amnesty colleagues previously, is that users regularly report content with which they disagree or in some cases, with the direct intent of trying to silence another user’s voice for political reasons. Moreover, there is a misperception that the volume of reports impacts our enforcement decisions, but this is not the case. In our meeting you discussed potential targets in this space, and we would be interested to hear any ideas you might have.”

Amnesty International agrees that this context is important to understand in the reporting process, however this does not negate Twitter's human rights responsibility to be transparent in how it is dealing with reports of violence and abuse on the platform nor does it take away from the fact that understanding how Twitter is responding to specific reports of abuse will only help strengthen and empower user's understanding of how such abuse is dealt with. Moreover, knowing the true scale of abuse on the platform ultimately allows everyone to help develop more concrete solutions to tackle the problem.

In a further letter to Amnesty International dated 15 March 2018 (see Annex II), Twitter reiterated that abuse and hateful conduct directed at women are prohibited on the platform, that the company had made changes to improve safety on the platform and had increased the number of actions taken against abusive accounts:

“Abuse and hateful conduct directed at women, including direct threats of violence, and harassment, are prohibited on Twitter. We have made more than 30 individual changes to our product, policies and operations in the past 16 months, all with the goal of improving safety for everyone. We now take action on 10 times the number of abusive accounts as the same time last year.”

Amnesty International acknowledges that Twitter has recently taken steps to improve women's experiences on the platform, however the organization's assessment is that these are not yet sufficient to adequately tackle the scale and nature of violence and abuse against women on Twitter.

³⁵ *UN Guiding Principles on Business and Human Rights*, 2011, available at http://www.ohchr.org/Documents/Publications/GuidingPrinciplesBusinessHR_EN.pdf (last accessed 12 March 2018).

³⁶ *UN Guiding Principles on Business and Human Rights*, 2011, available at http://www.ohchr.org/Documents/Publications/GuidingPrinciplesBusinessHR_EN.pdf (last accessed 12 March 2018), p. 19.

“We suck at dealing with abuse and trolls on the platform and we've sucked at it for years...We lose core user after core user by not addressing simple trolling issues that they face every day. I'm frankly ashamed of how poorly we've dealt with this issue during my tenure as CEO. It's absurd. There's no excuse for it.”

Dick Costello, former Twitter CEO (2015)³⁷

“Freedom of expression means little as our underlying philosophy if we continue to allow voices to be silenced because they are afraid to speak up.”

Vijaya Gadde, General Counsel (2015)³⁸

“Abuse is not part of civil discourse. It shuts down conversation and prevents us from understanding one another. No one deserves to be a target of abuse online and it doesn't have a place on Twitter.”

Jack Dorsey, current Twitter CEO (2016)³⁹

“We see voices being silenced on Twitter every day. We've been working to counteract this for the past 2 years...We prioritized this in 2016. We updated our policies and increased the size of our teams. It wasn't enough.”

Jack Dorsey, current Twitter CEO (2017)⁴⁰

“...We love instant, public, global messaging and conversation. It's what Twitter is and it's why we're here. But we didn't fully predict or understand the real-world negative consequences. We acknowledge that now, and are determined to find holistic and fair solutions.”

Jack Dorsey, current Twitter CEO (2018)⁴¹

³⁷ Nitasha Tiku and Casey Newton, *Twitter CEO: 'We suck at dealing with abuse'*, The Verge, 04 February 2015, available at <https://www.theverge.com/2015/2/4/7982099/twitter-ceo-sent-memo-taking-personal-responsibility-for-the> (last accessed 12 March 2018).

³⁸ Vijaya Gadde, *Twitter executive: Here's how we're trying to stop abuse while preserving free speech*, The Washington Post, 16 April 2015, available at https://www.washingtonpost.com/posteverything/wp/2015/04/16/twitter-executive-heres-how-were-trying-to-stop-abuse-while-preserving-free-speech/?utm_term=.fd5df2981517 (last accessed 12 March 2018).

³⁹ Ken Yeung, *Jack Dorsey on Twitter's harassment problem: 'Abuse is not part of civil discourse'*, Venture Beat, 26 July 2016, available at <https://venturebeat.com/2016/07/26/jack-dorsey-on-twitters-harassment-problem-abuse-is-not-part-of-civil-discourse/> (last accessed 12 March 2018).

⁴⁰ Jack Dorsey, Tweets, 13 October 2017, available at <https://twitter.com/jack/status/919028949434241024> (last accessed 12 March 2018), p. 19.

⁴¹ Jack Dorsey, Tweet, available at <https://twitter.com/jack/status/969234278167949313> (last accessed 15 March 2018).

SECTION 2: TRIGGERS OF VIOLENCE AND ABUSE AGAINST WOMEN ON TWITTER

Violence and abuse against women on Twitter comes in many forms and targets women in different ways. Women who experience multiple and intersecting forms of discrimination offline often find that violence and abuse online also targets their different identities. Non-binary individuals can also face targeted and misogynistic abuse online for not conforming to gender norms of male and female.

“When I was elected in 2015 and even during my election campaign, I found myself at the other end of horrific levels of abuse. And the question is: why might that be? Is everyone receiving the same levels of abuse? Is it women? Is it because I’m a BAME (Black Asian Minority Ethnic)⁴² woman?”

Tasmina Ahmed-Sheikh, former UK Politician⁴³

TRIGGERS OF VIOLENCE AND ABUSE

Women are targeted with violence and abuse on Twitter for a variety of different reasons. Sometimes it is for speaking out about certain (often feminist) issues and sometimes it is because they are public figures. A joint statement by the United Nations Special Rapporteurs on Violence against Women and Freedom of Expression highlighted how violence and abuse against women online can “*chill and disrupt the online participation of women journalists, activists, human rights defenders, artists and other public figures and private persons.*”⁴⁴

When asked about what triggers abuse against her on Twitter, UK writer Danielle Dash explained,⁴⁵

“The amount of abuse on Twitter depends on what you post. For example, if you talk about the EU, if you talk about race, if you talk about race and politics, if you talk about race and politics and gender, if you talk about rape – those are some of the triggering things that will attract the attention. It’s like a dog whistle to them.”

UK journalist Nosheen Iqbal emphasized how the abuse she receives on Twitter is often sparked by writing strong opinion pieces. She also emphasized how the abuse can spiral out of control if someone with a large following is the original perpetrator of the abuse. She notes,⁴⁶

“On Twitter, the general abuse I receive comes after I write opinion pieces. Expressing an opinion or a strong opinion will get you roasted online...The most memorable is when someone has a large following and all the followers jump in. You just don’t expect it – you shouldn’t expect it. Attitudes to women are heinous online.”

⁴² UK Government, *Ethnicity in the UK*, available at <https://www.ethnicity-facts-figures.service.gov.uk/ethnicity-in-the-uk> (last accessed 14 March 2018).

⁴³ Amnesty International interview with Tasmina Ahmed-Sheikh, 11 July 2017.

⁴⁴ UN Office of the High Commissioner, *UN experts urge States and companies to address online gender-based abuse but warn against censorship*, 08 March 2017, available at <http://www.ohchr.org/EN/NewsEvents/Pages/DisplayNews.aspx?NewsID=21317&LangID=E> (last accessed 12 March 2018).

⁴⁵ Amnesty International interview with Danielle Dash, 08 April 2017.

⁴⁶ Amnesty International interview with Nosheen Iqbal, 25 January 2018.

Sometimes, perpetrators of abuse will seek trending hashtags to target Twitter users who identify with feminist causes or concerns. For example, in October 2015, the popular feminist hashtags #TakeBackTheTech and #ImagineAFeministInternet were targeted by an organized Twitter attack with thousands of misogynistic tweets and memes.⁴⁷ The attack was in response to a tweet chat organised by the Internet Governance Forum Best Practice Forum on Countering Online Violence and Abuse meeting to discuss the impact of such violence and abuse. The organizer of the tweet chat also received an email in her personal inbox that stated the Twitter attack aimed to “destroy” the campaign.

Women from marginalized communities can also be targeted for violence and abuse on Twitter because they are viewed by some people to represent the opinions of an entire community. UK journalist and campaigner Shaista Aziz explains,⁴⁸

“I actually don’t think there are that many visible Muslim women with public platforms, so when you do have one, you become the individual that everything is targeted to.”

Women also experience violence and abuse online in response to opinions about a range of matters that are not necessarily related to feminism or gender equality. For example, UK women’s rights activist Sian Norris described how she once tweeted about how much she disliked an Amazon advertisement featuring British broadcaster Jeremy Clarkson and was told to drink floor polish as a response.⁴⁹



Example of abusive tweet sent to UK activist @sianushka

Women who are public figures are often particular targets of violence and abuse online, including on Twitter. A report⁵⁰ by the Association for Progressive Communications⁵¹ and Hivos⁵² found that prominent women bloggers, journalists and leaders are regularly subjected to violence and abuse online, especially when it is related to fields where men have been traditionally held as experts, such as gaming, politics and technology.

JOURNALISTS

The public nature of journalism and the dependency of journalists on social media platforms to report the news or express opinions about current affairs makes female journalists prime targets for targeted violence and abuse. In the words of Dunja Mijatović, former Organization for Security and Co-operation in Europe Representative on Freedom of the Media,

⁴⁷ APC WRP, *Take Action for #TakeBackTheTech and #ImagineAFeministInternet*, APC, October 2015, available at <https://www.apc.org/en/pubs/take-action-takebackthetech-and-imagineafeministin> (last accessed 12 March 2018).

⁴⁸ Amnesty International interview with Shaista Aziz, 23 October 2017.

⁴⁹ Amnesty International interview with Sian Norris, 13 April 2017.

⁵⁰ Jan Moolman, *Global Information Society Watch 2013*, APC and Hivos, available at https://www.giswatch.org/sites/default/files/violence_gisw13.pdf (last accessed 12 March 2018).

⁵¹ Association for Progressive Communications, Website, available at <https://www.apc.org> (last accessed 12 March 2018).

⁵² Hivos, Website, available at <https://hivos.org> (last accessed 12 March 2018).

“Female journalists and bloggers throughout the globe are being inundated with threats of murder, rape, physical violence and graphic imagery via email, commenting sections and across all social media... Male journalists are also targeted with online abuse, however, the severity, in terms of both sheer amount and content of abuse, including sexist and misogynistic vitriol, is much more extreme for female journalists.”⁵³

Many of the female journalists we spoke to emphasized how important Twitter, in particular, is to their work. Journalists are often expected to be active on Twitter to push stories out to their readers and to attract new audiences. In fact, journalists use Twitter to build their social capital by breaking news in their communities⁵⁴ - which translates into more readers and which can attract more advertisers and result in economic gains. Female freelance journalists we spoke to stressed that they are reliant on Twitter for networking and securing employment. For many female journalists, not being on Twitter simply isn't an option.

Due to the reliance on Twitter as a platform deeply connected to their profession, female journalists have had to come to terms with the fact that what they post on social media platforms will often be met with violent and abusive commentary. When asked about her experience of violence and abuse on Twitter, Scottish journalist Vonny Moyes told us,⁵⁵

“The majority of the abuse I receive is Twitter-based because I have a very active Twitter profile and following – part of which is necessary for the job I do. You can't really be a journalist without being on Twitter these days because it's where news breaks. Its where a lot of my work comes from...

...I guess I would say I have come to expect everything I post online, whether it's a tweet or a piece of writing to have some type of pushback. It's become as if I have had to develop combat navigation skills, not just to do my job but to be a woman occupying space on the internet.”

In some instances, threats of violence and abuse against female journalists are so severe that high-profile female journalists take multiple precautionary measures in case the threats made against them online transpire offline. For example, in September 2017, BBC Political Editor Laura Kuenssberg was assigned bodyguards while covering a political party conference.⁵⁶ In January 2018, UK Channel 4 news presenter Cathy Newman had to bring on specialist security experts to verify online threats made against her, including threats on Twitter.⁵⁷

POLITICIANS

Women in politics have also been particular targets of violence and abuse online. A study⁵⁸ by the Inter-Parliamentary Union showed that social media platforms, including Twitter, have become the number one place in which violence and abuse against women parliamentarians is perpetrated. By attempting to silence and exclude the voices of women in politics, violence and abuse online are yet another challenge to women's political engagement.⁵⁹ Violence and abuse against women Members of Parliament (MPs) in the UK have been of particular concern in recent years. For example, in January 2018, MP Anna Soubry submitted multiple

⁵³ Dunja Mijatović, *New Challenges to Freedom of Expression: Countering Online Abuse of Female Journalists*, OSCE Representative on Freedom of the Media, 2016, available at <https://www.osce.org/fom/220411?download=true> (last accessed 12 March 2018), p. 1.

⁵⁴ Alecia Swasy, *I studied how journalists used Twitter for two years. Here's what I learned*, Poynter, 22 March 2017, available at <https://www.poynter.org/news/i-studied-how-journalists-used-twitter-two-years-heres-what-i-learned> (last accessed 12 March 2018).

⁵⁵ Amnesty International interview with Vonny Moyes, 23 March 2017.

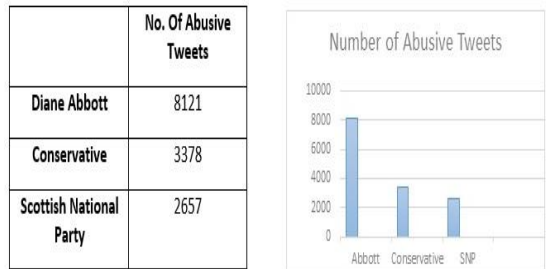
⁵⁶ Peter Walker, *BBC political editor given bodyguard for Labour conference*, the Guardian, 24 September 2017, available at <https://www.theguardian.com/media/2017/sep/24/bbc-political-editor-given-bodyguard-for-labour-conference> (last accessed 12 March 2018).

⁵⁷ Graham Ruddick, *C4 calls in security experts after presenter suffers online abuse*, the Guardian, 19 January 2018, available at <https://www.theguardian.com/society/2018/jan/19/channel-4-calls-in-security-experts-after-cathy-newman-suffers-online-abuse> (last accessed 12 March 2018).

⁵⁸ Inter-Parliamentary Union, *Sexism, harassment and violence against women parliamentarians*, October 2016, available at <https://www.ipu.org/file/2425/download?token=0H5YdXVB> (last accessed 12 March 2018).

⁵⁹ National Democratic Institute, *Stopping Violence Against Women In Politics, Program Guidance*, 2017, available at <https://www.ndi.org/sites/default/files/not-the-cost-program-guidance-final.pdf> (last accessed 12 March 2018) p. 14.

tweets to the police containing death threats against her.⁶⁰ In September 2017, Amnesty International used machine learning to measure and analyse online abuse against women MPs active on Twitter in the UK between 1 January and 8 June 2017, with a particular focus on the six weeks leading up to the UK General Election.⁶¹ The analysis reinforced the way in which online abuse targets different identities and found that Diane Abbott, the Shadow Home Secretary and first black female MP in the UK - alone - received almost half (45.14%) of all abuse against women MPs active on Twitter in the UK during this period.⁶²



Number of abusive Tweets received by Diane Abbott and two UK pollical parties.

Diane Abbott said,⁶³

“I welcome scrutiny, and I welcome engagement, and I welcome debate. That’s why I was so positive about these online platforms. But the problem is when people are not engaging in debate or scrutiny but just showering you with abuse—saying that you are a nigg*r, that you are a prostitute, threats against your safety. It’s just abuse which has no political content and which actually people wouldn’t say in a meeting or to your face.”

First Minister of Scotland, Nicola Sturgeon agrees. She told Amnesty International,⁶⁴

“I think we have to be very vigilant about genuine and legitimate criticism crossing that line and becoming unacceptable abuse. I think for politicians that threshold is higher... When that kind of commentary crosses a line and becomes threatening or just downright abusive, then I think politicians, just like anybody else, have a right to call that out and say that’s not acceptable... Online abuse poisons the well of political engagement for a vast majority of people who want to use those mediums in a really positive way.”

Amnesty International’s research also found that violence and abuse against women cuts across all political parties in the UK demonstrating that targeted violence and abuse against female politicians is a bi-partisan issue that does not pay heed to political boundaries.



NICOLA STURGEON, FIRST MINISTER OF SCOTLAND

There’s undoubtedly a gendered element to [online] abuse. Women in politics will be much more likely to get abuse about how they look, what they wear and that’s probably, to be frank, at the more innocent end

⁶⁰ Agnes Chambre, *Tory MP Anna Soubry reports Twitter threats to police after Telegraph ‘mutineers’ splash*, Politic Home, 15 November 2017, available at <https://www.politicshome.com/news/uk/foreign-affairs/brexit/news/90651/tory-mp-anna-soubry-reports-twitter-threats-police-after> (last accessed 12 March 2018).

⁶¹ Azmina Dhrodia, *Unsocial Media: Tracking Twitter Abuse against Women MPs*, Amnesty International, Medium, 04 September 2017, available at <https://medium.com/@AmnestyInsights/unsocial-media-tracking-twitter-abuse-against-women-mps-fc28aeca498a> (last accessed 12 March 2018).

⁶² Amnesty International Press release, *Diane Abbott talks about “sheer levels of hatred” she receives online*, 06 September 2017, available at <https://www.amnesty.org.uk/press-releases/diane-abbott-talks-about-sheer-levels-hatred-she-receives-online> (last accessed 12 March 2018).

⁶³ Amnesty International interview with Diane Abbott, 24 July 2017.

⁶⁴ Amnesty International interview with Nicola Sturgeon, 22 June 2017.

of it. Women in politics, and I probably speak for every woman in politics, will have seen commentary on social media which is of a sexually violent nature and that's unacceptable. It's unacceptable for women in politics, just as it's unacceptable for a woman anywhere to suffer that kind of abuse. It's not only just women, of course. I think ethnic minority people in politics will suffer abuse that is of a different nature and we've got to call that out. It's not acceptable.

The thing about people who hurl abuse on social media is that social media didn't create these people, they've always been there. It's just that social media gives them a way of getting to you that they never ever had before. We should be careful about not letting [online abuse] affect our behaviour, nor should we just shrug our shoulders and ignore it, because that contributes to the sense that somehow [abuse] is acceptable, when it's emphatically not.

What makes me angry when I read abuse about me is that I worry that it puts the next generation of young women off politics. So, I feel a responsibility to challenge it, not so much on my own behalf, but on behalf of young women out there who are looking at what people say about me and thinking, "I don't want to ever be in that position".

WOMEN'S RIGHTS ACTIVISTS

Threats, intimidation and harassment have been part and parcel of many women's experiences of standing up for human rights. The United Nations Declaration on Human Rights Defenders reaffirms rights that are essential for women's rights activists - including the rights to freedom of expression and opinion.⁶⁵ But as activism has evolved to include social media platforms as a powerful place to organize and share information about human rights abuses – threats, intimidation and harassment against women's rights activists have also moved to these same platforms. Multiple United Nations Resolutions have explicitly recognized the risk of violence and abuse and against women's rights activists online.

Women's rights activists are targeted on social media platforms like Twitter because they stand up to injustice. US blogger and reproductive rights activist Pamela Merritt shared multiple examples of how her work as a women's rights activist has led to her being targeted with threats of violence and abuse on social media platforms, including Twitter,⁶⁶

"When I did a series of blog posts about police killings of black people I got tweets and comments saying I am an 'ugly fat black b*tch and that I deserve to die' or 'you wouldn't be upset with the cops if some giant nigg*r rapes you.' If you write about abortion or about sexual violence then you are going to get a rape threat like 'you should be raped' or 'I wouldn't be upset if somebody raped you'. I did a post about black, queer women and got a bunch of comments about how 'they should f*ck the queer out of you' and 'you just haven't been f*cked well'."

Pamela Merritt's experiences are not isolated. A 2013 global monitoring survey⁶⁷ conducted by the Association for Progressive Communications⁶⁸ on sexual rights and the internet showed that while 98% of sexual rights activists see the internet as critical for their work, 51% of activists received violent and threatening messages online. A 2017 UK study conducted by the University of Northumbria titled 'Online Abuse of Feminists as An Emerging form of Violence Against Women and Girls'⁶⁹ found that Twitter was the social media platform used most commonly by respondents for feminist debate (80%). It also found that respondents experienced the most abuse⁷⁰ on Twitter with some respondents reporting that abuse started when they began to use Twitter.

⁶⁵ UN General Assembly, *Declaration on the Right and Responsibility of Individuals, Groups and Organs of Society to Promote and Protect Universally Recognized Human Rights and Fundamental Freedoms*, Resolution, 08 March 1999, A/RES/53/144 available at <https://documents-dds-ny.un.org/doc/UNDOC/GEN/N99/770/89/PDF/N9977089.pdf?OpenElement> (last accessed 13 March 2018).

⁶⁶ Amnesty International interview with Pamela Merritt, 07 July 2017.

⁶⁷ Association for Progressive Communications, *Take Back The Tech! Map it. End it.*, available at <https://www.takebackthetech.net/mapit/> (last accessed 13 March 2018).

⁶⁸ Association for Progressive Communications, Website, available at <https://www.apc.org> (last accessed 15 March 2018).

⁶⁹ Ruth Lewis, Michael Rowe and Clare Wiper, *Online Abuse Of Feminists As An Emerging Form Of Violence Against Women And Girls*, British Journal of Criminology Advance Access, 30 September 2016, available at <http://nrl.northumbria.ac.uk/27724/3/Br%20J%20Criminol-2016-Lewis-bjc-azw073.pdf> (last accessed 13 March 2018) p. 5.

⁷⁰ The study describes 'abuse' as 'harassment and sexual harassment, threats of physical and sexual violence, flaming and trolling, stalking, electronic sabotage, impersonation and defamation—and definitions, drawn from relevant contemporary research'. It also states that asking about 'general' and specific ('the last incident') experiences captured the range and specificity of abuse without focusing disproportionately

The study found that 88% of the respondents who use Twitter regularly for feminist debate had experienced abuse on the platform.

UK reproductive rights activist Dawn Purvis told Amnesty:⁷¹

“I have faced multiple layers of abuse on social media platforms, both on Twitter and on Facebook, both because I’m a woman and because of the stand I take on certain issues. On Twitter I was threatened and told that I deserved to be killed. I’ve been called a murderer, a killer, I’ve been called anti-woman - I’ve been called all sorts of names.”

It is important to stress that violence and abuse against women online is not only limited to women in the public eye. Women without a large or significant public following can also be targeted on Twitter. For example, Rachel*, a 19-year-old woman in the UK with just over 350 followers on Twitter described an experience where she posted a tweet about sexist behaviour towards her alongside a popular feminist hashtag. She detailed the level of abuse she received as a result,⁷²

“The hashtag is there to be a supportive tool and to make women feel like they are not alone. But the more popular the tweet becomes the more abuse you get. I had tweets saying ‘you are such a little b*tch’ or ‘Jimmy Saville⁷³ should have raped you’.”

THE INTERSECTIONAL NATURE OF VIOLENCE AND ABUSE AGAINST WOMEN ONLINE

Women have the right to live free from discrimination, both online and offline. The right to non-discrimination is a core provision in all international human rights standards. The UN Committee on the Elimination of Discrimination against Women states,⁷⁴

“Discrimination against women’ shall mean any distinction, exclusion or restriction made on the basis of sex which has the effect or purpose of impairing or nullifying the recognition, enjoyment or exercise by women, irrespective of their marital status, on a basis of equality of men and women, of human rights and fundamental freedoms in the political, economic, social, cultural, civil or any other field.”

Additionally, UN General Assembly Resolution 68/181 affirms that violence and abuse against women on social media platforms can be a form of systematic gender discrimination against women. It states,⁷⁵

“...information-technology-related violations, abuses, discrimination and violence against women, including women human rights defenders...with a view to discrediting them and/or inciting other violations and abuses against them, are a growing concern and can be a manifestation of systemic gender-based discrimination...”

However, any analysis of violence and abuse against women online should not solely be seen through a gender lens. Women who face discrimination because of their different identities offline often find that violence and abuse against them will target those same identities on Twitter. This is because an individual’s race, religion or sexual orientation, for example, can have just as much of an effect as gender—if not more—on how that person is treated both in the physical and digital world. In the case of online violence and abuse, women of colour, religious or ethnic minority women, lesbian, bisexual, transgender or intersex (LBTI) women, women

on experiences which might skew the data towards the ‘worst’ incidents. Responses indicate that abuse can be experienced over extended periods, so an individual ‘incident’ can consist of a single communication or of many, over weeks or months.

⁷¹ Amnesty International interview with Dawn Purvis, 11 January 2018.

⁷² Amnesty International interview with Rachel* (Pseudonym), 27 April 2017.

⁷³ Josh Halliday, *Jimmy Savile: timeline of his sexual abuse and its uncovering*, the Guardian, 26 June 2014, available at <https://www.theguardian.com/media/2014/jun/26/jimmy-savile-sexual-abuse-timeline> (last accessed 13 March 2018).

⁷⁴ UN Convention on the Elimination of All Forms of Discrimination against Women, *Article 1*, adopted on 18 December 1979, entry into force 03 September 1981, available at <http://www.ohchr.org/Documents/ProfessionalInterest/cedaw.pdf> (last accessed 13 March 2018).

⁷⁵ UN General Assembly, *Promotion of the Declaration on the Right and Responsibility of Individuals, Groups and Organs of Society to Promote and Protect Universally Recognized Human Rights and Fundamental Freedoms: protecting women human rights defenders*, Resolution, 30 January 2014, A/RES/68/181, available at <https://documents-dds-ny.un.org/doc/UNDOC/GEN/N13/450/31/PDF/N1345031.pdf?OpenElement> (last accessed 13 March 2018) p. 3.

with disabilities, or non-binary individuals who do not conform to traditional gender norms of male and female, will often experience abuse that targets them in unique or compounded way. US Professor Kimberlé Crenshaw, who coined the term ‘intersectionality’⁷⁶ explains:⁷⁷

“Intersectionality is a lens through which you can see where power comes and collides, where it interlocks and intersects. It’s not simply that there’s a race problem here, a gender problem here, and a class or LGBTQ problem there. Many times that framework erases what happens to people who are subject to all of these things.”

International human rights standards explicitly recognize the negative impact of intersecting forms of discrimination on the ability of women and girls to exercise their human rights.⁷⁸ Almost every single woman interviewed for this study who has experienced multiple or intersecting forms of discrimination offline stressed to Amnesty International that the violence and abuse they experience on Twitter reflect these same forms of discrimination.

US journalist Imani Gandy explains her experience as a woman of colour on Twitter,⁷⁹

“I get harassment as a woman and I get the extra harassment because of race and being a black woman. They will call white women a ‘c*nt’ and they’ll call me a ‘n*gger c*nt’. Whatever identity they can pick they will pick it and use it against you. Whatever slur they can come up with for a marginalized group – they use.”

Imani’s experience was echoed by many others. UK journalist Charlie Brinkhurst-Cuff simply stated,⁸⁰

‘I’ve never had abuse only because I’m a woman – it’s almost always had to do with my race’.

Scottish Parliamentarian and Leader of the Opposition Ruth Davidson told Amnesty International that the abuse she faces is both misogynistic and homophobic. She said,⁸¹

“..Because I’m openly gay – I was the first openly gay leader of a major political party in the UK- and particularly when I started, there was a lot of homophobic abuse. I have a lot of young gay followers on my Twitter, and for me it’s important to call that out. ...Every now and again, every month or so, I’ll retweet or push back on some of the homophobic abuse because I think it’s important that people see that sort of language is not acceptable – you don’t have to take it.”

US writer and presenter Sally Kohn, an openly lesbian woman, has had similar experiences on Twitter. She explained how her gender and sexual orientation both play a part in the abuse she receives. During the interview with Amnesty International she read out real-time comments she was receiving in response to a tweet she had posted a day earlier. The Twitter responses she received had nothing to do with the subject matter and included,⁸² **‘Are you a boy or a girl?’ and ‘Are you – are you just really a man?’**

⁷⁶ The concept of intersectionality was first coined by US legal academic Kimberlé Crenshaw in 1989 when she argued that the experiences of African-American women in the US were excluded from both feminist and anti-racist discourses. She argued that African-American women suffer different forms of discrimination than African-American men or white women. Understanding the experiences of African-American women requires analysing discrimination on the basis of both gender and race and how they interact, otherwise anti-racist analysis alone may risk narrowly focusing on (or prioritizing) the experiences of African-American men, and gender analysis alone may risk prioritizing the experiences of white women. As the concept has developed, intersectionality approaches recognise that everyone’s identity is made up of multiple components – including based on one’s sex, race, nationality, sexual orientation, gender identity, religion, language(s) spoken, caste, class, disability, age, marital status, place of birth, and place of residence. People may identify (or be identified) more strongly with some of these factors more than others, and this may change over time.

⁷⁷ Kimberlé Crenshaw interview with the African American Policy Forum, *Kimberlé Crenshaw on Intersectionality, More than Two Decades Later*, 08 June 2017, available at <http://www.law.columbia.edu/news/2017/06/kimberle-crenshaw-intersectionality> (last accessed 13 March 2018).

⁷⁸ UN Human Rights Council, *Impact of multiple and intersecting forms of discrimination and violence in the context of racism, racial discrimination, xenophobia and related intolerance on the full enjoyment of all human rights by women and girls*, 21 April 2017, A/HRC/35/10, available at <https://documents-dds-ny.un.org/doc/UNDOC/GEN/G17/097/49/PDF/G1709749.pdf?OpenElement> (last accessed 13 March 2018).

⁷⁹ Amnesty International interview with Imani Gandy, 15 August 2017.

⁸⁰ Amnesty International interview with Charlie Brinkhurst-Cuff, 21 April 2017.

⁸¹ Amnesty International interview with Ruth Davidson, 21 June 2017.

⁸² Amnesty International interview with Sally Kohn, 15 November 2017.

Former UK Politician Tasmina Ahmed-Sheikh also explained to Amnesty International how the abuse she receives on Twitter targets every aspect of her identity. She explained,⁸³

“I am from a Scottish Asian community. I am a Muslim. And I’m a woman. So it’s everything. It has an exponential effect, so people will pile on the abuse for a variety of different reasons. Some of them because you are all of these things, and some because you are one of these things, or two of these things, which makes it so much more difficult to deal with, because you just wonder where do I start with this?”

Irish Politician from Belfast Michaela Boyle told Amnesty International that the abuse she has received on social media platforms not only focuses on her appearance but also targets the fact that she has a disability. She told us,⁸⁴

“I have a physical disability and that has often been commented on - about how I should ‘get that disability fixed’. And that to me is derogatory. You know, I don’t care how my physical disability looks to others, I am comfortable in my own skin with it.”

Because intersectional discrimination is rooted in power and patriarchal structures, individuals who refuse to conform to gender binaries of male and female are also at risk of targeted abuse, especially if they are ‘femme’ presenting. Travis Alabanza, a trans, femme, non-binary UK actor and poet explains,⁸⁵

“I’m more interested in how we are all experiencing misogyny, and that’s why this violence online is happening to us. Not because of how we identify but because we’re all experiencing misogyny under the patriarchy.”

It is imperative that social media platforms like Twitter recognize the underlying factors that trigger online violence and abuse against people to ensure that their response both reflects and addresses these different experiences. Any approach to combatting violence and abuse on the platform that is not intersectional will only continue to further silence women from marginalized communities.

⁸³ Amnesty International interview with Tasmina Ahmed-Sheikh, 11 July 2017.

⁸⁴ Amnesty International interview with Michaela Boyle, 10 January 2018.

⁸⁵ Amnesty International interview with Travis Alabanza, 24 January 2018.

SECTION 3: WOMEN'S EXPERIENCES OF VIOLENCE AND ABUSE ON TWITTER

Violence and abuse against women on Twitter has become far too common an experience. Although people of all genders can experience violence and abuse online, the abuse experienced by women is often sexist or misogynistic in nature, and online threats of violence against women are often sexualized and include specific references to women's bodies. The aim of violence and abuse is to create a hostile online environment for women with the goal of shaming, intimidating, degrading, belittling or silencing women. The women interviewed by Amnesty International have experienced a wide spectrum of violence and abuse on Twitter which have negatively impacted on their human rights.

"Abuse on Twitter can include general nastiness or name calling (you b*tch, slut, c*nt). It can be more targeted harassment or can be more direct threats – which in the past I have had directed at my daughter. I've had my address, my tax information, as well as my phone number released."

Jessica Valenti, US journalist and writer⁸⁶

WHAT IS VIOLENCE AND ABUSE AGAINST WOMEN ONLINE?

According to the UN Committee on the Elimination of Discrimination against Women, discrimination against women includes gender-based violence, that is, "*violence which is directed against a woman because she is a woman or that affects women disproportionately, and, as such, is a violation of their human rights.*"⁸⁷ The Committee also states that gender-based violence against women includes (but is not limited to) physical, sexual, psychological or economic harm or suffering to women as well as threats of such acts.⁸⁸

International human rights standards emphasize that the concept of 'violence against women' is a form of gender-based violence. The UN uses the term 'gender-based violence against women' to explicitly recognize the gendered causes and impacts of such violence.⁸⁹ The term gender-based violence further strengthens the understanding of such violence as a societal - not individual - problem requiring comprehensive responses. Moreover, the UN Committee on the Elimination of Discrimination against Women states that a woman's right to a live free from gender-based violence is indivisible from, and interdependent on, other human rights, including the rights to freedom of expression, participation, assembly and association.⁹⁰

⁸⁶ Amnesty International interview with Jessica Valenti, 15 May 2017.

⁸⁷ UN Women, *General recommendations made by the Committee on the Elimination of Discrimination against Women*, General Recommendation No. 19, 11th session, 1992, available at <http://www.un.org/womenwatch/daw/cedaw/recommendations/recomm.htm> (last accessed 13 March 2018) para. 6.

⁸⁸ Committee on the Elimination of Discrimination against Women, *General recommendation No. 35 on gender-based violence against women, updating general recommendation No. 19*, 26 July 2017, CEDAW/C/GC.35, available at http://tbinternet.ohchr.org/_layouts/treatybodyexternal/Download.aspx?symbolno=CEDAW/C/GC/35&Lang=en (last accessed 13 March 2018) para. 14.

⁸⁹ Committee on the Elimination of Discrimination against Women, *General recommendation No. 35 on gender-based violence against women, updating general recommendation No. 19*, 26 July 2017, CEDAW/C/GC.35, available at http://tbinternet.ohchr.org/_layouts/treatybodyexternal/Download.aspx?symbolno=CEDAW/C/GC/35&Lang=en (last accessed 13 March 2018).

⁹⁰ Committee on the Elimination of Discrimination against Women, *General recommendation No. 35 on gender-based violence against women, updating general recommendation No. 19*, 26 July 2017, CEDAW/C/GC.35, available at http://tbinternet.ohchr.org/_layouts/treatybodyexternal/Download.aspx?symbolno=CEDAW/C/GC/35&Lang=en (last accessed 13 March 2018).

Violence and abuse against women on social media, including Twitter, includes a variety of experiences such as direct or indirect threats of physical or sexual violence, abuse targeting one or more aspects of a woman's identity (e.g., racism, transphobia, etc.), targeted harassment, privacy violations such as doxing – i.e. uploading private identifying information publically with the aim to cause alarm or distress, and the sharing of sexual or intimate images of a woman without her consent.⁹¹ Sometimes one or more forms of such violence and abuse will be used together as part of a coordinated attack against an individual which is often referred to as a 'pile-on'. Individuals who engage in a pattern of targeted harassment against a person are often called 'trolls'.⁹²

It is important to note that violence and abuse online can take place in many different contexts. In November 2017, Amnesty International commissioned an online poll with Ipsos MORI about women's experiences of abuse and harassment on social media platforms across eight countries including the USA and UK.⁹³ The findings showed that nearly a quarter (23%) of the women surveyed across the eight countries said they had experienced online abuse or harassment at least once, including 21% of women polled in the UK and 1/3 (33%) of women polled in the US. In both countries, 59% of women who experienced abuse or harassment said the perpetrators were complete strangers.⁹⁴

Although most of the women interviewed for Amnesty International's research received violent and abusive tweets from strangers or people unknown to them, online violence and abuse can also be used as a tactic by current or former intimate/domestic partners of women to control women and instil fear. A survey⁹⁵ conducted by the US organization National Network to End Domestic Violence⁹⁶ found that '97 percent of domestic violence programs reported that abusers use technology to stalk, harass, and control victims'. It also found that 86 percent of domestic violence programs reported that victims are harassed through social media.⁹⁷ In the UK, research on domestic online abuse⁹⁸ by domestic violence organization Women's Aid⁹⁹ found that 85% of respondents said the abuse they received online from a partner or ex-partner was part of a pattern of abuse they also experienced offline. Additionally, 50% of respondents stated that the online abuse they experienced also involved direct threats to them or someone they knew.¹⁰⁰ Of the women polled by Amnesty International who experienced abuse or harassment on social media platforms, 18% of women in the UK and 23% of women in the US said that the perpetrators of the abuse were current or former partners.¹⁰¹

THE MANY FORMS OF VIOLENCE AND ABUSE

Amnesty International's online poll found that women have experienced a variety of abuse and harassment on social media platforms, including Twitter.¹⁰² Of the women polled who had experienced abuse or harassment on social media platforms - 29% of women in the USA said they had experienced threats of physical or sexual violence, with 27% experiencing such threats in the UK. Around half of women polled who experienced abuse

⁹¹ Shiromi Pinto, *What is online violence and abuse against women?*, Amnesty International, 20 November 2017, available at <https://www.amnesty.org/en/latest/campaigns/2017/11/what-is-online-violence-and-abuse-against-women/> (last accessed 13 March 2018).

⁹² Shiromi Pinto, *What is online violence and abuse against women?*, Amnesty International, 20 November 2017, available at <https://www.amnesty.org/en/latest/campaigns/2017/11/what-is-online-violence-and-abuse-against-women/> (last accessed 13 March 2018).

⁹³ Azmina Dhrodia, *Unsocial Media: The Real Toll of Online Abuse against Women*, Amnesty International, Medium, 20 November 2017, available at <https://medium.com/amnesty-insights/unsocial-media-the-real-toll-of-online-abuse-against-women-37134ddb3f4> (last accessed 1 March 2018).

⁹⁴ Amnesty International, *Amnesty reveals alarming impact of online abuse against women*, 20 November 2017, available at <https://www.amnesty.org/en/latest/news/2017/11/amnesty-reveals-alarming-impact-of-online-abuse-against-women/> (last accessed 12 March 2018).

⁹⁵ National Network to End Domestic Violence, *Recognizing and Combating Technology-Facilitated Abuse*, 13 October 2016, available at https://nnedv.org/latest_update/combating-technology-facilitated-abuse/ (last accessed 13 March 2018).

⁹⁶ National Network to End Domestic Violence, website, available at <https://nnedv.org> (last accessed 13 March 2018).

⁹⁷ National Network to End Domestic Violence, *Recognizing and Combating Technology-Facilitated Abuse*, 13 October 2016, available at https://nnedv.org/latest_update/combating-technology-facilitated-abuse/ (last accessed 13 March 2018).

⁹⁸ Women's aid, *Online and digital abuse*, undated, available at <https://www.womensaid.org.uk/information-support/what-is-domestic-abuse/online-safety/> (last accessed 13 March 2018).

⁹⁹ Women's aid, website, available at <https://www.womensaid.org.uk> (last accessed 13 March 2018).

¹⁰⁰ Women's aid, *Online and digital abuse*, undated, available at <https://www.womensaid.org.uk/information-support/what-is-domestic-abuse/online-safety/> (last accessed 13 March 2018).

¹⁰¹ Azmina Dhrodia, *Unsocial Media: The Real Toll of Online Abuse against Women*, Amnesty International, Medium, 20 November 2017, available at <https://medium.com/amnesty-insights/unsocial-media-the-real-toll-of-online-abuse-against-women-37134ddb3f4> (last accessed 13 March 2018).

¹⁰² Azmina Dhrodia, *Unsocial Media: The Real Toll of Online Abuse against Women*, Amnesty International, Medium, 20 November 2017, available at <https://medium.com/amnesty-insights/unsocial-media-the-real-toll-of-online-abuse-against-women-37134ddb3f4> (last accessed 13 March 2018).

or harassment said that the abuse included sexist or misogynistic comments (53% in the USA and 47% in the UK).



Results of Amnesty International's poll from the UK and US on abuse and harassment against women online.

THREATS OF VIOLENCE

Threats of violence against women online includes both direct and indirect threats that can be physical or sexual in nature. Several women who spoke to Amnesty International about their experience of violence and abuse online reported receiving threats of violence on Twitter. For example, UK women's rights activist and writer Laura Bates has experienced multiple forms of sexually violent threats against her on Twitter. She told Amnesty,¹⁰³

“Online abuse began for me when I started the Everyday Sexism Project¹⁰⁴—before it had become particularly high-profile or I received many entries. Even at that stage, it was attracting around 200 abusive messages on the site per day. The abuse then diversified into other forums, such as Facebook and Twitter messages. These often spike if I’ve been in the media... You could be sitting at home in your living room, outside of working hours, and suddenly someone is able to send you a graphic rape threat right into the palm of your hand.”



Examples of abusive tweets mentioning @EverydaySexism directed to Founder of Everyday Sexism Project Laura Bates.

UK political comedian Kate Smurthwaite told Amnesty International about a pile-on of violence and abuse against her on Twitter following a media appearance on a television debate programme. She told us,¹⁰⁵

“..After the debate he continued to be rude about me on Twitter. That reached a whole new level. In the following 48 hours, I received 165 pages of Twitter abuse. Suddenly it went insane. In that, there were four or five death threats, rape threats, and things like that.”

¹⁰³ Amnesty International interview with Laura Bates, 25 January 2017.

¹⁰⁴ The Everyday Sexism Project, website, available at <http://everydaysexism.com> (last accessed 13 March 2018).

¹⁰⁵ Amnesty International interview with Kate Smurthwaite, 24 March 2017.

UK poet and actor Travis Alabanza told Amnesty International that much of the violence and abuse they experienced on Twitter was from people telling them to die. Travis explained,¹⁰⁶

“A lot of it was ‘die’. It was a mix, a lot of ‘I wish you people wouldn’t exist’. ‘Go create your own world’ or ‘set up your own country’. And ‘Die’. ‘Die’. ‘Die’. ‘Die.’ Lots of death.”



Example of abusive tweet mentioning UK poet and actor @travisalabanza

UK writer Danielle Dash summed up her experience of violent threats on Twitter. She said,¹⁰⁷

“The violence is at the intersection of everything that I am – for example - ‘I’m going to rape you, you black b*tch’. You have the misogyny, and you have the racism and you have the sexual violence all mixed up into one delicious stew of cesspit shit.”

UK journalist Allison Morris told Amnesty International about threats she has received against herself as well as her family. She stated,¹⁰⁸

“Some of the things that have been put on Twitter about me have had people say they know where I live, I’ve had people say that they’ll be outside my work, I’ve had people not just threaten me but also say things that, you know, are clearly veiled threats against my family.”

Threats of violence and abuse online can also have a profound impact on women’s sense of safety offline. Amnesty International’s online poll found that of the women who experienced abuse or harassment online, 42% in the USA and 36% in the UK said it made them feel that their physical safety was threatened. 1 in 5 of women in the UK (20%) and over 1 in 4 (26%) in the USA said they felt their family’s safety was at risk after experiencing abuse or harassment on social media platforms.

US reproductive rights activist Pamela Merritt echoed these findings. She said,¹⁰⁹

“After five years of online harassment coupled with offline harassment, I have basically reconciled with the fact that I’m prepared to die for the work I do. That might happen. If you get 200 death threats, it only takes one person who really wants to kill you.”

US journalist and writer Jessica Valenti told Amnesty International about how difficult it is to assess the seriousness of threats made against her online. She explained,¹¹⁰

“When you’re in real life you decipher what is a real threat and what is not. Should I cross to the other side of the street, or should I tell this person to f*ck off? You can make informed decisions in that moment. You can’t do that

¹⁰⁶ Amnesty International interview with Travis Alabanza, 24 January 2018.

¹⁰⁷ Amnesty International interview with Danielle Dash, 08 April 2017.

¹⁰⁸ Amnesty International interview with Allison Morris, 10 January 2018.

¹⁰⁹ Amnesty International interview with Pamela Merritt, 07 July 2017.

¹¹⁰ Amnesty International interview with Jessica Valenti, 15 May 2017.

online because you don't know where or who that person is. Is this person a real threat or is this person a 12 year old? You have no clue."

Women who have experienced threats of violence on social media platforms, including Twitter, also told Amnesty International about the precautionary measures they have taken to protect their families from this abuse. One woman told Amnesty that she changed her child's last name at school so the child could not be identified as someone related to her and targeted with abuse. Another woman told Amnesty International that she turned down media appearances once her pregnancy became visible because she was terrified of any abuse or violence online targeting the baby. Her particular fear of such violence and abuse was triggered by previous threats of sexual violence towards her sister on Twitter.

SEXIST AND MISOGYNISTIC ABUSE

Although different levels of abuse on Twitter will require different responses from the platform (see Section 7: Solutions), Amnesty International's research has found that all forms of abuse against women can have a harmful impact on women's rights online. Sexist and misogynistic abuse against women on Twitter was highlighted by almost every woman interviewed by Amnesty International. Such abuse includes offensive, insulting or abusive language or images directed at women on the basis of their gender and is intended to shame, intimidate or degrade women. Sexist or misogynistic abuse often includes references to negative and harmful stereotypes against women and can include gendered profanity.

Politician and Former Leader of the Labour Party in Scotland Kezia Dugdale told Amnesty International about the underlying misogyny in tweets she has received,¹¹¹

"It's definitely the case that I get more sexist commentary on Twitter and online than men. In Scotland the phrase would be 'Daft wee lassie complex'. It means she doesn't know what she's talking about - she's too young, too female to really understand what she's going on about. So people will question your intelligence by referring to your gender. That's probably the most common theme."

UK activist Alex Runswick also spoke about the variety of misogynistic abuse she has experienced on Twitter. She recounted a particular wave of abuse she received after she posted a photo of a letter that had assumed she was a man. She said,¹¹²

"Because my name is Alex, I often get mis-gendered. I took a photo of the letter and put it on Twitter and said 'Just because I'm a Director does not mean that I am male!' and used the Everyday Sexism hashtag. And then I got, for me, an enormous amount of abuse for something I didn't expect to be remotely controversial...It started with straightforward anti-women stuff around 'you shouldn't be doing politics', 'you should be in the kitchen', 'go make me a cup of tea love'...but then it quickly moved into what sort of sex acts I should be performing, my appearance and whether they 'would or wouldn't'".

Women also told Amnesty International that they often receive a spike in violence or abuse online following a television or media appearance. For example, UK science broadcaster, writer and educator Dr Emily Grossman told Amnesty International about a barrage of abuse she experienced on Twitter following an appearance in a TV debate. She explained to Amnesty International how she categorized the scale of abuse she received,¹¹³

"There were personal attacks on me and my appearance, there was sexually abusive and aggressive language – no rape threats or death threats – but certainly people talking about their cock and slapping it around my face, what they wanted to 'do' to me, tearing me a new arsehole. Then there were these comments on my qualifications and my career undermining me as a scientist. There was some really awful anti-Semitism saying that Hitler was right...And there was a category of messages that seemed to be attacking me as a representative of all women –

¹¹¹ Amnesty International interview with Kezia Dugdale, 22 June 2017.

¹¹² Amnesty International interview with Alex Runswick, 27 April 2017.

¹¹³ Amnesty International interview with Emily Grossman, 18 July 2017.

saying that women weren't clever enough to be scientists, that we were stupid, illogical, irrational, if you can't stand the heat get back in the kitchen, or if women aren't succeeding blame it on their DNA. And then there were comments saying I must be a feminist and be crazy and asking why I hate men or suggesting that maybe my uncle raped me..."

OTHER FORMS OF IDENTITY-BASED ABUSE

Sometimes, online abuse focuses solely on an aspect of a person's identity other than their gender. Under international human rights legal frameworks, the right to non-discrimination covers multiple protected characteristics of a person's identity including standards to tackle discrimination against women, minorities, racial and religious discrimination and discrimination based on sexual orientation or gender identity.¹¹⁴

A 2017 [report](#)¹¹⁵ by LGBT organization Stonewall¹¹⁶ commissioned by YouGov surveyed more than 5000 LGBT people across England, Scotland and Wales found that 10% of LGBT people experienced homophobic, biphobic and transphobic abuse or behaviour online in the last month. This number increased to one in four trans people (26 %) who experienced transphobic abuse or behaviour online in the last month. Moreover, the study¹¹⁷ found that 23% of LGBT young people aged 18 to 24 were personally targeted online in the last month, with this number rising to 34% of trans young people. The study also found that 20% of Black, Asian, Minority, Ethnic (BAME) LGBT people experienced abuse online in the last month compared to 9% of white LGBT people. Non-binary LGBT people were found to be significantly more likely than LGBT men and women to experience personal online abuse with 26% experiencing such abuse.



RUTH DAVIDSON, PARLIAMENTARIAN AND LEADER OF THE OPPOSITION

There is an issue, I think, in online debate and discussion where it does go from criticism of a policy to straight outright abuse; being dismissive of somebody because of their appearance, because of their sexuality or because of their gender. All the 'death threat' stuff, the 'wish you were dead' things that we've had – that's clearly crossed a line into abuse. I'm openly gay [and] particularly when I started, there was a lot of homophobic abuse. I have a lot of young gay followers on my twitter platform, and for me it's always been quite important to call that out. Every now and again... I'll retweet or push back on some of the homophobic abuse because I think it's important that people see that that sort of language is not acceptable – you don't have to take it.

The sheer volume [of the abuse] can sometimes make you feel a bit hunted online. At that point you just stop reading the mentions and just use it as a transmit function, rather than a transmit and receive function.

Have I bothered to report? I think in the beginning I probably did and not much happened, so I don't now, which maybe isn't the best example to set. I think that if we do want platforms to act in a more robust manner then we've got to keep up the level of complaint. I think that [online abuse] makes the whole online arena less warm, less engaging and less open for people – and that's a shame because it's a tremendous tool.

Just because you're saying something on a keyboard and not to someone's face, doesn't mean it doesn't matter. There needs to be an understanding of the seriousness of what this is – rather than the kind of frivolous 'Oh they only said it on Twitter, so it doesn't matter'. Actually, it does.

¹¹⁴ See for example, *Universal Declaration of Human Rights*, 10 December 1948, available at http://www.un.org/en/udhrbook/pdf/udhr_booklet_en_web.pdf (last accessed 14 March 2018), Article 2 ; *International Covenant on Civil and Political Rights*, adopted 16 December 1966, entry into force 23 March 1976, available at <https://treaties.un.org/doc/publication/unt/volume%20999/volume-999-i-14668-english.pdf> (last accessed 14 March 2018) Article 26 ; *International Convention on the Elimination of All Forms of Racial Discrimination*, adopted 21 December 1965, entry into force 04 January 1969, available at <http://www.ohchr.org/EN/ProfessionalInterest/Pages/CERD.aspx> (last accessed 14 March 2018) Article 1(1).

¹¹⁵ Stonewall AND YouGov, *LGBT IN BRITAIN: Hate Crime and Discrimination*, 2017, available at http://www.stonewall.org.uk/sites/default/files/lgbt_in_britain_hate_crime.pdf (last accessed 15 March 2018).

¹¹⁶ Stonewall, Website, available at <http://www.stonewall.org.uk> (last accessed 15 March 2018).

¹¹⁷ Stonewall AND YouGov, *LGBT IN BRITAIN: Hate Crime and Discrimination*, 2017, available at http://www.stonewall.org.uk/sites/default/files/lgbt_in_britain_hate_crime.pdf (last accessed 15 March 2018).

Additionally, a US study¹¹⁸ by the Pew Research Centre on online harassment found that 59% of black internet users said they had experienced online harassment compared with 41% of white internet users and 48% of Hispanic internet users. 38% of black internet users also said they had been called offensive names. US activist Shireen Mitchell described to Amnesty International how abuse against black people on Twitter often references animal names. She described how black people will sometimes be called 'ape', 'gorilla' or 'monkey' as a form of racial abuse online. Similarly, UK Politician and activist Seyi Akiwowo told Amnesty International that the abuse she experienced on Twitter and other social media platforms included racial slurs like 'n*gger', 'n*ggerress', 'negro', references to lynching and being hanged, as well as 'monkey', 'ape' and being told to 'die of an STI'.¹¹⁹



Examples of abusive tweets mentioning UK Politician and activist @seyiakiwowo and @DanielleDASH.

US-born and UK-based journalist Hadley Freeman, told Amnesty International that most of the abuse she experiences on Twitter is actually anti-Semitic - but that she receives misogynistic abuse on the platform as well.¹²⁰ In addition, Jaclyn Friedman, a US writer and activist of Jewish descent, spoke to Amnesty International about a tide of anti-Semitic abuse she experienced on Twitter in November 2016. She said,¹²¹

“I got numerous threats on Twitter threatening me with Zyclon B – that’s the gas they used to kill people in the gas chambers of the Holocaust.”

Rani Baker, a US writer and illustrator, told us about the abuse she experiences on Twitter as a trans woman on the platform. She said,¹²²

“People have made so many dehumanizing and humiliating assumptions about, references to, and descriptions of, my body, surgical results, sexual orientation and proclivities, general lifestyle and behaviours that it could fill a book. It’s shockingly common to see the most degrading descriptions of myself and my existence being bandied around by people trying to get under my skin.”

It is important to note that identity-based abuse can be used to target women from marginalized groups in different ways. A US sex worker and advocate who wished to remain anonymous, told Amnesty International that abuse on Twitter targeting sex workers often includes deliberately being ‘outed’ with a view to shaming or humiliating them. She explains,¹²³

¹¹⁸ Maeve Duggan, *1 in 4 black Americans have faced online harassment because of their race or ethnicity*, Pew Research Center, 25 July 2017, available at <http://www.pewresearch.org/fact-tank/2017/07/25/1-in-4-black-americans-have-faced-online-harassment-because-of-their-race-or-ethnicity/> (last accessed 13 March 2018).

¹¹⁹ Amnesty International interview with Seyi Akiwowo, 30 March 2017.

¹²⁰ Amnesty International interview with Hadley Freeman, 29 January 2018.

¹²¹ Amnesty International interview with Jaclyn Friedman, 14 July 2017.

¹²² Amnesty International interview with Rani Baker, 15 May 2017.

¹²³ Amnesty International interview with sex worker advocate, anonymous, 26 September 2017.

“When dealing with a criminalized and stigmatized population being attacked by people who are not in that population; there is always a question of power...Twitter being an open space is a problem for targeted abuse against sex worker advocates. For me the fear of being outed means I couldn't advocate effectively. Being outed is something I've seen over and over again on Twitter. People live in constant fear of being outed non-consensually...It's really hard to do advocacy when you are waiting for the other shoe to drop”

DOXING

Abuse against women on Twitter can also include ‘doxing’ (slang for ‘docs’ or ‘documents’) which involves revealing personal or identifying documents or details about someone online without their consent. This can include personal information such as a person’s home address, real name, children’s names, phone numbers and email address. Doxing is a violation of a person’s privacy and the aim is to distress, panic and otherwise cause alarm. In the USA, Amnesty International’s online poll found that almost 1 in 3 (29%) women who experienced abuse or harassment on social media platforms had been doxed.

Former UK Politician Tasmina Ahmed Sheikh told Amnesty,¹²⁴

“Somebody thought it was a really good idea to tweet out my home address with post code which meant the police then had to patrol my house. It was at a point where my husband and I were out and my kids were at home on their own so it was really worrying.”

US writer and activist Jaclyn Friedman told Amnesty International about the security measures she took before she published a report¹²⁵ about abuse on Twitter out of fear of being doxed,¹²⁶

“Before we launched our Twitter report,¹²⁷ I got a new security system on my house. We lived in a really safe neighbourhood and I'd never thought about it before – but I didn't want to go to bed at night and think ‘what if they dox me at 3am when I'm sleeping? It felt too vulnerable....

...It's time and energy and actual money that goes into making sure that I can say to Amnesty in this interview – ‘oh sure, publish my name, publish my face’. I consistently resent that...It's absolutely a tax on women's speech’.”

The fear of being doxed is a particular concern for women in the USA where publishing a person’s address online can lead to ‘swatting’ – which is when someone makes a fake emergency call to trigger a large police response at the target’s home. The response often includes a heavily armed SWAT (Special Weapons and Tactics) team who arrive expecting a hostile situation. For example, in January 2016, US Congresswoman Katherine Clark, who is a vocal proponent of legislation against online abuse in the USA was ‘swatted’ - soon after she sponsored a bill to combat swatting.¹²⁸

SHARING SEXUAL AND PRIVATE IMAGES WITHOUT CONSENT

Sharing sexual or private images without consent is a violation of women’s right to privacy and is usually carried out by an ex-partner with the aim of distressing, humiliating or blackmailing a woman. While a woman may have initially consented to having images taken and voluntarily shared them with an individual, she may not have given that person permission to share the images more widely. It is the non-consensual aspect of this form of abuse which makes it distinct from sexually explicit content online more broadly.

¹²⁴ Amnesty International interview with Tasmina Ahmed-Sheikh, 11 July 2017.

¹²⁵ WAM! *Twitter's Abuse Problem: Now With Actual Solutions And Science*, undated, available at <https://womenactionmedia.org/twitter-report/> (last accessed 15 March 2018).

¹²⁶ Amnesty International interview with Jaclyn Friedman, 14 July 2017.

¹²⁷ Matias, J. N., Johnson, A., Boesel, W. E., Keegan, B., Friedman, J., & DeTar, C., *Reporting, Reviewing, and Responding to Harassment on Twitter*, Women, Action, and the Media, 13 May 2015, available at <https://womenactionmedia.org/cms/assets/uploads/2015/05/wam-twitter-abuse-report.pdf> (last accessed 16 March 2018).

¹²⁸ Hannah Levintova, *This Congresswoman Has Plans to Stop Online Harassment*, Mother Jones, September/October 2016 issue, available at <https://www.motherjones.com/politics/2016/09/katherine-clark-fight-against-internet-trolls-gamergate/> (last accessed 13 March 2018).

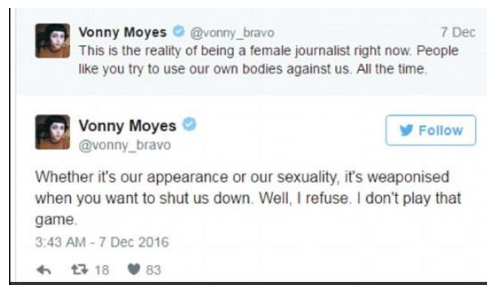
Of the women polled by IPSOS Mori for Amnesty International who experienced abuse or harassment on social media platforms, 10% in the USA and 8% in the UK said that intimate images had been posted of them online without their consent.¹²⁹

When UK journalist Vonny Moyes had private images of her shared on Twitter without her consent, she used the same platform to hold the perpetrator to account. Her response included a series of tweets that stated,

“So @*** has just found and posted nudes of me. I would very much appreciate your help in reporting him for this.”**

“The thing is @*** – this only works as weaponry if I accept the shaming. I did not give you those or permission to look at me.”**

“This is the reality of being a female journalist right now. People like you try to use our own bodies against us. All the time.”



Tweets by UK journalist @vonny_bravo highlighting violence and abuse online against women journalists.

TARGETED HARASSMENT

Targeted harassment online involves one or more people working together to repeatedly target a woman with violence or abuse over a short or coordinated period of time with the aim of humiliating her or otherwise causing distress.

Games Developer Zoe Quinn experienced what has become one of the most well-known cases of targeted harassment online. The term ‘Gamergate’¹³⁰ was coined in 2014 following a relentless flood of online violence and abuse that she and other prominent women in the gaming industry were subjected to on multiple social media platforms, including Twitter. The extensive violence and abuse targeting Zoe began after her ex-partner posted an article about their relationship and accused her of an affair. The post was picked up on platforms like Reddit and 4Chan and resulted in a barrage of violence, abuse and harassment against her.

¹²⁹ Azmina Dhrodia, *Unsocial Media: The Real Toll of Online Abuse against Women*, Amnesty International, Medium, 20 November 2017, available at <https://medium.com/amnesty-insights/unsocial-media-the-real-toll-of-online-abuse-against-women-37134ddab3f4> (last accessed 13 March 2018).

¹³⁰ Jennifer Vineyard, *You Don't Need to Care About Video Games to Care About Gamergate*, Vulture, 16 October 2014, available at <http://www.vulture.com/2014/10/why-you-should-care-about-gamergate.html> (last accessed 13 March 2018).



ZOE QUINN, US GAMES DEVELOPER

I had been dating this guy for 5 months. We broke up and he wrote a manifesto – a sort of call to arms – about me and posted it in online communities that had a reputation for targeting women. The abuse I [then] experienced on Twitter included rape threats and death threats – people calling me ‘a nasty slut’, telling me to ‘die slowly’, saying that I ‘would f*ck anybody’. I got tweets from a guy who said that he was carrying a bottle of disinfectant for me to swallow in case he were to see me at a games conference. People were sharing drives full of images and stuff about me to post online – to hurt me. They mass reported my Twitter account and work account [to get them shut down]. My blog was hacked and it was linked to my Twitter account. They posted my home address, phone number, dad’s address and dad’s phone number on Twitter. They changed my passwords [to lock me out of my account]. People found nude photos of me that they spread onto Twitter and other platforms – and people would print these photos off, masturbate on them and send them to me. People would also text me photos of my house or tweet about dead animals they put in my mailbox.

Every single aspect of my life was impacted by the abuse. People [online] pushed me really hard to kill myself. My partner at the time didn’t leave my side for more than a few hours. For the first few days I couldn’t eat or sleep or drink water. All I could do was watch everything collapse around me. It was and it still is hard to get closer to new people [after going through that].

People online pushed me really hard to kill myself. My partner at the time didn’t leave my side for more than a few hours. For the first few days I couldn’t eat or sleep or drink water. All I could do was watch everything collapse around me. It was and it still is hard to get closer to new people [after going through that].

I don’t report shit to Twitter anymore, the reports are just ignored. I gave up reporting to Twitter a very long time ago.

The very nature of the Internet allows content posted on social media platforms to ‘go viral’ and be shared across platforms almost instantly - which means that violent and abusive content can be easy to share or repeat and difficult to contain. One abusive or violent tweet against an individual, for example, can quickly multiply into hundreds or thousands of abusive or violent tweets against that individual within minutes.

UK actor and poet Travis Alabanza told Amnesty International about the targeted harassment they experienced in November 2017 after a tweet they posted went viral. Travis recounted,¹³¹

“I tweeted before I went to bed and woke up and was like ‘why do I have 700+ notifications on Twitter? A few days later Twitter sent me a tweet saying ‘Congratulations your tweet is featured in the Moments!’ ...I’m used to Twitter traffic but something felt a bit different about this. We’re talking, I think 100 tweets per minute. There were points where you could scroll my name and you could scroll 7 or 8 times and you’d only be 3 minutes in...

...I remember just looking at the abuse and being so shocked and just sitting there and my friends being like ‘What do you need?’ and me just saying ‘I don’t know.’... A lot of the abuse was to die and much of it said ‘You’re a man’. They also took photos of me and then circled parts of my face that are ‘manly’.”

Many women we spoke to emphasized that the sheer volume of violent and abusive tweets they receive on Twitter is specifically what they find so overwhelming about the platform.

US writer and presenter Sally Kohn explained,¹³²

“The abuse on Twitter is sort of constant. And it’s disturbing when you recognize that is so kind of constant and normalized that you don’t even notice it anymore...I can’t find the reasonable stuff amongst the trolls, I really can’t.”

¹³¹ Amnesty International interview with Travis Alabanza, 24 January 2018.

¹³² Amnesty International interview with Sally Kohn, 15 November 2017.

When UK Shadow Home Minister Diane Abbott told Amnesty International about the difference between offline abuse she received when she first became a Member of Parliament, and the abuse she receives on social media platforms 30 years later, she stated,¹³³

“It’s the volume of it which makes it so debilitating, so corrosive, and so upsetting. It’s the sheer volume. And the sheer level of hatred that people are showing.”



DIANE ABBOTT, UK POLITICIAN AND SHADOW HOME SECRETARY

I welcome scrutiny, and I welcome engagement and I welcome debate that’s why I was so positive about these online platforms. But the problem is when people are not engaging in debate or scrutiny, but just showering you with abuse. That you are a n*gger, that you are a prostitute, threats against your safety... It’s just abuse which has no political content and which actually people wouldn’t say in a meeting or to your face. I think the distinction between abuse and genuine political debate is, would they say it if they met you in the street? No they wouldn’t.

[The online abuse I get] is highly racialized and it’s also gendered because people talk about rape and they talk about my physical appearance in a way they wouldn’t talk about a man. I get a double whammy. I’m abused as a female politician and I’m abused as a black politician. And also the volume of abuse is much greater. It’s the volume of it which makes it so debilitating. So corrosive and so upsetting. It’s the sheer volume. And the sheer level of hatred that people are showing.

Online abuse does damage you, it damages your confidence, it corrodes your self-esteem and it can make you second guess yourself as to what you talk about and what you write about. In terms of Twitter, I now spend much less time on Twitter than I used to, because the abuse is so terrible.

[Online abuse] is not free speech, it’s actually limiting everyone else’s free speech, because in my experience there are many women, and many women of colour who don’t participate online in the way that they would want to, because they’re really repelled by the level of abuse you get. I think the abuse I get online makes younger women of colour very hesitant about entering the public debate and thinking about going into politics. And after all, why should you have to pay that price for being in the public space?

¹³³ Amnesty International interview with Diane Abbott, 24 July 2017.

SECTION 4: THE REPORTING PROCESS

Twitter has a responsibility to take concrete steps and actions to address any human rights abuses occurring on its platform. Part of this requires Twitter to have policies that are compliant with international human rights standards and to ensure that it has a robust reporting process that enables users to easily report any instances of violence and abuse. Yet, despite having policies that explicitly state that hateful conduct and abuse will not be tolerated on the platform, Twitter appears to be inadequately enforcing these policies when women report violence and abuse. At times, it fails to even respond to women who have taken the time to report abusive content about what action, if any, has been taken. Twitter's inconsistency and inaction on its own rules not only creates a level of mistrust and lack of confidence in the company's reporting process, it also sends the message that Twitter does not take violence and abuse against women seriously - a failure which is likely to deter women from reporting in the future.

"I gave up on reporting to Twitter a long time ago."

Zoe Quinn, US games developer¹³⁴

TWITTER'S DUE DILIGENCE RESPONSIBILITIES

According to the United Nations Guiding Principles on Business and Human Rights,¹³⁵ Twitter must identify, prevent, address and account for human rights abuses in its operations. In doing so, it must express and embed its commitment to human rights both through its policies *and* practices. The company must take a number of steps to meet its human rights responsibilities - the first being the development of policy commitments that incorporate and recognize international human rights standards.

TWITTER POLICIES

Twitter currently has no publically available human rights policy stating its commitment to respect human rights in its operations. The UN Guiding Principles state¹³⁶ that such a policy must also identify who is in charge of overseeing, implementing and monitoring Twitter's human rights commitments. Twitter does have a number of policies¹³⁷ related to addressing violence and abuse on the platform. In December 2015, Twitter introduced its policy on 'Hateful Conduct and Abuse'. More recently, following the suspension of actor Rose McGowan's account in November 2017 and #WomenBoycottTwitter going viral, current CEO Jack Dorsey promised a "more aggressive stance" on tackling abuse. He expedited the expansion and enforcement of Twitter rules to include unwanted sexual advances, intimate media, hateful imagery and display names, and violence.

¹³⁴ Amnesty International interview with Zoe Quinn, 13 December 2017.

¹³⁵ *UN Guiding Principles on Business and Human Rights*, 2011, available at http://www.ohchr.org/Documents/Publications/GuidingPrinciplesBusinessHR_EN.pdf (last accessed 13 March 2018), p. 24.

¹³⁶ *UN Guiding Principles on Business and Human Rights*, 2011, available at http://www.ohchr.org/Documents/Publications/GuidingPrinciplesBusinessHR_EN.pdf (last accessed 13 March 2018).

¹³⁷ Twitter, *Help Center, Rules and policies*, available at <https://help.twitter.com/en/rules-and-policies#twitter-rules> (last accessed 13 March 2018).

The Hateful Conduct and Abuse policy explicitly states that users ‘*may not promote violence against or directly attack or threaten other people on the basis of race, ethnicity, national origin, sexual orientation, gender, gender identity, religious affiliation, age, disability, or disease.*’¹³⁸

Hateful conduct policy

Freedom of expression means little if voices are silenced because people are afraid to speak up. We do not tolerate behavior that harasses, intimidates, or uses fear to silence another person’s voice. If you see something on Twitter that violates these rules, please report it to us.

How our policy works

As explained in the Twitter Rules,

- **Hateful conduct:** You may not promote violence against or directly attack or threaten other people on the basis of race, ethnicity, national origin, sexual orientation, gender, gender identity, religious affiliation, age, disability, or disease. We also do not allow accounts whose primary purpose is inciting harm towards others on the basis of these categories.

Examples of what we do not tolerate includes, but is not limited to behavior that harasses individuals or groups of people with:

- violent threats;
- wishes for the physical harm, death, or disease of individuals or groups;
- references to mass murder, violent events, or specific means of violence in which/with which such groups have been the primary targets or victims;
- behavior that incites fear about a protected group;
- repeated and/or non-consensual slurs, epithets, racist and sexist tropes, or other content that degrades someone.

Twitter’s Hateful Conduct Policy.

The policy¹³⁹ also provides an overview of the types of behaviours that are not allowed on the platform and encourages users to report content on the platform that they believe is in breach of Twitter’s community standards. However, Twitter does not state who is responsible for the oversight and implementation of the policy.

When a reported tweet or account is found to be in violation of the Twitter Rules, the company states it will either:¹⁴⁰ require the user to delete prohibited content before posting new content or interacting with other Twitter users; temporarily limit the user’s ability to create posts or interact with other Twitter users; ask the user to verify account ownership with a phone number or email address; or permanently suspend account(s). Twitter has a ‘philosophy’ on how it enforces its own rules but it does not provide concrete examples or guidance on this. It therefore remains unclear how the category or severity of violence or abuse reported to Twitter is assessed by company moderators to determine which of the aforementioned resolutions are applicable. Twitter does have a specific policy on abusive profiles¹⁴¹ that states accounts with ‘*abusive profile information usually indicate abusive intent and strongly correlate with abusive behavior*’ and describes how the platform reviews and enforces action against account profiles that violate the Twitter rules.

Twitter is taking some pro-active steps to address violence and abuse on the platform. In a letter to Amnesty International (see Annex I), Twitter detailed how it is using machine learning to identify and collapse potentially abusive and low-quality replies so that the most relevant conversations are brought forward. It also stated that such low-quality replies will still be accessible to those who seek them out. Twitter also stated (see Annex I)

¹³⁸ Twitter, *Help Center, Twitter Rules and policies*, available at <https://help.twitter.com/en/rules-and-policies/hateful-conduct-policy> (last accessed 13 March 2018).

¹³⁹ Twitter, *Help Center, Twitter Rules and policies*, available at <https://help.twitter.com/en/rules-and-policies/hateful-conduct-policy> (last accessed 13 March 2018).

¹⁴⁰ Twitter, *Help Center, Rules and policies*, available at <https://help.twitter.com/en/rules-and-policies#twitter-rules> (last accessed 13 March 2018).

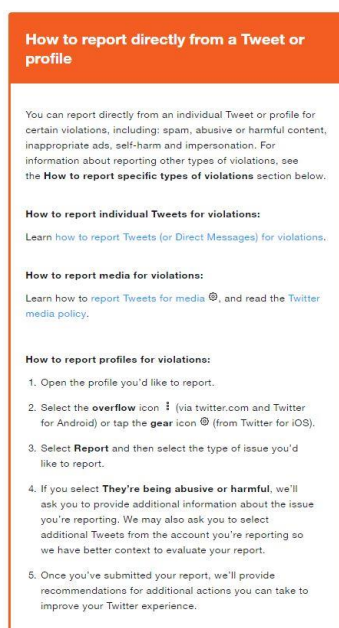
¹⁴¹ Twitter, *Help Center, General guidelines and policies*, available at <https://help.twitter.com/en/rules-and-policies/abusive-profile> (last accessed 13 March 2018).

that they are using machine learning to “bolster their approach to violence and abuse in a number of areas, from better prioritising report signals to identifying efforts to circumvent suspensions” – however, no further details or assessments were given about these specific efforts.

THE REPORTING PROCESS

Twitter’s policies, for the most part, contain definitions that could be used to address the range of violence and abuse that women experience on the platform. However, in addition to incorporating specific human rights policy commitments, Twitter also has a responsibility¹⁴² to ensure that women who experience violence and abuse on the platform have access to an effective complaints mechanism that they know about, are easily able to use and also trust.

Twitter’s reporting system relies on users to report violence and abuse on the platform. Individuals can report abusive or harmful content directly from a tweet or profile.



Twitter’s instructions on how to report tweets to the platform.

HOW TWITTER INTERPRETS AND IMPLEMENTS ITS RULES

Twitter states that it takes a number of factors into account when determining to take enforcement action, such as whether “the behaviour is directed at an individual, group or protected category of people; the report has been filed by the target of the abuse or a bystander; the user has a history of violating our policies; the severity of the violation; the content may be a topic of legitimate public interest.” However, the company does not clarify how these different factors determine which level of enforcement option moderators will apply to specific reports of violence and abuse.

¹⁴² UN Guiding Principles on Business and Human Rights, 2011, available at http://www.ohchr.org/Documents/Publications/GuidingPrinciplesBusinessHR_EN.pdf (last accessed 13 March 2018), p. 35.

When Twitter determines that a tweet is in violation of the rules, it requires the user to delete the specific tweet before they can tweet again. The user will then need to go through the process of deleting the violating tweet or appealing for review if they believe Twitter made an error. In October 2017, Twitter committed¹⁴³ to notify users whose accounts were locked for breaking the Twitter rules against abusive behaviour with details about the offending tweet and an explanation of which policy it violated, via email and in-app. On 9 March 2018, Twitter stated¹⁴⁴ that they will now also email account owners whose Twitter accounts have been suspended both the content of the tweet and which rule was broken to share more context around what led to the suspension. Whilst these improvements are a welcome step in Twitter communicating the types of content it finds to be in breach of its own policies on the platform, the company continues to fail in exercising wider transparency and communicating to its hundreds of millions of Twitter users about the behaviour it finds unacceptable on its platform.

Twitter does not share with users specific examples of content on the platform that would violate the Twitter rules nor does it provide specific information on how content moderators are trained to interpret the Twitter rules when dealing with reports of abuse. In response to a request from Amnesty International about how content moderators are trained to understand and enforce rules against abusive behaviour on the platform, Twitter stated (see Annex I), “Our review teams are empowered to use their judgement and take appropriate action on accounts that violate our rules.” This lack of clarity creates an uncertainty about what Twitter will act on and therefore leads women to place less faith in the reporting system.

Amnesty International also asked Twitter to share details about the content moderation process such as figures of the number of moderators employed per region and details about how content moderators are trained on gender and other identity based discrimination, as well as international human rights standards more broadly. Twitter declined to share this information and stated (see Annex I), “*We are well aware that ensuring our teams are taking into account gender-based nuance in behavior is essential. Every agent handling abuse reports receives in-depth trainings around the nuanced ways historically marginalized groups are harassed, intersectionality, and the impact that abuse online has on people's everyday lives, across a range of issues.*” However, Twitter did not share details about what this training includes, nor did they share the number of moderators employed per region and language.

HOW TWITTER RESPONDS TO REPORTS OF ABUSE

Twitter does not publicly share any specific data on how it responds to reports of abuse or how it enforces its own policies. Amnesty International requested that Twitter share disaggregated data about the company's reporting process and response rate on three separate occasions but our requests were refused. Twitter stated in their response (Annex I) that the absolute numbers of reports and the proportion of accounts that are actioned can be both uninformative and potentially misleading. In addition, they stated that users regularly report content with which they disagree or, in some cases, with the direct intent of trying to silence another users' voice for political reasons. Although context is incredibly important to understand in the reporting process, it does not negate Twitter's human rights responsibilities to be transparent in how it is dealing with reports of violence and abuse on the platform. The lack of meaningful and disaggregated data leaves little information by which to assess how the reporting process is working for users who report violence and abuse to the platform.

Twitter also states that it can take up to 24 hours for moderators to confirm receipt of a report but it does not explicitly guarantee that it will respond to reports of abuse nor does it stipulate how long it will take to respond to reports. It also does not set any public benchmarks or targets to record or improve response times to reports of abuse. There are currently no features on the platform for users to check on the status of any reports made to Twitter. In January 2018, Twitter introduced its latest reporting feature which provides an in-app notification to users about the progress of any submitted reports and specifies which particular rule an account or tweet is found to have violated. Although such a feature, in principle, provides users with more access to decisions

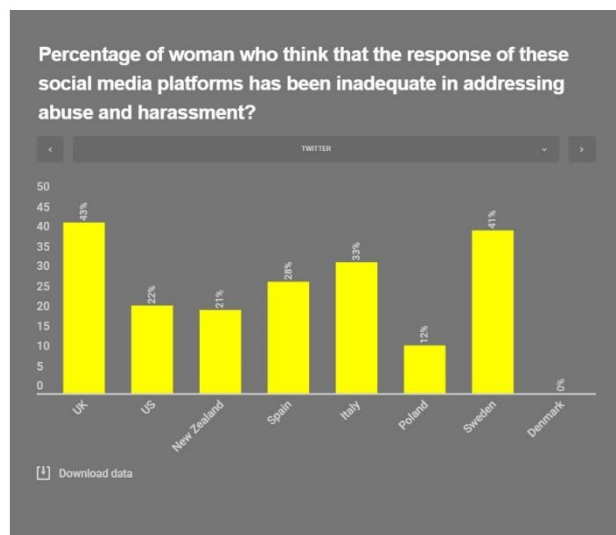
¹⁴³ Twitter Safety, Tweet, 19 October 2017, available at <https://twitter.com/TwitterSafety/status/921133556054687745> (last accessed 13 March 2018).

¹⁴⁴ Twitter Safety, Tweet, 08 March 2018, available at <https://twitter.com/TwitterSafety/status/971882517698510848> (last accessed 13 March 2018).

on reports that have been made and greater transparency about which Twitter rule has been broken, this update is only helpful when Twitter makes a decision on reports of violence and abuse.

Twitter also states¹⁴⁵ that users can appeal a decision based on their review of a report of violence and abuse if the user believes that they made an error. A detailed overview of the appeals process, including an explicit commitment to respond to all appeals or a timeframe of when to expect a response is not included in any of Twitter's policies. Additionally, although Twitter's policies state that users can report violence or abuse on the platform that they see happening to users, it does not seem to be investing in awareness-raising campaigns to ensure that users know that they are able to do so.

Amnesty International's online poll¹⁴⁶ showed that many women feel that Twitter's response to online abuse has been inadequate. 43% of women polled in the UK and 22% of women polled in the USA, who are Twitter users, stated that the company's response to abuse or harassment was inadequate. Almost 30% of women across all countries polled, excluding Denmark, who are Twitter users stated the company's response to abuse or harassment was inadequate. These figures show that Twitter has a lot more work to do when it comes to making its reporting mechanisms more effective for women who experience violence and abuse on the platform.



Results of Amnesty International's online poll across 8 countries on 'Percentage of women who think that the response of Twitter has been inadequate in addressing abuse or harassment'.

WOMEN'S EXPERIENCES: INCONSISTENT ENFORCEMENT OF TWITTER RULES

The Twitter rules are meant to provide guidance and clarity on which behaviours the company deems acceptable on the platform. Although Twitter's hateful conduct policy covers many forms of abuse that affect women rights, it means little for women on Twitter if such policies are not consistently enforced in practice.

¹⁴⁵ Twitter, *Help Center, General guidelines and policies*, available at <https://help.twitter.com/en/rules-and-policies/enforcement-options> (last accessed 13 March 2018).

¹⁴⁶ Azmina Dhrobia, *Unsocial Media: The Real Toll of Online Abuse against Women*, Amnesty International, Medium, 20 November 2017, available at <https://medium.com/amnesty-insights/unsocial-media-the-real-toll-of-online-abuse-against-women-37134ddb3f4> (last accessed 13 March 2018).

UK Politician Lisa Nandy told Amnesty International that she has little faith in Twitter's reporting process after a tweet she believed was threatening was found not to be in breach of the Twitter rules. She recounted,¹⁴⁷

"...The next tweet I got was from a guy who said that every time he saw my face it made his knuckles itch, so I thought, 'Well that's not right, is it?'. So I pressed Report and straight away I got this automated response saying 'We will investigate this and get back to you as soon as possible' and I thought that was brilliant. Then, about two hours later, I got a message saying 'We've looked at this and we don't consider it violated our guidelines around abuse so no further action will be taken'.... I thought – well this is just a total waste of time then, isn't it?"

In January 2018, UK journalist Ash Sarkar tweeted a screenshot of abuse¹⁴⁸ that she had received on Twitter which was found not to be in violation of the community standards. The racist and sexist language explicitly used in this tweet demonstrates how, even when reported, abusive content can remain on the platform due to Twitter's inconsistent interpretation and enforcement of its own rules.



Example of abusive tweet sent to UK journalist @ayoCaesar reported to Twitter and found to be not in violation of Twitter rules on 18 January 2018

Other women interviewed by Amnesty International shared similar concerns about Twitter's failure to protect its users online when they report abuse. UK journalist Allison Morris told Amnesty International that she feels Twitter has let her down after experiencing continuous and targeted harassment on Twitter.¹⁴⁹ Twitter's policy states¹⁵⁰ that "You may not engage in the targeted harassment of someone, or incite other people to do so. We consider abusive behavior an attempt to harass, intimidate, or silence someone else's voice."

As a verified Twitter user, Allison has access to additional features when filing a report on Twitter and is allowed to include a short explanation as to why she believes a specific tweet is in breach of Twitter's community standards. She told us,¹⁵¹

"Twitter, I find, don't remove anything. I think I've maybe managed, out of reporting probably over 100 posts to Twitter, I think they've removed two – one was a threat and the other had a pornographic image. There are 50 or

¹⁴⁷ Amnesty International interview with Lisa Nandy, 25 October 2017.

¹⁴⁸ Ash Sarkar, Tweet, 18 January 2018, available at <https://twitter.com/AyoCaesar/status/953959860403687424/photo/1> (last accessed 13 March 2018).

¹⁴⁹ Amnesty International interview with Allison Morris, 10 January 2018.

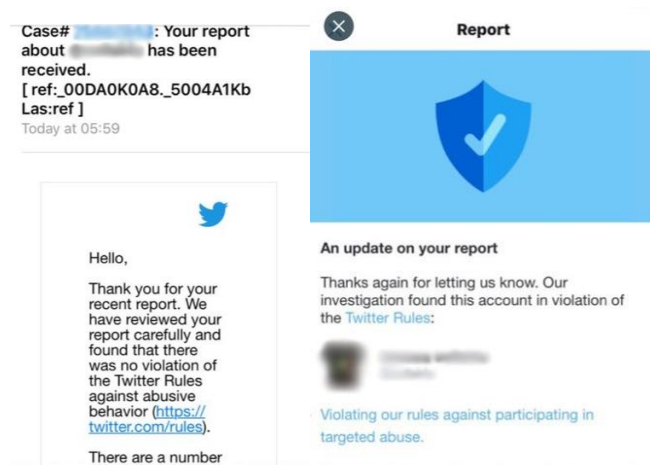
¹⁵⁰ Twitter, *Help Center, Twitter Rules and Policies*, available at <https://help.twitter.com/en/rules-and-policies/twitter-rules> (last accessed 13 March 2018).

¹⁵¹ Amnesty International interview with Allison Morris, 10 January 2018.

60 posts where I have specifically explained the tweets were malicious and the person that I believe is sending them has already been criminally convicted [for harassment but they won't remove any of the tweets.

Twitter says they will remove any accounts which are specifically set up to target a person. But I've had accounts set up where a person has only posted three times, all three posts were directed at me and were malicious and had comments about my children and my father who is dying. When I reported to Twitter, they said it didn't breach community standards."

In January 2018, Allison reported yet another account which she believed was in violation of Twitter's rules on 'targeted harassment'. Twitter initially found that the account was not in breach of their community standards on 'targeted harassment'. However, after Amnesty International contacted Twitter later that month about the specific account to clarify how the rules on 'targeted harassment' are interpreted, the account was found to be in breach of Twitter's rules upon further investigation by the company.



Report of abusive tweet by found not to be in violation of Twitter's rules. (left)

Report of same abusive tweet found to be in violation of Twitter's rules following contact by Amnesty International (right)

UK Politician Sophie Walker also spoke about her disappointment in Twitter failing to take violence and abuse against her seriously. She told Amnesty International,¹⁵²

I have had tweets saying 'We are going to tie you up, we are going to make you drink bleach, you will be sorry when you are in a burkha'. I have also had someone incite others to rape me, with the words: 'Who wouldn't rape Sophie Walker?' When I reported these to Twitter, the response came back so quickly it was almost like an 'out of office' reply. It was so fast that it felt automated, rather than considered. They said: 'We've investigated and there is nothing to see here.' I've tried escalating the reports, but the reply comes back just as fast and just the same. In some respects, Twitter's response to abuse is more hurtful than the abuse. It feels like they are saying: 'You're on your own if you participate in this forum.'"

¹⁵² Amnesty International interview with Sophie Walker, March 2018.



Tweet sent by Sophie Walker to Jack Dorsey, current Twitter CEO.

UK writer and activist Laura Bates also spoke about the discrepancies between Twitter’s policies and practices and how reporting abuse directly to more senior Twitter staff was the only way she received results. She explains,¹⁵³

“Social media companies aren’t stepping up and being held accountable for protecting the safety of their users. They talk a good game when something comes up in the press, but they’re not taking that action. In my experience of reporting accounts to Twitter, there’s a safety gap in terms of how their terms and conditions — which are quite vaguely worded — are interpreted. When I reported things to Twitter, it very rarely resulted in anybody being suspended. But when I was put in touch with someone who was higher up in the company, they took action and removed the harassment.

I think that’s significant because it shows that there’s a real gap between the small number of people who are engaged with these tech companies and everyday users. And it’s really problematic that social media companies are only taking action when they’re under pressure, because that amplifies already privileged voices and continues to push marginalized voices off these platforms.”

Sana Saleem, a Pakistani journalist who has been living in the USA for the last two years, also spoke about the lack of consistent enforcement she has witnessed after reporting abuse on the platform. She explained,¹⁵⁴

“It’s really weird. When I was working in Pakistan – and reporting accounts as abusive – they were being taken down and suspended. Here in the US, hardly ever has an account been suspended. I receive a lot of hate from Neo Nazis and white supremacists and Twitter hasn’t taken it down.”

UK journalist Hadley Freeman also spoke about the inconsistency of Twitter’s reporting system. She told us,¹⁵⁵

“Sometimes I report abuse to Twitter but most of the time it does nothing. Twitter is really unpredictable about what it considers abuse. I could report a virulently abusive message and they say it doesn’t violate their guidelines. And then I’ll report something relatively minor and they’ll take action.”

UK Politician Naomi Long also spoke about Twitter’s failure to take reports of abuse against her seriously. She told us,¹⁵⁶

“In respect of reporting abuse on Twitter I think it’s a wholly ineffective process – I have to be honest. I have reported abuse and unless it is very specific, they don’t take it seriously.”

The inconsistent enforcement of its own policies reinforces the need for Twitter to be far more transparent about how it interprets and enforces its own community standards. It is also important to note that an increased level of transparency from Twitter about the reporting process benefits all users on the platform. One way to

¹⁵³ Amnesty International interview with Laura Bates, 25 January 2017.

¹⁵⁴ Amnesty International interview with Sana Saleem, 20 February 2018.

¹⁵⁵ Amnesty International interview with Hadley Freeman, 29 January 2018.

¹⁵⁶ Amnesty International interview with Naomi Long, 11 January 2018.

ensure that Twitter is a less toxic place for women is to clearly communicate and reinforce to users which behaviours are not tolerated on the platform and to consistently apply its own rules.

WOMEN'S EXPERIENCES: TWITTER'S INACTION ON REPORTS OF ABUSE

Many women who spoke to Amnesty International told us that when they report abuse on Twitter it is often met with inaction and silence from the platform. These experiences reflect a similar trend found in a 2017 study¹⁵⁷ by the Fawcett Society¹⁵⁸ and Reclaim the Internet¹⁵⁹. The study reviewed a range of abusive content on Twitter which was against Twitter's policies and then reported to the platform by anonymous accounts. One week after the tweets were reported, the posts remained on Twitter, the accounts which reported the abuse received no further communication from the platform and no action had been taken on any of the accounts that were reported.

Multiple women interviewed by Amnesty International echoed these findings. US journalist and writer Jessica Valenti said,¹⁶⁰

"I have not gotten a lot of movement on anything unless it is a very direct, obvious threat. That's part of the problem. When someone says 'Someone should shoot Jessica in the head', that's very obvious. But harassment can be savvy and they know what they can say that's not going to get them kicked off a site or not illegal so I don't even bother with stuff like that... But I've reported tons of stuff and nothing happens".

Scottish Parliamentarian and Leader of the Opposition Ruth Davidson shared a similar experience. She explained,¹⁶¹

"Have I bothered to report? I think in the beginning I did and not much happened, so I don't know, maybe it isn't the best example to set. I think that if we do want platforms to act in a more robust manner then we've got to keep up the level of complaint".

US reproductive rights activist¹⁶² and blogger Erin Matson also told Amnesty International how she never heard back from Twitter after reporting abuse to the platform. She told us,¹⁶³

"Most of the harassment I received is on Twitter. It's very fast paced and people send you horrible images. One time I tweeted that white people need to be accountable towards racism and I started getting images of death camps... I've reported people to Twitter but I never heard back".

In 2015, US organization Women, Action, Media! (WAM!) was temporarily elevated to 'Trusted Reporter Status'¹⁶⁴ by Twitter. Being a 'Trusted Reporter' allowed WAM! to assess and escalate reports of abuse (as necessary) to Twitter for special attention. US writer and activist Jaclyn Friedman, who was also the Executive Director of the organization during this time, explained that many of the women who reported abuse through WAM! said they had previously made reports of abuse to Twitter but never heard back from the platform. Jaclyn explained,¹⁶⁵

¹⁵⁷ Fawcett Society and Reclaim the Internet, *TWITTER IS 'FAILING WOMEN' EXPERIENCING ONLINE THREATS AND HARASSMENT*, 22 August 2017, available at <https://www.fawcettsociety.org.uk/News/twitter-failing-women-experiencing-online-threats-harassment> (last accessed 13 March 2018).

¹⁵⁸ Fawcett Society, Website, available at <https://www.fawcettsociety.org.uk> (last accessed 15 March 2018).

¹⁵⁹ Reclaim The Internet, Website, available at <http://www.reclaimtheinternet.com> (last accessed 15 March 2018).

¹⁶⁰ Amnesty International interview with Jessica Valenti, 15 May 2017.

¹⁶¹ Amnesty International interview with Ruth Davidson, 21 June 2017.

¹⁶² Reproaction, Website, available at <https://reproaction.org> (last accessed 14 March 2018).

¹⁶³ Amnesty International interview with Erin Matson, 26 June 2017.

¹⁶⁴ Women , Action, Media!, *REPORTING, REVIEWING, AND RESPONDING TO HARASSMENT ON TWITTER*, 13 May 2015, available at <http://womenactionmedia.org/cms/assets/uploads/2015/05/wam-twitter-abuse-report.pdf> (last accessed 13 March 2018).

¹⁶⁵ Amnesty International interview with Jaclyn Friedman, 14 July 2017.

“...The things that stays with me...is just the number of people we heard from who said ‘I reported this over and over and over and got no response’ - and that we were able to get a response. That gap, I think, is really instructive.”

Jaclyn also shared multiple examples of abusive tweets that she has reported to Twitter, but which remain on the platform.



Example of abusive tweet mentioning US writer and activist @jaclynf

Rachel*, a young woman in the UK without a large Twitter following, told us about her experience of reporting abuse.¹⁶⁶

“I reported the abusive tweets and accounts. From that, one account got suspended. Everything else I reported to Twitter, they said, we are taking your complaint seriously and then nothing else was done. There was no further contact from Twitter. There was nothing.

It almost feels that when you are filling in the reporting questions that they don’t believe you. It’s like a recurring theme with women – people don’t instantly believe you.”

UK Politician and activist, Seyi Akiwowo explained to Amnesty International how she also felt let down by Twitter when they initially failed to take any action on the multiple reports she made after receiving a torrent of racist and sexist abuse in February 2017. She recounted,¹⁶⁷

“I reported around 75 comments on both YouTube and Twitter and did not receive feedback or acknowledgment. It took around two hours to go through the comments and it was really upsetting.

Twitter never actually contacted me when I reported it – they only started suspending those accounts and deleting those tweets after I appeared on the ITV’s London Tonight and did media interviews with BuzzFeed and the BBC. So if I didn’t go into fighter mode and make media appearances and gain public support I very much doubt there would be any action from Twitter.

I was really frustrated with] their lack of response and that is what led me to start Glitch! UK.¹⁶⁸ At the time, I felt like Twitter were leaving these tweets there for me to be abused and there was no one to help me. I felt kind of betrayed. I feel silly about saying that - it’s obviously just a company, they don’t know me owe me anything...but, actually, they do. For it to turn like that and to be abandoned by Twitter at my time of need, my hour of need, it was just really sad. I was more frustrated at feeling let down by Twitter than these idiots that don’t know me calling me the N-word.”

¹⁶⁶ Amnesty International interview with Rachel* (Pseudonym), 27 April 2017.

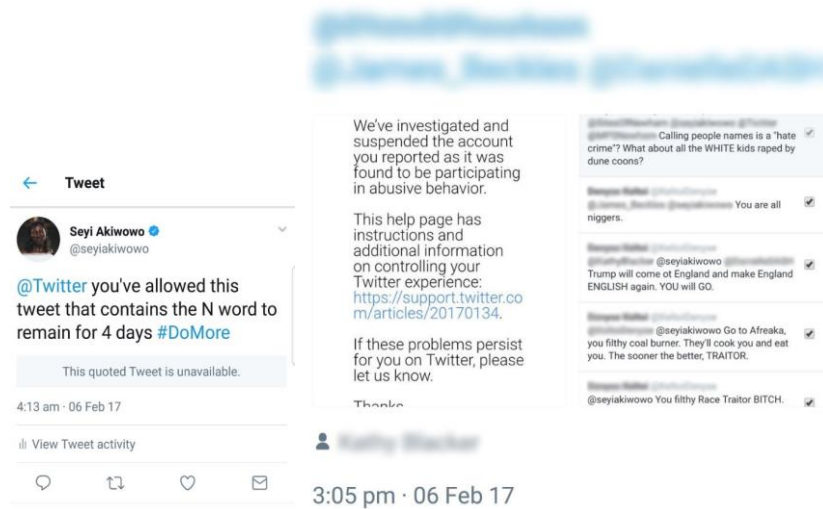
¹⁶⁷ Amnesty International interview with Seyi Akiwowo, 30 March 2017.

¹⁶⁸ Glitch! UK, website, available at <https://seyiakiwowo.com/glitchuk/> (last accessed 15 March 2018).



Seyi Akiwowo
@seyiakiwowo

Four days later @Twitter have finally suspended this awful troll's account. #smallvictories



Tweet by UK Politician and activist Seyi Akiwowo to @Twitter about an abusive tweet that remained on platform 4 days after reporting. (left)

Tweet by UK Politician and activist Seyi Akiwowo about suspension of account that sent her abusive tweet. (right)



Laura Bates, UK Writer And Activist

Online abuse began for me when I started the Everyday Sexism Project – before it had become particularly high-profile or received many entries. Even at that stage, it was attracting around 200 abusive messages a day. The messages included detailed, graphic, and explicit descriptions of rape and domestic violence. When it was through Twitter, it was clearly in response to news stories where I was one of the people speaking out about online abuse. It was like there was a backlash – people making threats about you because you are speaking up about rape threats.

I found it difficult not to be scared about my safety initially. The psychological impact of reading through someone's really graphic thoughts about raping and murdering you is not necessarily acknowledged. You could be sitting at home in your living room, outside of working hours, and suddenly someone is able to send you an incredibly graphic rape threat right into the palm of your hand. There tends to be a general reaction of, 'if you know they're trying to scare you, why would you be worried?' But when you're receiving the abuse on such a large scale, it's hard not think that it only takes one person to follow through.

Social media companies aren't stepping up and being held accountable for protecting the safety of their users. In my experience of reporting accounts to Twitter, there's a safety gap in terms of how their terms and conditions — which are quite vaguely worded — are interpreted. When I reported things to Twitter, it very rarely resulted in anybody being suspended. But when I was put in touch with someone who was higher up in the company, they took action and removed the harassment. I think that's significant because it shows that there's a real gap between the small number of people who are engaged with these tech companies and everyday users.

We are seeing young women and teenage girls experiencing online harassment as a normal part of their existence online. It's an invisible issue right now, but it might be having a major impact on the future political

participation of those girls and young women. We won't necessarily see the outcome of that before it's too late.

WOMEN STOP REPORTING ABUSE

One of the most pernicious impacts of the combination of inaction and inconsistency in responding to reports of abuse is the detrimental effect this has on women reporting experiences of abuse in the future. Women who are the targets of abuse bear the burden of reporting it. This not only takes time, but also takes an emotional toll on women. When women have had – or hear of people who have had – negative experiences reporting abuse to Twitter, they are often less likely to feel it is worth the effort to undertake the work of reporting abuse.

US abortion rights activist Renee Bracey Sherman explained,¹⁶⁹

“I am way over Twitter. If I am dealing with a lot of hate, I can't bear to be in there, so I don't even report it.”

US activist Pamela Merritt stressed how exhausting the reporting process can be after experiencing abuse on the platform. She said,¹⁷⁰

“When possible, I report it. By possible, I mean if I have the time and the emotional bandwidth to be disappointed about the report. In my experience, they rarely take action. I have only ever had them take one report seriously.”

US activist Shireen Mitchell told us,¹⁷¹

“I don't bother reporting anymore for myself because it doesn't matter. All the solutions they came up with are ridiculous. The problem is they do not have enough diverse staff – enough to understand what the threats look like for different groups. They think these are jokes and they have allowed this to be part of the discourse.”

When Twitter's reporting processes fail women who report abuse, others may be dis-incentivized from attempting to use the processes. UK journalist Siobhan Fenton told us,¹⁷²

“I wrote an article about transgender rights a couple of years ago and someone on Twitter sent me an image of a transgender woman who had been raped and killed and said that the same thing should happen to me. I think I mentioned it to a few friends at the time, just for support, but it didn't occur to me to report it to Twitter itself because I know that there are journalists who have experienced online abuse and whenever they've reported things to Twitter they almost invariably don't get a positive response and the materials tends to stay up.”

Twitter has a responsibility to ensure that any potential human rights abuses - such as the silencing or censoring of women's expression on their platform - are addressed through prevention or mitigation strategies,¹⁷³ and the reporting process is a key factor in this. However, in order for the reporting process to work effectively, Twitter must be much more transparent in how it interprets its own rules. It is also important to note that most Twitter users do not have verified accounts,¹⁷⁴ and as a result are unable to provide any additional context or a written explanation when they report a tweet or account as abusive.

¹⁶⁹ Amnesty International interview with Renee Bracey Sherman, July 2017.

¹⁷⁰ Amnesty International interview with Pamela Merritt, 07 July 2017.

¹⁷¹ Amnesty International interview with Shireen Mitchell, July 2017.

¹⁷² Amnesty International interview with Siobhan Fenton, 10 January 2018.

¹⁷³ *UN Guiding Principles on Business and Human Rights*, 2011, available at http://www.ohchr.org/Documents/Publications/GuidingPrinciplesBusinessHR_EN.pdf (last accessed 13 March 2018), p. 13.

¹⁷⁴ Twitter, *Help Center, Verified Accounts*, available at <https://help.twitter.com/en/managing-your-account/about-twitter-verified-accounts> (last accessed 13 March 2018).

Given the complexities of assessing abuse in different regions and the importance of context when determining whether certain content is abusive, providing a rigid definition of what constitutes online abuse is not a simple task. However, although there are ‘grey areas’ around the parameters of online abuse– this should not be an excuse for Twitter’s inaction and inconsistency in dealing with reports of abuse. It would be much easier to tackle the ‘grey areas’ of abuse if Twitter were, first and foremost, clear about the specific forms of violence and abuse that it will not tolerate. A paper¹⁷⁵ by the Association for Progressive Communications on Due Diligence and Accountability for Online Violence against Women recommends that companies should create appropriate record keeping systems specific to violence against women and classify and share the ways in which they have responded to it. Ultimately, Twitter’s lack of transparency about how it interprets and enforces its own rules undermines the importance of human rights that are implied in its policy commitments.

In the words of US Black Lives Matter Activist, Miski Noor,¹⁷⁶

“Twitter needs to hone in on their responsibilities and their values. I’m tired of tech companies or social media companies thinking they are exempt from living their values. If Twitter values women and femmes, if they value our safety, then they need to have practices that they actually develop and implement in real ways that will protect us.”

TWITTER’S HUMAN RIGHTS FAILURES

In its response to Amnesty International (See Annex II), Twitter highlighted several positive changes to their policies and practices in response to violence and abuse on the platform over the past 16 months, including a tenfold increase in the number of abusive accounts against action has been taken.

Amnesty International’s findings indicate that Twitter’s inconsistent enforcement and application of the rules as well as delays or inaction to reports of abuse when users breach the Twitter rules mean that the company’s response is still insufficient. Together with the lack of specific human rights policy commitments and seemingly ineffective reporting mechanisms, this clearly demonstrates a failure of the company to adequately meet its corporate responsibility to respect human rights in this area. Based on Amnesty International’s research and publically available materials, and given Twitter’s refusal on three occasions to disclose comprehensive data on the reporting process, Amnesty International’s conclusion is that these failures result in Twitter contributing to the harms associated with women’s experiences of violence and abuse on the platform. It is worrying that a social media platform of Twitter’s importance does not appear to have adequate human rights compliant policies and processes in place to tackle this problem adequately, efficiently and transparently. If Twitter had disclosed the requested information about reports of abuse, this may have revealed additional insight that would have impacted our conclusion.

¹⁷⁵ Zarizana Abdul Aziz, *Due diligence and accountability for online violence against women*, Association for Progressive Communications, July 2017, available at <https://www.apc.org/en/pubs/du-diligence-and-accountability-online-violence-against-women> (last accessed 13 March 2018).

¹⁷⁶ Amnesty International interview with Miski Noor, 07 February 2018.

SECTION 5: THE SILENCING EFFECT

Women have the right to use Twitter equally, freely and without fear. But Twitter's inadequate response to violence and abuse against women is leading women to self-censor what they post, limit or change their interactions online, or is driving women off the platform altogether. At times, the threat of violence and abuse against women on Twitter, alone, leads to a chilling effect on women speaking out online. The silencing and censoring impact of violence and abuse against women on Twitter can have far-reaching and harmful repercussions on how younger women, women from marginalized communities, and future generations fully exercise their right to participate in public life and freely express themselves online.

"In my experience there are many women, and many women of colour, who don't participate online in the way that they would want to because of online abuse."

Diane Abbott, UK Politician and Shadow Home Secretary¹⁷⁷

WHY VIOLENCE AND ABUSE AGAINST WOMEN ON TWITTER IS A FREEDOM OF EXPRESSION ISSUE

Ensuring that everyone can freely participate online and without fear of violence and abuse is vital to ensuring that women can effectively exercise their right to freedom of expression.

The rights to freedom of expression and to non-discrimination are guaranteed under major international human rights instruments including the Universal Declaration of Human Rights and the International Covenant on Civil and Political Rights.

The United Nations Human Rights Council has stated that "the same rights people have offline must also be protected online, in particular freedom of expression, which is applicable regardless of frontiers and through any media of one's choice, in accordance with articles 19 of the Universal Declaration of Human Rights and the International Covenant on Civil and Political Rights".¹⁷⁸

Additionally, in a joint-statement,¹⁷⁹ the United Nations experts on Freedom of Expression and Violence against Women, respectively, commented on the negative impact of online abuse on the right to freedom of expression online for women. Special Rapporteur on Freedom of Expression, David Kaye states,

"The internet should be a platform for everyone to exercise their rights to freedom of opinion and expression, but online gender-based abuse and violence assaults basic principles of equality under international law and freedom of expression. Such abuses must be addressed urgently, but with careful attention to human rights law".

¹⁷⁷ Amnesty International interview with Diane Abbott, 24 July 2017.

¹⁷⁸ UN Human Rights Council, *The promotion, protection and enjoyment of human rights on the Internet*, 27 June 2016, A/HRC/32/L.20, available at <https://documents-dds-ny.un.org/doc/UNDOC/LTD/G16/131/89/PDF/G1613189.pdf?OpenElement> (last accessed 13 March 2018) p. 3, para. 1.

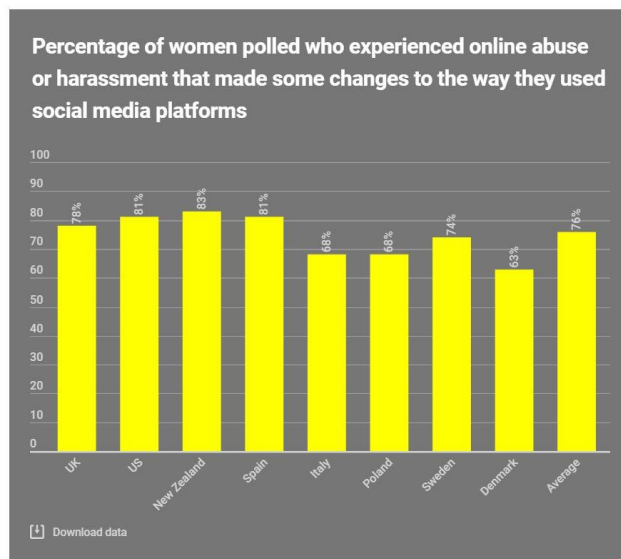
¹⁷⁹ UN Office of the High Commissioner, *UN experts urge States and companies to address online gender-based abuse but warn against censorship*, 08 March 2017, available at <http://www.ohchr.org/EN/NewsEvents/Pages/DisplayNews.aspx?NewsID=21317&LangID=E> (last accessed 13 March 2018).

UN Special Rapporteur on Violence against Women Dubravka Simonoci adds,

“Ensuring an internet free from gender-based violence enhances freedom of expression as it allows women to fully participate in all areas of life and is integral to women’s empowerment”.

Many of the women interviewed by Amnesty International described changing their behaviour on the platform due to Twitter’s failure to provide adequate remedy when they experienced violence and abuse. The changes women make to their behaviour on Twitter ranges from self-censoring content they post to avoid violence and abuse, fundamentally changing the way they use the platform, limiting their interactions on Twitter, and sometimes, leaving the platform completely.

Amnesty International’s online poll¹⁸⁰ about women’s experiences of abuse on social media platforms confirmed that the experiences of women interviewed in this investigation are a reflection of the silencing and censoring impact of online abuse on women more generally. It found that of the women polled who experienced online abuse or harassment, between 63% and 83% women in the eight countries polled¹⁸¹ made some changes to the way they used social media platforms, with the figures for the USA and UK being 81% and 78% respectively. The specific ways in which women modified their online interactions after experiencing abuse or harassment varied from women increasing privacy and security settings to women making changes to the content they post. However, the number of women who made changes to the posting or sharing of expression or opinions is deeply troubling. Amnesty International’s online poll¹⁸² showed that across the 8 countries polled, 32% of women who experienced abuse or harassment online said they had stopped posting content that expressed their opinion on certain issues, including 31% of women in the UK and 35% of women in the USA.



Results of Amnesty International’s online poll across 8 countries on the percentage of women polled who experienced online abuse or harassment and made some changes to the way they used social media platforms.

¹⁸⁰ Azmina Dhrodia, *Unsocial Media: The Real Toll of Online Abuse against Women*, Amnesty International, Medium, 20 November 2017, available at <https://medium.com/amnesty-insights/unsocial-media-the-real-toll-of-online-abuse-against-women-37134ddab3f4> (last accessed 13 March 2018).

¹⁸¹ Amnesty International, *Amnesty reveals alarming impact of online abuse against women*, 20 November 2017, available at <https://www.amnesty.org/en/latest/news/2017/11/amnesty-reveals-alarming-impact-of-online-abuse-against-women/> (last accessed 12 March 2018).

¹⁸² Azmina Dhrodia, *Unsocial Media: The Real Toll of Online Abuse against Women*, Amnesty International, Medium, 20 November 2017, available at <https://medium.com/amnesty-insights/unsocial-media-the-real-toll-of-online-abuse-against-women-37134ddab3f4> (last accessed 13 March 2018).

LIMITING INTERACTIONS AND CHANGING BEHAVIOUR

Getting started is easy

Find a bunch of things you love. And then find people to follow. That's all you need to do to see and talk about what's happening. Congratulations! You've just mastered Twitter. So let's get started.

Instructions on how to start using Twitter: https://about.twitter.com/en_us/lets-go-twitter.html

For many women, using Twitter is not easy; it means adapting their online behaviour and presence, self-censoring the content they post and limiting interactions on the platform out of fear of violence and abuse.

For example, Scottish Parliamentarian and Leader of the Opposition Ruth Davidson told Amnesty International about how an influx of abuse can change the way she uses the platform. She explains,¹⁸³

“...The sheer volume of abuse can make you sometimes feel hunted online. At that point you just stop reading the mentions and you use it as a transmit function rather than a transmit and receive platform.”

US journalist and writer Jessica Valenti told Amnesty International that although security features like ‘block’ and ‘mute’, as well as modifying Twitter notifications, have helped filter out abuse, it also means the platform and the way she uses it has fundamentally changed the way she interacts.¹⁸⁴

“Twitter has a thing where you can turn off notifications from anyone you don't follow back. It's a good thing and a bad thing. It's vastly improved my day to day experience, but it's bad in terms of I don't actually get to hear from my readers, so that sucks.”

Women, like US journalist Imani Gandy, also stressed to Amnesty International how violence and abuse against women on the platform simply becomes normalized after a while. Imani states,¹⁸⁵

“The abuse I receive on Twitter is mostly filtered out and I have trained myself to not search my name because then I get really angry. I've been online for 8 years and political on Twitter for 7 years so after 7 years I guess you just get used to it...it has become a part of my life, which is kind of sad, I guess. The filtering on my part and using third party apps have made it better but I'm also trying to reduce my time on Twitter as well.”

For many women, the inability to fully participate and express themselves equally online means that they are absent from public conversations they would like to be part of, and sometimes, *need* to be part of. To not engage or comment on an issue out of fear of violence and abuse means that certain women's voices are not represented on Twitter and that women are no longer part of the debate. For women in the public eye, in particular, this can have a detrimental effect on their career and building networks. The silencing effect of online abuse on women, including on Twitter, may also send a worrying message to younger generations that women's voices are not welcome.

¹⁸³ Amnesty International interview with Ruth Davidson, 21 June 2017.

¹⁸⁴ Amnesty International interview with Jessica Valenti, 15 May 2017.

¹⁸⁵ Amnesty International interview with Imani Gandy, 15 August 2017.

SELF-CENSORSHIP TO AVOID ABUSE

Many of the women who spoke to Amnesty International stressed that freely expressing themselves on Twitter is not worth the risk of violence and abuse.

UK activist Alex Runswick told Amnesty,¹⁸⁶

“The abuse has made me much, much more reluctant to comment on things. I mean, I still do tweet sometimes, but I think about it very, very careful before I do. Recently there was a case in the media about comments a judge had made in a rape case...and I was just thinking, okay, I want to tweet about this. And then I actually sat down for five minutes and had a conversation with my husband and said ‘Do I tweet this or not’ knowing full well what the response would be...”

US writer and presenter Sally Kohn echoed similar sentiments. She said,¹⁸⁷

“Once in a while there will be a tweet that I think I want to send and I’ll go, ‘Oh no, it’s not worth the trolls’.”

UK activist Sian Norris recalled a time that she posted a tweet on Transgender Day of Remembrance but took it down a few minutes later. She explained,¹⁸⁸

“I had posted something like ‘Today we remember transgender women being killed by male violence’. Within minutes around 5 accounts all with really horrible naked pictures of a man in the profile photo retweeted it. The re-tweets didn’t say anything but just had a really confrontational account photo and so I deleted my tweet. I was cross because it was an important day and I should be able to send solidarity.”

UK journalist Allison Morris explained to Amnesty International how she no longer posts about topics that she would have tweeted about in the past.¹⁸⁹

“First off, I don’t get Twitter notifications on my phone, I don’t want to see them. Also, I definitely self-censor what I post. There are things that I don’t tweet about even though I would have a couple of years ago. I think it’s just not worth the hassle, it’s not worth the abuse, and it’s not worth having to deal with hundreds of people all piling in and re-posting and re-tweeting it.”

Others talked about the time and energy they spend in carefully crafting and curating their Twitter posts to minimize the risk of violence and abuse on the platform.

UK poet and actor Travis Alabanza told Amnesty,¹⁹⁰

“I think before every tweet I send. Every single tweet now – it’s all completely crafted. My online persona is always crafted but now it’s to the point that there is nothing real about my Twitter anymore...”

I stay well clear of trans stuff most of the time. And if I’m going to bring up trans things now, I preface it so much in a thread, and as soon as I see it gain traction I look and decide whether or not I want it, and if I don’t, I delete. Before I would say I was far more political online and that’s what I was getting known for. Now I’ve moved a lot of this onto another social media platform. Everything about my Twitter now is just heavily constructed.”

¹⁸⁶ Amnesty International interview with Alex Runswick, A27 April 2017.

¹⁸⁷ Amnesty International interview with Sally Kohn, 15 November 2017.

¹⁸⁸ Amnesty International interview with Sian Norris, 13 April 2017.

¹⁸⁹ Amnesty International interview with Allison Morris, 10 January 2018.

¹⁹⁰ Amnesty International interview with Travis Alabanza, 24 January 2018.

US games developer Zoe Quinn faces a similar process every time she thinks about posting on Twitter. She explains,¹⁹¹

“I have to think three times about everything I post on Twitter. A lot of it relies on how ‘fuck it’ I’m feeling that day – which is exhausting. I also have to be very careful about who I visibly support online. If they haven’t been briefed on what me giving them a platform or visibility means [in terms of potential abuse], it is kind of unfair to them. People have set up bots to archive everything that I do and I’m not sure everyone is prepared to handle [the level of abuse I’m used to].”

WOMEN LEAVING TWITTER

The silencing impact of violence and abuse against women on Twitter manifests in different ways for different women online. Although the degree to which women are silencing themselves may differ, the impact of violence and abuse on the right to freedom of expression is a cause for concern not only for Twitter, but society more widely.

In July 2017, UK Member of Parliament and Shadow Home Secretary, Diane Abbott, told Amnesty International, *“Well the abuse that I get online does to a degree limit my freedom of expression. In terms of Twitter, I now spend much less time on Twitter than I used to because the abuse is so terrible.”* More recently in January 2018, she commented,¹⁹²

“I hardly go on Twitter any more...It’s a shame really, I used to enjoy Twitter.”

In recent years Twitter has been plagued with high-profile women leaving the platform after experiencing violence and abuse. For many, this is not an easy decision but seems to be the only option given the company’s failure to adequately tackle this issue and provide appropriate remedies. In August 2016, US actor Leslie Jones temporarily left Twitter¹⁹³ after a wave of racist abuse on the platform following the release of the film *Ghostbusters 2*. In January 2017, US journalist Lindy West wrote an op-ed about leaving the platform after years of enduring violence and abuse on Twitter.¹⁹⁴



Tweet posted by US actor Leslie Jones (@Lesdoggg) about leaving Twitter, 19 July 2016.

¹⁹¹ Amnesty International interview with Zoe Quinn, 13 December 2017.

¹⁹² Michael Segalov, *Diane Abbott: ‘The abuse and the attacks have never made me falter’*, the Guardian, 27 January 2018, available at <https://www.theguardian.com/lifeandstyle/2018/jan/27/diane-abbott-the-abuse-never-made-me-falter> (last accessed 15 March 2018).

¹⁹³ Jacob Bryant, *Leslie Jones Calls Out Racist ‘Ghostbusters’ Haters on Twitter*, Variety, 18 July 2016, available at <http://variety.com/2016/film/news/leslie-jones-ghostbusters-twitter-hate-1201816809/> (last accessed 13 March 2018).

¹⁹⁴ Lindy West, *I’ve left Twitter. It is unusable for anyone but trolls, robots and dictators*, the Guardian, 03 January 2017, available at <https://www.theguardian.com/commentisfree/2017/jan/03/ive-left-twitter-unusable-anyone-but-trolls-robots-dictators-lindy-west> (last accessed 13 March 2018).

US writer Chelsea Cain told Amnesty International about her decision to temporarily leave Twitter after experiencing abuse on the platform. She recounted,¹⁹⁵

“I had written a comic book called Mockingbird - a series - and the cover of the last issue featured the character wearing a shirt that says, 'Ask Me About My Feminist Agenda.' I knew it was provocative. That was the point. But I was surprised at the level of hysteria. Let's just say that not everyone appreciated my feminist gesture, and most of them had Twitter accounts.

I posted something on Twitter about my intention to delete my account. Then I left my computer and watched an episode of Buffy with my daughter, which seemed like a far better use of my time. I didn't go back online until the next morning. By then Twitter had exploded. So many comments were appearing in my feed that they were coming and going faster than I could read them. I deleted my account, closed my laptop, backed away, went out the front door, and kept walking. It was like leaving a burning house.”

US abortion rights activist Renee Bracey Sherman also explained how she left Twitter and other social media platforms following the wave of abuse and violence she received after writing an open letter in media outlet Refinery 29. She told us,¹⁹⁶

“The harassment on Twitter lasted two to two and a half weeks. I was dealing with so much hate and I had never experienced so much. Someone tweeted at me saying they hoped I would get raped over and over again...that's when I was like this isn't the normal shit I am used to...I threw my phone under the couch and hid in bed. I left social media for 6 weeks.”

US journalist and writer Jessica Valenti spoke about her experience of leaving Twitter after someone posted a rape threat against her daughter on another platform. She explained,¹⁹⁷

“Someone posted ‘I'm going to rape and gut your daughter’...It was a breaking point for me and I needed to step back. I needed to take stock of what I wanted to do and I needed some space to be able to do that.”

UK journalist Vonny Moyes also spoke about the need to leave Twitter when abuse or violence against her becomes too much to handle. She told us,¹⁹⁸

“I'm being really disciplined about self-care and taking time out to re-charge my batteries so I deactivate Twitter for a few days. Or maybe a week at a time. Even last weekend, I had a steady drip of harassment and dismissiveness on Twitter so I thought ‘Let's get out of the city, go to the beach and turn off our phones’...it's crazy that I have to do all these things in my offline life to balance out the effects of being online. It's really elbowed its way into my – I don't want to say ‘real’ life – because my digital life is my real life. It's everyone's real life now.”

IMPACT OF VIOLENCE AND ABUSE ON WOMEN'S PARTICIPATION IN PUBLIC LIFE

Ensuring the Internet is free from violence and abuse against women enhances freedom of expression and allows women to participate on an equal basis in public life. US activist Pamela Merritt talked to Amnesty International about the importance of women staying online. She told us,¹⁹⁹

“I think it is so important that women, and women of colour, particularly, stay online. It's important that we create our own spaces where we can be safe and share ideas. We should see this harassment for what it is: an extension of the patriarchy and oppression. The goal of online harassment is to erase women and women of colour from the

¹⁹⁵ Amnesty International interview with Chelsea Cain, 09 June 2017.

¹⁹⁶ Amnesty International interview with Renee Bracey Sherman, July 2017.

¹⁹⁷ Amnesty International interview with Jessica Valenti, 15 May 2017.

¹⁹⁸ Amnesty International interview with Vonny Moyes, 23 March 2017.

¹⁹⁹ Amnesty International interview with Pamela Merritt, 07 July 2017.

public dialogue. So the internet is a space where we are having an impact and that is where the harassers want to silence us.”

The UN Special Rapporteur on Violence against Women warns²⁰⁰ that online violence and abuse,

“can lead women and girls to limit their participation and sometimes withdraw completely from online platforms”.

UK writer and activist Laura Bates re-iterated these concerns. She explains,²⁰¹

“We are seeing young women and teenage girls experiencing online harassment as a normal part of their existence online. Girls who dare to express opinions about politics or current events often experience a very swift, misogynistic backlash. This might be rape threats or comments telling them to get back in the kitchen. It’s an invisible issue right now, but it might be having a major impact on the future political participation of those girls and young women. We won’t necessarily see the outcome of that before it’s too late.”

Laura Bates’ words were echoed by young women. Rachel*, a 19-year-old woman in the UK without a large Twitter following, told Amnesty,²⁰²

“The main thing that goes through my head every time I tweet anything feminist in nature is I’ll probably get death threats if this gets any traction. It’s sad and depressing to think that every time you tweet something opinionated it may come back with something horrible. Being online should be somewhere you feel safe, if you can’t feel safe on your Twitter account – then what’s the point? Sometimes I feel like just leaving Twitter”.

A report²⁰³ by the National Democratic Institute²⁰⁴ highlights just how serious the silencing and censoring impact of violence and abuse against women online can be. It states,

“By silencing and excluding the voices of women and other marginalized groups, online harassment fundamentally challenges both women’s political engagement and the integrity of the information space....In these circumstances, women judge that the costs and danger of participation outweigh the benefits, and withdraw from or choose not to enter the political arena at all”.

This was also a particular concern for female politicians interviewed by Amnesty International. Former UK Politician Tasmina Ahmed-Sheikh summed up her concerns by saying,²⁰⁵

“If online abusers are not held to account, if they are not reported, if we don’t do that, then young women are not going to come forward, young women from minority communities are not going to come forward, disabled people are not going to come forward, people from LGBTI communities are not going to come forward, and then what kind of society are we going to be? What will our Parliament look like?”

First Minister of Scotland Nicola Sturgeon, raised a similar worry about the impact of online abuse against women in politics deterring a future generation of leaders from entering politics. She explains,²⁰⁶

“What makes me angry when I read that kind of abuse about me is, I worry about that it’s putting the next generation of young women off politics. So, I feel a responsibility to challenge it, not so much on my own behalf,

²⁰⁰ UN Office of the High Commissioner, *UN experts urge States and companies to address online gender-based abuse but warn against censorship*, 08 March 2017, available at <http://www.ohchr.org/EN/NewsEvents/Pages/DisplayNews.aspx?NewsID=21317&LangID=E> (last accessed 13 March 2018).

²⁰¹ Amnesty International interview with Laura Bates, 25 January 2017.

²⁰² Amnesty International interview with Rachel* (Pseudonym), 27 April 2017.

²⁰³ National Democratic Institute, *Stopping Violence Against Women In Politics, Program Guidance*, 2017, available at <https://www.ndi.org/sites/default/files/not-the-cost-program-guidance-final.pdf> (last accessed 13 March 2018) p. 14.

²⁰⁴ National Democratic Institute, Website, available at <https://www.ndi.org> (last accessed 15 March 2018).

²⁰⁵ Amnesty International interview with Tasmina Ahmed-Sheikh, 11 July 2017.

²⁰⁶ Amnesty International interview with Nicola Sturgeon, 22 June 2017.

but on behalf of young women out there who are looking at what people say about me and thinking ‘ I don’t ever want to be in that position’.”

It is imperative that Twitter respects the right of all women to exercise their human rights online, including the right to freedom of expression. A failure to do so will have serious consequences for women’s participation in public life now and in the generations to come.

As UK Politician and Shadow Home Secretary Diane Abbott points out,²⁰⁷

“I think the online abuse I get makes younger women of colour very hesitant about entering the public debate and going into politics. And after all, why should you have to pay that price for being in the public space?”

²⁰⁷ Amnesty International interview with Diane Abbott, 24 July 2017.

SECTION 6: THE PSYCHOLOGICAL HARMS OF VIOLENCE AND ABUSE AGAINST WOMEN ONLINE

The psychological consequences of violence and online abuse remain under-researched, and as a result, understated. However, almost every woman interviewed by Amnesty International spoke about the adverse impact of violence and abuse on Twitter on their mental health. The findings of Amnesty International's online poll support the experiences of the women we interviewed, showing that the majority of women polled across the 8 countries²⁰⁸ who experienced abuse or harassment on social media platforms reported stress, anxiety, panic attacks, powerlessness and loss of confidence as a result.

"The abuse has definitely increased my anxiety. There are definitely days where I can't even work and I can't even focus because I'm in such a bad mood."

Imani Gandy, US journalist²⁰⁹

ONLINE VERSUS OFFLINE REALITIES

Some people may believe that when the violence and abuse women face is online, it can simply be ignored or shrugged off – but almost every woman who spoke to Amnesty International emphasized that this simply is not the case. In particular, women highlighted the interlinkages between their online and offline identities and stressed how violence and abuse online impacts their lives offline.

US reproductive rights activist and blogger Erin Matson explains,²¹⁰

"The distinction between our online and offline lives is a false distinction. There is an attitude that it's just name calling and the online world is not real and one's feelings getting hurt online is not a real problem. But the online world is real and our offline lives are absolutely integrated with the online space. The idea that we don't need online platforms to survive is false."

US activist Shireen Mitchell agrees. She states,²¹¹

"The conversation that the online and offline worlds are different is a key piece of the problem. The online equals the offline and the offline equals the online."

A majority of women in the UK and USA agree that abuse that takes place online can have an impact on women's lives more widely. In fact, in both countries, more than 3 times as many women polled²¹² disagree (63% and 61%) than agree (19% in both countries) that online abuse and harassment can be stopped by just ignoring it.

²⁰⁸ Amnesty International, *Amnesty reveals alarming impact of online abuse against women*, 20 November 2017, available at <https://www.amnesty.org/en/latest/news/2017/11/amnesty-reveals-alarming-impact-of-online-abuse-against-women/> (last accessed 12 March 2018).

²⁰⁹ Amnesty International interview with Imani Gandy, 15 August 2017.

²¹⁰ Amnesty International interview with Erin Matson, 26 June 2017.

²¹¹ Amnesty International interview with Shireen Mitchell, July 2017.

²¹² Azmina Dhrodia, *Unsocial Media: The Real Toll of Online Abuse against Women*, Amnesty International, Medium, 20 November 2017, available at <https://medium.com/amnesty-insights/unsocial-media-the-real-toll-of-online-abuse-against-women-37134ddab3f4> (last accessed 13 March 2018).

UK politician Ruth Davidson stressed to Amnesty,²¹³

“Just because you say something on a keyboard and not to someone’s face, it doesn’t mean it doesn’t matter. It does matter. It can have a huge impact on people. There needs to be an understanding of the seriousness of what this is – rather than the kind of frivolous ‘Oh they only said it on Twitter so it doesn’t matter’. Actually, it does”.



IMANI GANDY, US JOURNALIST

I think Twitter has become the new public square. I've found Twitter to be a really good platform for people who normally don't have as much of a say in the political process. I'm talking primarily young people and people of colour. There is a reason that 'Black Twitter' is a thing. It has been really powerful that black people have been able to come together to reach out to other black people across the country. Whether it's dishing about the latest episode of Scandal or organizing around Black Lives Matter – it has really become a powerful tool. And elite politicians are not used to having to answer to these types of people

I get harassment as a woman and I get the extra harassment because of race and being a black woman. They will call white women a 'c*nt' and they'll call me a 'n*gger c*nt'. Whatever identity they can pick they will pick it and use it against you. Whatever slur they can come up with for a marginalized group – they use.

The abuse I receive on Twitter is mostly filtered out and I have trained myself to not search my name because then I get really angry. I've been online for 8 years and political on Twitter for 7 years so after 7 years I guess you just get used to it...it has become a part of my life, which is kind of sad, I guess. The filtering on my part and using third party apps have made it better but I'm also trying to reduce my time on Twitter as well.

It's not fun, you get into this place where you think everybody hates me, what am I doing wrong? It definitely increased my anxiety and I'm already an anxious person. The first time someone called me n*gger on Twitter I literally cried...and then five minutes later I laughed. It's just one of those things you get used to, at this point I am mostly used to it.

THE TOLL OF VIOLENCE AND ABUSE AGAINST WOMEN ONLINE

It is important to remember that the United Nations Declaration on the Elimination of Violence against Women defines violence against women²¹⁴ as “any act of gender-based violence that results in, or is likely to result in, physical, sexual, *psychological* or economic harm or suffering to women...whether occurring in public or in private life”. Although little research exists about the psychological impact of violence and abuse online, Amnesty International’s qualitative and quantitative research demonstrates that violence and abuse against women on Twitter risks causing psychological harm and suffering to women. A study by UNESCO on ‘Building Digital Safety for Journalists’²¹⁵ also acknowledges that, “*Despite a lack of aggregated data and specific studies on the issue, experts say there is enough evidence to suggest that the online harassment of journalists, including threats of violence, has a serious psychological impact that may result in self-censorship.*”

Women who spoke to Amnesty International during the course of this investigation told us about increased anxiety, a loss of self-confidence, trouble sleeping, and an overall feeling of disempowerment after experiencing violence and abuse on Twitter. Interviews with mental health experts also detailed the negative

²¹³ Amnesty International interview with Ruth Davidson, 21 June 2017.

²¹⁴ UN General Assembly, *Declaration on the Elimination of Violence against Women, Resolution*, 20 December 1993, A/RES/48/104, available at <http://www.un.org/documents/ga/res/48/a48r104.htm> (last accessed 13 March 2018).

²¹⁵ International Press Institute, *Online abuse of journalists has deep psychological impact, experts say*, 10 April 2015, available at <https://ipi.media/online-abuse-of-journalists-has-deep-psychological-impact-experts-say/> (last accessed 13 March 2018).

psychological implications of online violence and abuse on women. Licensed mental health specialist and former Director of Online SOS,²¹⁶ Samantha Silverberg, told Amnesty International,²¹⁷

“There is little research examining the psychological toll online abuse has on individuals. Anecdotally, we can see the toll when individuals are fearful of opening their emails, unable to return to work, or are making other changes to their daily lives based on fears related to the abuse.

The lack of certainty around the practicality of threats made makes an individual that much more cautious as to how he or she interacts with the world around him or her. These threats do not just exist on the internet, they exist in reality, placing the threat anywhere due to this missing information on whether the threat simply exists online or may also exist in one's physical world. This feeling of not-knowing is pervasive and can drastically change how an individual engages in society.”

Samantha has also spoken²¹⁸ about other specific psychological consequences of online abuse, including: a decreased ability to concentrate, difficulty in making day-to-day decisions and increased levels of anxiety.

Many of the women interviewed by Amnesty International during the course of this investigation spoke in detail about the negative psychological impact of online abuse and the consequences on their day-to-day lives.

For example, US writer Chelsea Cain told Amnesty International how she felt after experiencing abuse on Twitter and deciding to leave the platform,²¹⁹

“I was a wreck -- a total wreck. It feels bad enough to get just one hate tweet...that is why I had to detox from Twitter in the first place. It was a positive mental health choice. Like exercising more, or giving up smoking. A lifestyle decision. Ironically, the fact that I had the audacity to delete my account is what really set everyone off - that's when things escalated. It felt like I was under siege in every way. I was afraid to get on the computer, to talk to anybody. I was afraid of what I was going to find.”

Similarly, UK Politician and activist Seyi Akiwowo also spoke about the harmful psychological impact a wave of abuse on Twitter had on her. She states,²²⁰

“I felt overwhelmed. Looking back on it, I wasn't okay. I was on auto-pilot and I didn't take the time to do so some self-care. I was constantly on my phone checking [Twitter]. The support helped drown out the hate but I wasn't sleeping well.”

US writer and blogger Alyssa Royce emphasized that even years after experiencing abuse on Twitter, she still feels apprehensive when receiving notifications on the platform,²²¹

“It felt scary going back online after the abuse I experienced. It was anxiety filled. If I see that I have a Twitter notification, I still get nervous, and its years later. It's often just my friends making jokes with me but literally seeing that I have a Twitter notification makes me so nervous...So, I'm not sure that anxiety ever went away. It's some form of social media PTSD. I laugh at it because it's usually nothing but the feeling is very real”.

²¹⁶ Online SOS, Website, available at <https://www.onlinesosnetwork.org> (last accessed 15 March 2018).

²¹⁷ Amnesty International interview with Samantha Silverberg, 02 June 2017.

²¹⁸ Tiffany McLain, *How to Cope With the Trauma of Online Harassment*, Psychology Today, 09 March 2017, available at <https://www.psychologytoday.com/blog/living-between-worlds/201703/how-cope-the-trauma-online-harassment> (last accessed 13 March 2018).

²¹⁹ Amnesty International interview with Chelsea Cain, 09 June 2017.

²²⁰ Amnesty International interview with Seyi Akiwowo, 30 March 2017.

²²¹ Amnesty International interview with Alyssa Royce, 06 June 2017.

Other women spoke to Amnesty International about how the abuse they receive on Twitter can come in waves and can therefore leave them feeling unprepared when they do experience it. UK activist Alex Runswick explains,²²²

“The abuse went on for months. I mean there was an immediate response to [my tweet], and I dealt with that. But a few weeks later I was on holiday and suddenly I had 25 notifications on my Twitter account and it was these people again. And that was what created anxiety for me because it was coming back repeatedly.”

Some women spoke specifically about what it felt like to have someone try to break them down via a social media platform.

UK journalist Vonny Moyes states,²²³

“It’s almost as if sometimes people are using your emotional state as sport. That you’re just a leisure pursuit for them and they enjoy watching someone being slowly broken down by it. Sometimes there have been occasions where I’ve just burst into tears when I think about it. It’s ruined the enjoyment of the online sphere for me.”

US games developer Zoe Quinn also told Amnesty International about the toll online abuse had on her life,²²⁴

“People online pushed me really hard to kill myself. My partner at the time didn’t leave my side for more than a few hours. For the first few days I couldn’t eat or sleep or drink water. All I could do was watch everything collapse around me. It was and it still is hard to get closer to new people [after going through that].”

UK actor and poet Travis Alabanza also spoke about becoming more guarded after experiencing abuse on Twitter and also described the overwhelming feeling of panic coming over them. They said,²²⁵

“Panic was the overwhelming feeling; that the abuse will ruin all future job prospects. Then sadness was definitely there too. The hardest is when you’re tweeting stuff that’s irrelevant to trans issues and they still pop up... My friends have noticed a huge difference in me since. I’m way more nervous outside, I’m way more guarded about my protection. I don’t disclose as much with new people and when I meet new people I’m way more suspicious...”

Comics writer Kelly Sue DeConnick spoke to Amnesty International about her fear of the impact abuse on the platform could have on her family. She told us,²²⁶

“What I fear is the kind of action that can be taken from behind a keyboard to jeopardize my safety or that of my family, our sense of well-being. ... I fear my children or my husband losing time to anxiety or fear because I wouldn’t keep my mouth shut on the internet. I fear feeling like it’s my fault for speaking up. I understand how backwards that is, but it’s a real thought.”

Amnesty International’s online poll²²⁷ reflects many of the experiences brought forward by the women interviewed when they were asked how violence and abuse on Twitter makes them feel. Alarming, of the women polled across the 8 countries²²⁸ who experienced abuse or harassment on social media platforms, 41% of women said that on at least one occasion, these online experiences made them feel that their physical safety was threatened. In addition, between well over half to two-thirds of women (55% to 67%) polled in the UK and the USA who experienced abuse or harassment on social media platforms stated that they were less

²²² Amnesty International interview with Alex Runswick, 27 April 2017.

²²³ Amnesty International interview with Vonny Moyes, 23 March 2017.

²²⁴ Amnesty International interview with Zoe Quinn, 13 December 2017.

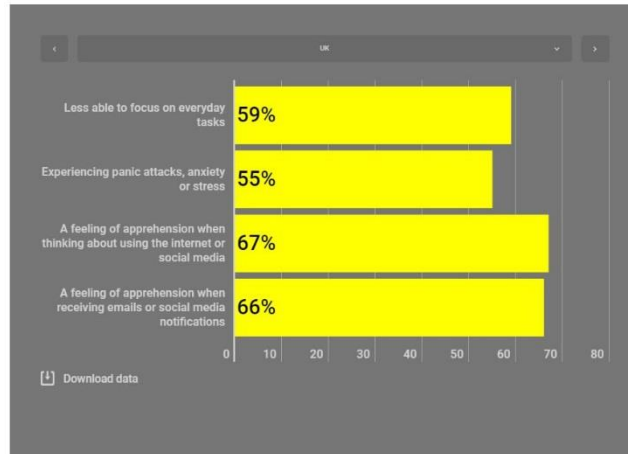
²²⁵ Amnesty International interview with Travis Alabanza, 24 January 2018.

²²⁶ Amnesty International interview with Kelly Sue DeConnick, 28 June 2017.

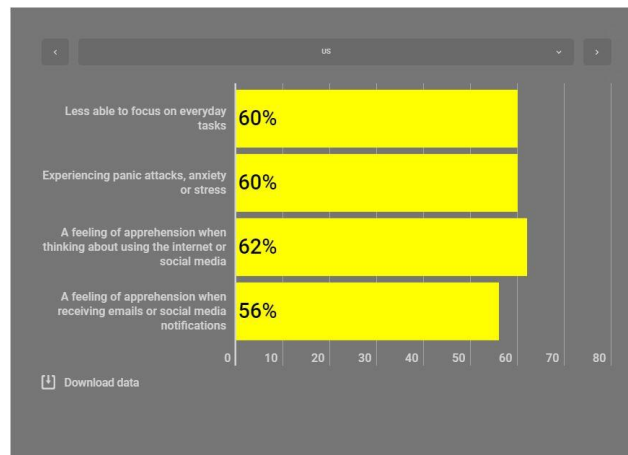
²²⁷ Azmina Dhrodia, *Unsocial Media: The Real Toll of Online Abuse against Women*, Amnesty International, Medium, 20 November 2017, available at <https://medium.com/amnesty-insights/unsocial-media-the-real-toll-of-online-abuse-against-women-37134ddab3f4> (last accessed 13 March 2018).

²²⁸ Amnesty International, *Amnesty reveals alarming impact of online abuse against women*, 20 November 2017, available at <https://www.amnesty.org/en/latest/news/2017/11/amnesty-reveals-alarming-impact-of-online-abuse-against-women/> (last accessed 12 March 2018).

able to focus on everyday tasks, had experienced stress, anxiety or panic attacks, and had a feeling of apprehension when thinking about social media or receiving social media notifications.



Results of Amnesty International’s online poll of the psychological impact on women in the UK who experienced abuse or harassment on social media platforms the



Results of Amnesty International’s online poll of the psychological impact on women in the UK who experienced abuse or harassment on social media platforms.



JESSICA VALENTI, US JOURNALIST AND WRITER

It is sort of hard to know where to begin, because it has been going on so long. I think Twitter is the worst of the social media platforms, just because of the quickened and masked flow of abuse that happens. The abuse on Twitter feels like a constant stream. It can include general nastiness or name calling (you b*tch, slut, c*nt). It can be more targeted harassment or can be more direct threats – which in the past I have had directed at my daughter. I’ve had my address released, my tax information released as well as my phone number released.

Twitter has this thing where you can turn off notifications from anyone who you don’t follow back. It’s a good and a bad thing. It vastly improved my day to day experience, but it’s bad in terms of I don’t get to hear

from my readers. So that sucks. The connection with my readers was a big part of my work for 10 years. The reason I am online is to connect with people, so it does feel like a big loss to me.

What's hard about online harassment is the response that you tend to get will be, "just don't be on that platform. Just don't be on Twitter, etc." Which is really akin to saying to someone "well just don't walk alone on the street at night". These are public spaces, this is the new town square. In order for women to fully participate in life we need to be in these spaces. Often times, especially if you're a writer, it's almost informally required. To tell women to just not be there is absurd.

Knowing what I know now I don't know if I would have chosen to write under my real name. I think I wouldn't have. The abuse doesn't just impact me, it impacts my family. On a really bad day when I've gotten a lot of threats it's not like I can just go to the police and report it. I can't get out of bed, I'm devastated, and my husband has to take care of me. It has a rippling effect that goes beyond your Twitter account.

Amnesty International's online poll²²⁹ also found that across all countries surveyed, around 2/3 of women who experienced abuse or harassment on social media platforms stated feeling a sense of powerlessness after experiencing online abuse or harassment.

UK Science Broadcaster, Writer and Educator Dr Emily Grossman explained to Amnesty International how online abuse on Twitter took a toll on her self-confidence. She said,²³⁰

"The abuse made me feel like I was an embarrassment, not a valid person, and that I don't have valid opinions. It was disempowering and humiliating – it was a removal of my agency."

US writer and blogger Imani Gandy told Amnesty International about the self-doubt she feels after receiving abuse on Twitter. She told us,²³¹

"It's not fun, you get into this place where you think everybody hates me, what am I doing wrong? It definitely increased my anxiety and I'm already an anxious person. The first time someone called me n*gger on Twitter I literally cried...and then five minutes later I laughed. It's just one of those things you get used to, at this point I am mostly used to it."

Amnesty International's research indicates that the psychological impact of online abuse is both real and harmful for women. Dr Emma Short, a psychologist and reader in cyber psychology at the University of Bedfordshire agrees. She explains,²³²

"I think the impact of online abuse is greater because your victimization is broadcast for everyone to see. It's often joined by a third party so the crowd or pack is going after you. So, very quickly, it feels as though the whole world is after you. There might be positive tweets, you might have lots of friends on the outside, but if the crowd has turned against you and is after you, it feels like the world wishes you harm."

She also highlighted a study²³³ she conducted in 2010 on the impact of cyberstalking—that is, when a person is engaged in a persistent course of conducting online abuse that causes alarm, distress and fear to another person. She stated,

"A third of people who experienced cyber stalking and who were still online reported symptoms of post-traumatic stress disorder, which is enduring and very rarely alleviates without professional support.... The fear described by people who have been cyberstalked is undoubtedly a very real fear."

²²⁹ Azmina Dhrodia, *Unsocial Media: The Real Toll of Online Abuse against Women*, Amnesty International, Medium, 20 November 2017, available at <https://medium.com/amnesty-insights/unsocial-media-the-real-toll-of-online-abuse-against-women-37134ddab3f4> (last accessed 13 March 2018).

²³⁰ Amnesty International interview with Emily Grossman, 18 July 2017.

²³¹ Amnesty International interview with Imani Gandy, 15 August 2017.

²³² Amnesty International interview with Emma Short, 07 February 2018.

²³³ National Centre for Cyberstalking Research, *Cyberstalking in the United Kingdom*, 2011, available at https://www.beds.ac.uk/_data/assets/pdf_file/0003/83109/ECHO_Pilot_Final.pdf (last accessed 13 March 2018).

More research into the psychological impact of violence and abuse on Twitter, and social media platform more generally, is needed to further understand the full extent of the problem as well as the appropriate and adequate remedies required. However, it is clear that violence and abuse against women on Twitter risks contributing to psychological harms and this should undoubtedly be a serious cause for concern for the platform.

SECTION 7: HUMAN RIGHT'S RESPONSIBILITIES

TWITTER'S FAILURE TO RESPECT WOMEN'S HUMAN RIGHTS

Companies, wherever they operate in the world, have a responsibility to respect all human rights. This is an internationally endorsed standard²³⁴ of expected conduct. The corporate responsibility to respect requires Twitter to take concrete steps to avoid causing or contributing to human rights abuses and to address human rights impacts with which they are involved, including by providing effective remedy for any actual impacts. It also requires them to seek to prevent or mitigate adverse human rights impacts directly linked to their operations or services by their business relationships, even if they have not contributed to those impacts. In practice, this means Twitter should be assessing – on an ongoing and proactive basis – how its policies and practices impact on users' rights to non-discrimination, freedom of expression and opinion, as well other rights, and take steps to mitigate or prevent any possible negative impacts.

Amnesty International acknowledges that Twitter has recently taken steps towards addressing the problem of violence and abuse against women on the platform. In a response to the organization (see Annex II), Twitter reiterated that abuse and hateful conduct directed at women are prohibited on the platform, and characterized that this issue is one that they are “energized and motivated” to address. The company also highlighted several positive changes to their policies and practices in response to violence and abuse on the platform over the past 16 months, including a tenfold increase in the number of abusive accounts against which action has been taken.

However, Amnesty International believes that Twitter is failing to adequately meet its responsibility to respect human rights in the context of violence and abuse against women on the platform as the steps it has taken are not sufficient to tackle the scale and nature of the problem. Women have the right to live free from discrimination and violence. They also have the right to freely express themselves, both online and offline. Twitter's policies and – in particular – its practices clearly fail to respect these rights. As one of the world's leading social media platforms with over 330 million monthly users, this failure has a serious impact as it contributes to the silencing of women's voices online.



MISKI NOOR, US BLACK LIVES MATTER ACTIVIST

I manage our like Black Lives Matters Twitter page. I usually see a ton a ton of violence come in there on a daily basis and just awful things that people say to black organisers. There are folks who have dedicated accounts that troll other people. That's just how they spend their time. I think that's why I'm like “oh the abuse” - it's sort of normalised. There are people who dedicate their existence to sending messages to f*ck with my mental health and black organisers' mental health to tear us down.

There are people who always focus on ‘you're an immigrant, you're queer, you're trans, you're Muslim, or you're disabled’. They know that we hold these identities and they are actively trying to destroy folks. Abuse online is not just a ‘thing’ that's happening, it's a tactic, it's a strategy that they're using to stop us from building power. It's just the nature of Twitter and the way that it allows hateful folks to just fester on the platform to the point where it's like, “oh, this is what it is”.

²³⁴ UN Guiding Principles on Business and Human Rights, 2011, available at http://www.ohchr.org/Documents/Publications/GuidingPrinciplesBusinessHR_EN.pdf (last accessed 13 March 2018).

Twitter needs to hone in on their responsibilities and their values. I'm tired of tech companies or social media companies thinking they are exempt from living their values. If Twitter values women and femmes, if they value our safety, then they need to have practices that they actually develop and implement in real ways that will protect us.

Twitter is going to have to say whether they're for the people or they're not. Twitter has the power to change the way women and femmes are experiencing abuse, or even if we experience abuse, on their platform. After all, they are the convenors of the space and they have the power to change our experiences.

In a response to the organization (see Annex II), Twitter said it disagreed with Amnesty International's findings and that it "cannot delete hatred and prejudice from society".

LACK OF TRANSPARENCY

Transparency is a key component of Twitter's human rights responsibilities. In this regard, Twitter's reporting mechanisms must be accessible and transparent.²³⁵ It is impossible to assess the effectiveness of these mechanisms for social media platforms more generally when companies like Twitter give little information about their internal review processes, including, how complaints are dealt with, the ratio of company moderators to the volume of reports, the type and level of gender-and other identity based human rights training the staff receives and the time limits and targets for reviewing reports.²³⁶

According to the UN Guiding Principles on Business and Human Rights,²³⁷ the responsibility to respect human rights involves having both policies and processes through which businesses can both "know and show that they respect human rights in practice". 'Showing' includes "providing a measure of transparency and accountability to individuals or groups who may be impacted (such as users) and to other relevant stakeholders, including investors."²³⁸ The Guiding Principles also state that companies should communicate how they respect human rights in practice in a number of ways, such as formal reporting that provides indicators on how they identify and address adverse impacts on human rights.

In July 2017, Twitter stated²³⁹ they were taking action on 10 times the number of abusive accounts every day compared to the same time last year. They also said they had removed twice the number of accounts of repeat offenders who created new accounts after being suspended. In terms of user behaviour, Twitter also stated that accounts that had been put into a period of limited functionality following a violation of the Twitter rules generate 25% fewer abuse reports, and approximately 65% of these accounts only enter this state once. While these statistics indicate an improved response to violence and abuse on the platform, they do not provide sufficient information to understand how significant this progress is relative to the overall scale of the problem.

In January 2018 Amnesty International wrote to Twitter asking for disaggregated figures in relation to the number of reports of abuse on the platform and the number of reports found to be in violation of Twitter rules among other requests for statistics. In their response (Annex I), Twitter outlined a number of steps it is taking

²³⁵ Due Diligence for Online Violence against Women

<https://www.apc.org/sites/default/files/DueDiligenceAndAccountabilityForOnlineVAW.pdf> page 20

²³⁶ Transparency is also an important component of the right to remedy for those users whose content is taken down or otherwise censored, and allows for more meaningful comparisons to be made as to how social media platforms respond to different types of reports, such as those by governments, or related to different topics and thus help avoid a resort to overly restrictive policies around certain types of content.

²³⁷ *UN Guiding Principles on Business and Human Rights*, 2011, available at

http://www.ohchr.org/Documents/Publications/GuidingPrinciplesBusinessHR_EN.pdf (last accessed 13 March 2018) p. 16.

²³⁸ *UN Guiding Principles on Business and Human Rights*, 2011, available at

http://www.ohchr.org/Documents/Publications/GuidingPrinciplesBusinessHR_EN.pdf (last accessed 13 March 2018) p. 24.

²³⁹ Ed Ho, *Our Safety Work: Results Update*, Twitter blog, 20 July 2017, available at

https://blog.twitter.com/official/en_us/topics/product/2017/Our-Safety-Work-Results-Update.html (last accessed 13 March 2018).

to combat violence and abuse on the platform but ultimately refused our request. Twitter stated that this information can be both uninformative and potentially misleading because users regularly report content with which they disagree or, in some cases, with the direct intent of trying to silence another user's voice for political reasons (see Annex I). Twitter also stated that there is a misperception that the volume of reports impacts their enforcement decisions. Whilst Amnesty International agrees that such figures must not be taken as providing a complete picture of violence and abuse on the platform – detailed statistics on reports of abuse can help set a baseline, and potentially targets, for response times to reports of abuse.

ENABLING AND EMPOWERING USERS

Enabling and empowering users to create a safer and less toxic Twitter experience is a key component of Twitter's responsibility to respect human rights on the platform. Part of this responsibility means that Twitter must enable and empower users to understand and utilize individual security and privacy measures such as blocking, muting and content filtering so women are easily able to curate a less toxic and harmful online experience.

Twitter has introduced a number of security and privacy features²⁴⁰ to help users protect themselves from violence and abuse. Users can block accounts, mute notifications or conversations, and filter out tweets containing specific language they prefer not to see. It is important that Twitter not only develops such features but also ensures that they are accessible and easy for users to utilize. Twitter should pay particular attention to educating and empowering users who may be targeted with abuse on the basis of their gender or other forms of their identity. However, it is important to stress that Twitter must pay sufficient attention to both equipping users to confidently use security and privacy features and also consistently enforcing the Twitter Rules to avoid placing undue burden on users to keep themselves safe on the platform.

ENSURING FREE EXPRESSION FOR EVERYONE

Ensuring that women can express themselves online on the basis of equality may require restricting some forms of expression. It is crucial that states and companies ensure that these efforts do not result in unlawful censorship.

The right to freedom of expression may legitimately be subject to restrictions, provided that such restrictions comply with the requirements of international human rights law, including that they serve a legitimate purpose, are provided by law and are necessary and proportionate. Many forms of violence and abuse against women, such as direct threats of physical or sexual violence, are widely considered to be illegal in many domestic systems, and this is generally consistent with the right to freedom of expression. Where such acts are considered crimes under national law, companies and respective governments must work together to address this. Twitter should also ensure that in so doing, they do not reveal sensitive user data except in response to valid court orders that comply with international human rights law.

Companies have a responsibility to respect free expression, which encompasses expression which may be offensive or disturbing. The International Covenant on Civil and Political Rights, for example, requires states to prohibit – though not necessarily through the criminal law – only “any advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence.” Many other forms of expression, even those which shock or offend, may not lawfully be restricted.

This means that not all forms of online abuse against women may be legitimately subject to criminal sanctions or take-down measures. Laws or policies aimed at so-called “hate speech” must be carefully crafted to ensure they do not lead to unlawful censorship, including of the very groups they may seek to protect.²⁴¹ This means,

²⁴⁰ Twitter, *Help Center, How to control your Twitter experience*, <https://help.twitter.com/en/safety-and-security/control-your-twitter-experience> (last accessed 13 March 2018).

²⁴¹ See for example, UN Human Rights Council, *Annual report of the United Nations High Commissioner for Human Rights*, 11 January 2013, A/HRC/22/17/Add.4, available at http://www.ohchr.org/Documents/Issues/Opinion/SeminarRabat/Rabat_draft_outcome.pdf (last accessed 13 March 2018).

among other things, that laws or policies that restrict the right to free expression must not be overly vague, but rather, “must be formulated with sufficient precision to enable an individual to regulate his or her conduct accordingly.”²⁴² Additionally, people whose content is taken down or otherwise restricted must be given clear reasons why and a meaningful opportunity to appeal against these restrictions.²⁴³

Even forms of violence and abuse against women online that may be not be lawfully prohibited, can negatively impact on women’s human rights. As a result, social media companies like Twitter have varying levels of responsibility depending on the degree of violence or abuse against women occurring on their platforms.

However, the imposition by states of legal liability for companies who fail to remove abusive content sets a dangerous precedent and risks causing more harm instead of addressing the core of the issue. Such penalties risk unintended consequences such as the overbroad application of existing company policies to avoid liability which can lead to negative repercussions for the right to freedom of expression of all individuals, including the censorship of legitimate expression.²⁴⁴ In this case, attempting to solve one freedom of expression issue but creating another is simply not the answer.

THE DUTIES OF STATES UNDER INTERNATIONAL HUMAN RIGHTS LAW

While private companies have responsibilities under human rights law, states are considered the primary duty bearers. Specifically, states are obliged to respect, protect and fulfil all human rights for everyone. This means that states must not only refrain from interfering with the exercise of human rights, but must also protect the exercise of rights from interference by private parties, and take proactive measures to ensure the enjoyment of human rights.²⁴⁵

These obligations include the duty not only to tackle violence and abuse online, but to address the underlying causes of such abuse, including by ensuring the right to non-discrimination in the enjoyment of all human rights. This task should be addressed with a broad set of policy initiatives aimed at promoting minority and under-represented voices, fostering tolerance and understanding and condemning discrimination and intolerance wherever it arises.²⁴⁶ Overall, states must ensure that there are adequate laws in place to prevent and end online violence and abuse against women and must also combat negative and harmful gender stereotypes against women that contribute to the manifestation of violence and abuse against women online.

²⁴² UN Committee on the Elimination of Discrimination against Women, *General recommendation No. 35 on gender-based violence against women, updating general recommendation No. 19*, 14 July 2017, CEDAW/C/GC/35, available at http://tbinternet.ohchr.org/Treaties/CEDAW/Shared%20Documents/1_Global/CEDAW_C_GC_35_8267_E.pdf (last accessed 12 March 2018).

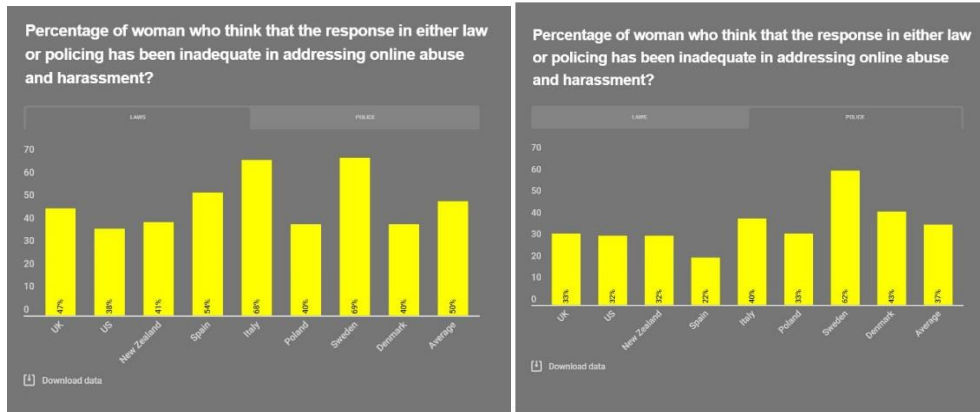
²⁴³ UN Human Rights Council, Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, 11 May 2016, A/HRC/32/38, available at http://www.ohchr.org/EN/HRBodies/HRC/RegularSessions/Session32/Documents/A_HRC_32_38_EN.docx (last accessed 13 March 2018).

²⁴⁴ Special Rapporteur David Kaye states, “Provisions imposing high fines for non-compliance with the obligations set out in the bill raise concerns, as these obligations as mentioned above may represent undue interference with the right to freedom of expression and privacy. The high fines raise proportionality concerns, and may prompt social networks to remove content that may be lawful.” David Kaye, *Mandate of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression*, 01 June 2017, available at <http://www.ohchr.org/Documents/Issues/Opinion/Legislation/OL-DEU-1-2017.pdf> (last accessed 13 March 2018) p. 4.

²⁴⁵ UN Human Rights Committee, General Comment No. 31 The Nature of the General Legal Obligation Imposed on States Parties to the Covenant, 26 May 2004, CCPR/C/21/Rev.1/Add.13, available at <http://docstore.ohchr.org/SelfServices/FilesHandler.ashx?enc=6QkG1d%2FPPrICAqhKb7yhsjYoiCfMKoIRv2FVaVzRkMiTnjRO%2Bfud3cPVrcM9YR0IW6Txaq3f9kUJfPWoq%2FhW%2FTpKI2tPhZsbEJw%2FGeZRASjdFuuJQRnbJEaUhby31WlQPI2mLFD6ZSwMMvmQGVHA%3D%3D> (last accessed 13 March 2018); *UN Guiding Principles on Business and Human Rights*, 2011, available at http://www.ohchr.org/Documents/Publications/GuidingPrinciplesBusinessHR_EN.pdf (last accessed 13 March 2018).

²⁴⁶ See Rabat Plan of Action on the prohibition of advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence http://www.ohchr.org/Documents/Issues/Opinion/SeminarRabat/Rabat_draft_outcome.pdf

Amnesty International's online poll²⁴⁷ shows that many women believe their governments have a lot more work to do on this issue: in the eight countries polled, half of all women polled stated the current laws to deal with online abuse or harassment were inadequate. In the UK and USA, around 1 in 3 women stated the police response to abuse and harassment online was inadequate.



Results of Amnesty International's online poll across 8 countries on percentage of women who think response in law or policing in addressing abuse and harassment online has been inadequate.

²⁴⁷ Azmina Dhroia, *Unsocial Media: The Real Toll of Online Abuse against Women*, Amnesty International, Medium, 20 November 2017, available at <https://medium.com/amnesty-insights/unsocial-media-the-real-toll-of-online-abuse-against-women-37134ddab3f4> (last accessed 13 March 2018).

SECTION 8: SOLUTIONS

Violence and abuse against women on this scale does not have to exist on Twitter. The company's failure to adequately meet its human rights responsibilities regarding violence and abuse will continue to silence women on the platform unless Twitter undertakes, with urgency, concrete steps to effectively tackle this problem.

"Twitter is going to have to say whether they're for the people or they're not. Twitter has the power to change the way women and femmes are experiencing abuse, or even if we experience abuse, on their platform. After all, they are the convenors of the space and they have the power to change our experiences."

Miski Noor, US Black Lives Matter activist

Summary of Recommendations:

- Twitter should publicly share comprehensive and meaningful information about the nature and levels of violence and abuse against women, as well as other groups, on the platform, and how they respond to it.
- Twitter should improve its reporting mechanisms to ensure consistent application and better response to complaints of violence and abuse.
- Twitter should provide more clarity about how it interprets and identifies violence and abuse on the platform and how it handles reports of such abuse
- Twitter should undertake far more proactive measures in educating users and raising awareness about security and privacy features on the platform that will help women create a safer, and less toxic Twitter experience.

Amnesty International is asking Twitter to:

1. Publish meaningful data on how they handle violence and abuse. This should include:

- The number of reports of abusive or harmful conduct Twitter receives per year. This should include how many of these reports are for directing 'hate against a race, religion, gender, or orientation', 'targeted harassment' and 'threatening violence or physical harm'. Twitter should also specifically share these figures for verified accounts on the platform.
- Of the aforementioned reports of abuse, the number of reports that are found to be – and not be - in breach Twitter's community guidelines, per year and by category of abuse. Twitter should also specifically share these figures for verified accounts on the platform.

- The number of reports of abuse it receives per year that failed to receive any response from the company, disaggregated by the category of abuse reported.
- The average time it takes for moderators to respond to reports of abuse on the platform, disaggregated by the category of abuse reported. Twitter should also specifically share these figures for verified accounts on the platform.
- The proportion of users who have made complaints against accounts on the platform and what proportion of users have had complaints made against them on the platform, disaggregated by categories of abuse;
- The above information should be shared in an easy and accessible way on its platform.

2. Improve reporting mechanisms, for example by::

- Ensuring that decisions to restrict content are consistent with international human rights law and standards, are transparent, and allow for effective appeal
- Adding an optional question for users who receive a notification about the outcome of any reports on whether or not they were satisfied with Twitter's decision. Twitter should annually share and publish these figures, disaggregated by category of abuse
- Giving users the option to provide a limited character count of context when making reports of violence or abuse to help moderators understand why a report has been made. Twitter should eventually test user satisfaction against reports with an added context and reports without an added context.
- Providing clear guidance to all users on how to appeal any decisions on reports of abuse and clearly stipulating in its policies how this process will work.
- Sharing information with users who have filed a report of violence and abuse with links and resources for support and suggestions on how to cope with any negative or harmful effects.
- Creating public campaigns and awareness amongst users about the harmful human rights impacts of experiencing violence and abuse on the platform, particularly violence and abuse targeting women and/or marginalized groups. This should include sending a notification/message to users who are found to be in violation of Twitter's rules about the silencing impact and risk of mental health harms caused by sending violence and abuse to another user.
- Notifying users found to be in violation of the Twitter rules when their conduct may also violate the domestic law of their country.
- Creating public campaigns on Twitter encouraging users to utilize reporting mechanisms on behalf of others experiencing violence and abuse. This can help foster and reiterate Twitter's commitment to ending violence and abuse on the platforms and recognize the emotional burden the reporting process can have on users experiencing the abuse.
- Sharing and publishing the number of appeals received for reports of abuse, and the proportion of reports that were overturned in this process, disaggregated by category of abuse.

3. Provide more clarity about how abuse reports are handled, for example by:

- Share specific examples of violence and abuse that Twitter will not tolerate on its platform to both demonstrate and communicate to users how it is putting its policies into practice.
- Share with users how moderators decide the appropriate penalties when accounts users are found to be in violation of the Twitter Rules.
- Share and publish the number of content moderators Twitter employs, including the number of moderators employed per region and by language.
- Share how moderators are trained to identify gender and other identity-based violence and abuse against users, as well as how moderators are trained about international human rights standards and

Twitter's responsibility to respect the rights of users on its platform, including the right for women to express themselves on Twitter freely and without fear of violence and abuse.

4. Improve security and privacy features, for example by:

- Create public campaigns and awareness on Twitter about the different safety features users can enable on the platform. Such campaigns could be promoted to users through various channels such as: promoted posts on Twitter feeds, emails, and in-app notifications encouraging users to learn how to confidently use various safety tools.
- Offer personalized information and advice based on personal activity on the platform. For example, share useful tips and guidance on privacy and security settings when users make a report of violence and abuse against them. This should be tailored to the specific category of abuse users report. For example, a person reporting against targeted harassment could be advised how to protect themselves against fake accounts.
- Clearly communicate any risks associated with utilizing security features alongside simple ways to mitigate against such risks. For example, if users are taught how to mute notifications from accounts they do not follow - the risk of not knowing about any threats made against them from such accounts should be explained alongside practical ways to mitigate against such risks (e.g. having a friend monitor your Twitter account).
- Provide tools that make it easier for women to avoid violence and abuse, such as a list of abusive key words associated with gender or other identity-based profanity or slurs that users can choose from when enabling the filter function. An additional feature could allow users to easily share keywords from their mute lists with other accounts on Twitter.

Amnesty International is calling on States to:

- Enact and implement adequate legislation, including, where appropriate, criminal penalties (in line with international human rights law and standards) in relation to violence and abuse against women online
- Prioritize and invest in capacity building and training of law enforcement bodies on relevant legislation, gender equality, the harms of online violence and abuse, and best practices to support those who have experienced online violence and abuse
- Invest in public awareness raising campaigns about violence and abuse online and public campaigns to promote gender equality and combat sex- and gender-based stereotypes
- Governments should ensure that sex and gender stereotyping online is included in comprehensive sexuality/sex and relationships education and that teachers are trained to such education
- Invest in specialist public services for women who have experienced violence and abuse online.

ANNEX I: LETTER FROM TWITTER TO AMNESTY INTERNATIONAL (14 FEBRUARY 2018)



Twitter, Inc.
1355 Market Street
Suite 900
San Francisco, CA 94103

14 February 2018

Dear Joshua,

Thank you for your letter regarding user safety and for your colleagues taking the time to meet with us at Twitter on Monday. As you may know, Twitter's co-founder and CEO Jack Dorsey said last year that safety is Twitter's top corporate priority. This continues to be the case. Your letter raises some important questions and relates to a number of ongoing areas of work for Twitter.

Meaningful transparency is a question we share your passion for. The question of absolute numbers of reports and the proportion of accounts that are actioned is one on which we continue to reflect. However, it is important to note that this type of information can be both uninformative and potentially misleading. One of the challenges, as we have discussed with your Amnesty colleagues previously, is that users regularly report content with which they disagree or in some cases, with the direct intent of trying to silence another user's voice for political reasons. Moreover, there is a misperception that the volume of reports impacts our enforcement decisions, but this is not the case. In our meeting you discussed potential targets in this space, and we would be interested to hear any ideas you might have.

For example, when we launched our violent extremist groups policy in December 2017, we saw calls to bulk report certain accounts which did not contravene this policy yet were politically opposed to by other users. As such, the volume of reports versus the proportion of accounts suspended ratio is likely to be skewed by such malicious reports. As we strive to defend our users against these efforts, suggesting a low proportion of removals is a sign of inaction on safety would avoid critical context and be misleading to the wider public debate.

We are currently considering how to better inform the conversation about the impact of our safety work, following on from our safety results update [blog post](#) last summer. Since that blog we have instituted a number of [new policies](#) and platform updates to fight abuse and harassment and to take action on accounts that promote or celebrate violence. As we are able to share relevant metrics related to those and other updates we will keep you updated.

We have teams of staff working around the world to provide timely responses to reports 24/7. Our goal is to respond to all reports promptly

and we are investing in the teams and technology required to make even more progress. For example, by leveraging technology we are able to take action on ten times the number of abusive accounts every day compared to the same time last year. This combination of technology and staff now embedded into our safety approach allows us to scale our efforts more quickly. We continue to consider how we can better inform our users and stakeholders on our safety efforts.

Our review teams are empowered to use their judgement and take appropriate action on accounts that violate our rules. This can involve requiring a user to confirm a phone number, delete a specific Tweet, placing their account in a locked time-out, or permanent suspension. This range of enforcement approaches allows us to respond proportionately and retain the ability to immediately remove those whose behaviour warrants it. As part of this process the teams take into account important and relevant context, for example considering any previous activity taken, while evaluating the severity and nature of the breach.

We use technology to supplement this process and our approach is showing positive signs of improving behaviour. Accounts that demonstrate abusive behavior are now limited for a time and told why. Accounts that we put into this period of limited functionality generate 25% fewer abuse reports, and approximately 65% of these accounts are in this state just once.

You are right to highlight the need for cultural sensitivity and awareness across our safety organization. We recognize that diverse perspectives matter and it is of critical importance that our teams have training and understanding of human rights issues and how they intersect with content on the platform.

With respect to privacy, we do not currently require users to disclose their gender as part of using a Twitter account. Our approach to privacy is one of privacy-by-design, including data minimization when it comes to the collection of personal information, an approach which we hope Amnesty supports. However, we are well aware that ensuring our teams are taking into account gender-based nuance in behavior is essential. Every agent handling abuse reports receives in-depth trainings around the nuanced ways historically marginalized groups are harassed, intersectionality, and the impact that abuse online has on people's everyday lives, across a range of issues.

Furthermore, our Trust and Safety Council encompasses organizations from around the world with an extremely broad range of expertise and knowledge. We share a variety of information with members, sometimes with a small group based on a specific subject matter, for example suicide prevention, and other topics we may share with the full Council.

The feedback we receive is incorporated at the highest levels of our decision making processes. We continue to consider new members, particularly from areas of the world where Twitter is more recently being used.

The Council's Impact is highlighted in a [blog](#) published outlining the policy development process, published in October 2017, where we explain how we seek the council's views ahead of decisions being made on policy changes. While we do not always take every piece of feedback onboard and, in some circumstances members do not share the same perspective, it is an invaluable part of our safety effort.

With regard to 'fake' accounts, there are a number of issues here. Firstly, Twitter does not require the use of a legal identity or name. This allows journalists and human rights activists, as well as citizens of oppressive regimes and members of vulnerable groups, to express themselves around the world every day. We prohibit accounts that impersonate another person or organization, but do allow parody accounts.

We will take action on any account that breaks our rules irrespective of the state of the identity of the account.

You are correct to highlight that there are repeat offenders who create new accounts after being suspended for violations. We continue to invest in technology and tools to spot these individuals and our new systems removed twice the number of these types of accounts between April and July 2017. Overall, as of December 2017, we are identifying and challenging 6.4 million suspicious accounts per week.

With regard to machine learning, this is one of the areas of greatest potential for tackling abusive users. We announced last year that our team has also been working on identifying and collapsing potentially abusive and low-quality replies so the most relevant conversations are brought forward. These Tweet replies will still be accessible to those who seek them out. We are also using machine learning to bolster our approach in a number of areas, from better prioritising report signals to identifying efforts to circumvent suspensions.

Our approach to policy development involves a number of processes that ensure we carefully consider the impact. We posted [a blog](#) last year that detailed the process of making policy changes. We work with external partners and internal stakeholders to evaluate issues and potential impact of any changes. Specifically we consider the impact on free expression, both from the perspective of voices being silenced by others, and where the boundaries on any limitations on speech should fall. Our Trust and Safety Council is an invaluable part of this process but we are always looking for insight from a range of voices, from users

to organisations with specific expertise.

Where these processes identify potential issues, or issues arise following any changes, we ensure that such impacts are either mitigated, reversed or properly acknowledged and explained.

In 2015, Twitter [introduced the ability](#) for users to generate a single email, following a report of a violent threat, with the information required for law enforcement in one place. This includes details on how information can be requested as part of any subsequent investigation. We respond to such requests every day and you can read more about the process [here](#). We take the responsibility to protect the privacy of our users while interacting with law enforcement with due legal process seriously.

Our [Transparency Report](#) details where and how often we produce information to law enforcement agencies.

As you can appreciate, our strongest tool to promote our tools and safety updates to our users is directly through Twitter, often utilizing tools not commercially available - for example, emailing our users or prompting them to review something upon opening the Twitter app. Accordingly, we may not need to spend money to reach our users in the way a brand would. Ensuring our users are aware of our rules and tools is a key objective for us and one that we have recently

These are complex issues and face different challenges according to the unique characteristics of each platform, as I expect will be highlighted by other platforms in response to your questions. We are committed to doing more to protect the voice of our users and in 2018 will redouble our efforts to make Twitter a safe place for free expression.

Twitter admires and supports the work that Amnesty International performs around the world in the field of human rights and we value our partnership with your organization.

We look forward to discussing these issues with you further. As always, we would be happy to meet your team to discuss safety-related issues, ongoing research, and future plans.

Best wishes,



Vijaya Gadde
General Counsel

ANNEX II: LETTER FROM TWITTER TO AMNESTY INTERNATIONAL (15 MARCH 2018)



Sherif El-Sayed Ali
Director, Global Issues and Research
Amnesty International
International Secretariat
Peter Benenson House
1 Easton Street
London WC1X 0DW
United Kingdom

15 March 2018

Dear Sherif,

Thank you for taking the time to talk with me yesterday afternoon.

As I outlined in our conversation, we do not believe that the findings of your upcoming report about Twitter and violence against women adequately reflect our engagement with Amnesty International on this important issue, including our extensive written reply and a meeting with our General Counsel.

I note your assurances that the report itself is fair and reflects the significant progress we have made to enhance the safety and security of all users. However, the wording and tone of the findings you have shared raises concern on our part about whether that progress and commitment is indeed reflected in the report. It is disappointing that you have declined to share the full draft report and enable us to offer a meaningful reply and, more importantly, an opportunity to engage constructively towards finding solutions to a challenge that is experienced across all online platforms and society at large.

Twitter is an open platform and often holds a mirror up to human behaviours - both the good and the bad. Twitter cannot delete hatred and prejudice from society, however we do remain committed every day to building on the major steps we have already taken to make Twitter safer.

Abuse and hateful conduct directed at women, including direct threats of violence, and harassment, are prohibited on Twitter. We have made more than 30 individual changes to our product, policies and operations in the past 16 months, all with the goal of improving safety for everyone. We now take action on 10 times the number of abusive accounts as the same time last year. Our policy, product and engineering teams continue to work collaboratively to find ways to innovate to protect our users and enhance their experience, particularly as those in our society intent on harm find new ways to hurt and abuse. We are also hiring a Director of Social Science research, to better explore opportunities to improve behaviour and partner with outside academics and leading organisations.

Twitter International Company	One Cumberland Place Fenian Street Dublin 2, D02 AX07 Ireland	Registered Number 503351	Directors L. O'Brien, R. O'Shea, V. Gadde (US)
-------------------------------	--	-----------------------------	---

In recent months, we have seen extraordinary engagement supporting women on Twitter. The rise of movements like #MeToo, #WomensMarch, and #PositionOfStrength are testimonies to the power of Twitter as a platform for women and their allies to share stories, offer support, and organize to advocate for change and growth. Twitter has been proud to support these initiatives, and we would expect that your report notes these movements, their real-world impacts, and our corporate support for their work.

Everyone has a part to play in building a more compassionate, empathetic, and resilient society. Twitter works with organisations around the world to highlight and reinforce what civil, healthy online engagement looks like, to publicise and encourage use of our safety tools, and to stress the importance of reporting breaches of our rules so we can take action. In an effort to meet an industry-leading level of transparency, we have opened up a process to request proposals from experts to join an initiative to promote a healthy public sphere and to assess the state of public discourse on our platform. It is important to note that our public API is accessible to all; this is not true of most other technology companies.

On behalf of the company, our global policy team partners with human rights organisations around the world. Twitter has stood up to nation states, consistently refused unfounded requests for user data, sued the U.S. government, partnered with the U.N. as part of our Data for Good initiatives, and funded and supported women's rights organisations worldwide. Despite the fact that it often presents reputational challenges, we stand proudly by our commitment to anonymity and have a strong track record of countering efforts to surveil and undermine those who use our platform to shine a light on human rights abuses.

Twitter doesn't claim to have all the answers. Indeed, in March our CEO Jack Dorsey stated this during a live Periscope in which he openly and transparently addressed many of these challenges. However, we are a passionate and committed group of people who believe in the power of the platform but understand also that it often reflects the worst as well as the best of society. We know we have a road to travel and that it is unlikely that we will ever reach a position where we can guarantee to prevent any single person from experiencing abuse on our platform. However, we are energised and motivated by the challenge and as committed as ever to doing right by our users, particularly as we reach inflection points around the rights of women and as the world deals with the increasingly polarised state of national and supranational politics.

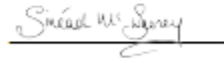
We are always open to learning from the experience and research of others. In particular, we will be attentive to Amnesty's recommendations on how we, and other Internet companies, can and should address the challenge of online abuse and violence against women and others. We will also welcome your insight into how we can meaningfully measure progress. As I explained yesterday, raw data around numbers of reports received or reports actioned is not informative as reporting tools are often used inappropriately or even sinisterly to undermine political viewpoints or to flag disagreement. At Twitter we are working to find a way in which we can offer more transparency at the user level but also at a macro level so we can have real and constructive conversations about the extent of the problem and ensure that we are finding sustainable solutions. Your proposals in this respect will be greatly appreciated.

The assertion that Twitter is consciously unengaged on these issues is flagrantly unrepresentative of the facts, the ethos of this team, and the core mission of Twitter as a company.

Twitter International Company	One Cumberland Place Fenian Street Dublin 2, D02 AX07 Ireland	Registered Number 503351	Directors L O'Brien, R. O'Shea, V. Gadde (US)
-------------------------------	--	-----------------------------	--

While we deeply respect Amnesty International's mission and work around the world, we cannot help but feel disappointed at the tenor of the findings you have shared with us. They do not take account of the substantial company-wide work we have done to protect and promote the voices of women on Twitter, and to offer better structural support to all of our users. We should all aspire to a meaningful conversation in which we can be held to account for our work while learning from your expertise and insights. It is a matter of significant regret to us that this is not the path that Amnesty International has chosen on this occasion, and we hope that our future engagements can be more productive and constructive, as we all seek to solve a complex societal problem.

Yours sincerely,



Sinéad McSweeney
VP, Public Policy and Communications, EMEA

Twitter International Company	One Cumberland Place Fenian Street Dublin 2, D02 AX07 Ireland	Registered Number 503351	Directors L. O'Brien, R. O'Shea, V. Gadde (US)
-------------------------------	--	-----------------------------	---

**AMNESTY INTERNATIONAL
IS A GLOBAL MOVEMENT
FOR HUMAN RIGHTS.
WHEN INJUSTICE HAPPENS
TO ONE PERSON, IT
MATTERS TO US ALL.**

CONTACT US



info@amnesty.org



+44 (0)20 7413 5500

JOIN THE CONVERSATION



www.facebook.com/AmnestyGlobal



[@Amnesty](https://twitter.com/Amnesty)

#TOXICTWITTER

VIOLENCE AND ABUSE AGAINST WOMEN ONLINE

Twitter is a social media platform used by hundreds of millions of people around the world to debate, network and share information with each other. From high-level female politicians to journalists, activists, writers and bloggers, to women who simply want to know what's happening around them - Twitter can be a powerful tool for women to make connections and express themselves. In fact, the company has touted itself as a place where 'every voice has the power to impact the world'. But for many women, Twitter is a platform where violence and abuse against them is allowed to flourish, often with little accountability. As a company, Twitter is failing to respect women's rights online by inadequately investigating and responding to reports of violence and abuse in a transparent manner. The violence and abuse many women experience on the platform has a detrimental effect on their right to express themselves equally, freely and without fear. Instead of Twitter strengthening women's voices, the violence and abuse women experience on the platform means that women are self-censoring what they post, limiting their interactions, or being driven off Twitter completely. At a watershed moment when women around the world are using their collective power to speak out and amplify their voices through social media platforms, Twitter's failure to respect human rights and tackle violence and abuse means that instead of women using their voices 'to impact the world', many women are instead being pushed backwards towards a culture of silence.