


Tracking experiences of online harms and attitudes towards online safety interventions

Findings from a large-scale, nationally representative survey of the British public



Florence Enock
Pica Johansson
Jonathan Bright
Helen Margetts

Content

Authors	3
Executive summary	4
Introduction	5
Key findings	7
> Self-reported exposure to online harms amongst the British public	7
> What should social media platforms do to tackle online harms?	8
> What should the government do to tackle online harms?	10
Discussion and conclusion	13
Acknowledgements	14
Supplementary Information	16
> Methods	16
> Participants	16
> Survey materials and procedure	18
> Supplementary figures	19
> Preferred social media platform and government actions for tackling online harms by demographic groups	19
> Number of social media accounts held and internet use frequency by demographic groups	20

Funding

This work was supported by funding from the Engineering and Physical Sciences Research Council (EPSRC) and the Arts and Humanities Research Council (AHRC).

Authors

Florence Enock is a Postdoctoral Research Associate in Online Safety at The Alan Turing Institute.

<https://www.turing.ac.uk/people/researchers/florence-enock>

Pica Johansson is a Research Associate in Online Safety at The Alan Turing Institute

<https://www.turing.ac.uk/people/researchers/pica-johansson>

Jonathan Bright is Head of Online Safety and Head of AI for Public Services at The Alan Turing Institute.

<https://www.turing.ac.uk/node/1794>

Helen Margetts is Director of the Public Policy Programme at The Alan Turing Institute and Professor of Society and the Internet at the Oxford Internet Institute, University of Oxford.

<https://www.turing.ac.uk/people/researchers/helen-margetts>

Executive summary

In light of current policy conversations around online safety, we sought to understand experiences of online harms and attitudes towards their mitigation amongst the British public.

To do so, we asked a nationally representative sample of over four thousand people the extent to which they had experienced content which they consider to be harmful online (such as hate speech, misinformation, bullying or violence), as well as what they thought social media platforms and the government should do to tackle harmful content online. Our findings show that exposure to online harms amongst the British public is high and demonstrate that people strongly welcome action to tackle such content. These findings come at a time of heightened national attention to a myriad of topics concerning the next phase of internet regulation, and highlight the importance of efforts from researchers, practitioners and policy-makers in working towards a safer online environment.

- **Our results suggest that exposure to online harms amongst the British public is high.** Two thirds (66%) of all adults in the sample reported that they had witnessed harmful content online before, whilst for participants aged 18-34 this was almost 9 in 10 (86%). Participants in the youngest age bracket reported the highest exposure to harm, with 41% of 18-24 year olds indicating that they had been exposed to harmful content many times.

- **Participants across all demographic groups strongly welcomed action from social media platforms to tackle online harms.** Almost 80% of respondents thought that social media platforms should ban or suspend users who create harmful content, and almost 75% thought that platforms should remove harmful content. This was consistent across age, gender, educational background, income and political ideology.
- **The majority of respondents support increased action from the government to tackle online harms.** More than 70% of respondents said that the government should be able to issue large fines for platforms that fail to deal with harmful content online, while 66% thought that legal action should be taken against platforms that fail to deal with harmful content online.

If you have questions about this report or would like more information about The Alan Turing Institute's research, please contact Florence Enock (fenock@turing.ac.uk).

Introduction

Currently in the UK, concerns surrounding people's exposure to harmful content online, such as hate speech, misinformation, harassment, violence and self-harm promotion, are heightening. Consequences of exposure to such content are potentially severe. For example, while online hate speech and abuse can cause harm to the psychological wellbeing of targets and can provoke and justify violent attacks offline, exposure to certain kinds of misinformation may unfairly manipulate voting behaviours or encourage people to put themselves or others at risk of extreme physical harm.

Researchers, policy-makers and practitioners working in the broad arena of online safety are largely in agreement that more must be done to protect people from online harms whilst also protecting freedom of speech. The Online Safety Bill emerged in response to some of these concerns, and was introduced to UK Parliament in 2019 with the aim to regulate companies that offer online services and facilitate user-to-user interactions, such as social media companies and search engines. The legislation aims to ensure that comprehensive measures are in place to systematically tackle harmful online content (though since its introduction, the Bill has sparked debate on where to draw the limits on the right to freely express oneself online, and some policymakers have argued that this piece of legislation might not be the solution to the problems it seeks to address). Under the Bill, Ofcom is provided with regulatory pow-

ers to ensure that online services are taking a proactive approach to managing the risk of online harm and will be able to conduct risk assessments and examine transparency reports, as well having the ability to make information requests to services about their products.¹ The Bill also places special emphasis on protecting children. While the Bill is still going through parliament with the aim of passing before the end of 2023, and the scope of the Bill is also changing, conversations surrounding this new piece of legislation draw attention to the importance of understanding the extent to which the British public are exposed to content which they consider to be harmful, and how they would like to see such content dealt with.

It is difficult to estimate the extent to which people are typically exposed to harmful content online because platforms at this time do not typically make this information available. Studying online harms is also made difficult by the differing definitions of what constitutes harm. Some studies seek to quantify exposure to lists of specific harms, while other researchers instead choose to account for the subjective nature of what respondents themselves consider harmful. One study by Ofcom suggested that 62% of UK-based internet users had encountered potentially harmful content or behaviour online in the past four weeks at the time of asking, with almost half of these respondents indicating that their most recent encounter with such content was on social media.²

¹ Ofcom. (2022). *Online Safety Bill: Ofcom's roadmap to regulation*. https://www.ofcom.org.uk/_data/assets/pdf_file/0016/240442/online-safety-roadmap.pdf

² Ofcom. (2022). *Online Nation: 2022 report*. https://www.ofcom.org.uk/_data/assets/pdf_file/0023/238361/online-nation-2022-report.pdf

Further, while many policy debates contemplate important issues in online safety interventions such as balancing free speech with content moderation, little is known about general public opinion on these matters.³ One recent

study suggests that people believe social media platforms should do more to tackle harms online,⁴ but which specific actions people are in favour of remain unclear.

³ For recent findings suggesting that US participants choose removing harmful misinformation over free speech expression, see: Kozyreva, A., Herzog, S. M., Lewandowsky, S., Hertwig, R., Lorenz-Spreen, P., Leiser, M., & Reifler, J. (2023). Resolving content moderation dilemmas between free speech and harmful misinformation. *Proceedings of the National Academy of Sciences*, 120(7), e2210666120. <https://www.pnas.org/doi/10.1073/pnas.2210666120>

⁴ Ipsos. (2022). *Quantitative research about the impact of harmful, but legal, content online - Wave 2 UK general public*. <https://www.ipsos.com/en-uk/new-poll-finds-seven-ten-adults-want-social-media-firms-do-more-tackle-harmful-content>

Key findings

Self-reported exposure to online harms amongst the British public

To understand general levels of exposure to online harms amongst people living in Great Britain, we asked a large, nationally representative sample (N = 4012) the extent to which they had experienced content which they consider to be harmful online. We note here that the harms we asked about are potentially more wide ranging than those covered by the Online Safety Bill referenced above.

Participants were initially briefed: *Some people are concerned that social media platforms such as Twitter and Facebook are exposing their users to harmful content, including hate speech, false information, fake images, bullying, violence and self-harm promotion.*⁵

Participants were then asked to indicate both the extent to which they had witnessed content which they considered to be harmful on social media platforms in the past, and the extent to which they had directly received such content. 'Witnessing' was defined as observing harmful content not intended directly for the participant, for example in a social media post that did not explicitly name them or their username. 'Directly receiving' was defined as content which was directly intended for the participant, such as content sent to them by direct message (DM) or explicitly naming them or their username.

For each type of harm exposure, participants could choose one from the following response options: *Many times; Occasionally (from time to time); Very rarely (only once or twice); Never; Not sure; Prefer not to say.*

Self-reported exposure to harmful content online was high. Across all participants, 66% reported that they had witnessed harmful content online before, while 23% of those had witnessed such content many times. 29% of all participants reported that they had directly received harmful content before, with 1 in 20 (5%) of those stating that they had directly received harmful content many times.

Younger participants reported the greatest exposure to online harms, with 85% of 18-24 year olds and 87% of 25-34 year olds reporting that they had previously witnessed harmful content online. Of these, 41% of 18-24 year olds and 39% of 25-34 year olds reported that they had witnessed such content many times. Additionally, 45% of 18-24 year olds and 42% of 25-34 year olds reported having directly received harmful content before, while of the 18-24 year olds, almost 1 in 10 (9%), said they had directly received such content many times. Self-reported exposure to harmful content decreased with age though was still notable, with 47% of 65-74 year olds and 27% of over 75s having witnessed harmful content before, and 16% of 65-74 year olds and 11% of over 75s having directly received such content.⁶

⁵ These examples are not informed by specific existing or proposed legislation, but instead reflect types of content that are commonly understood to pose a risk of harm.

⁶ It is possible that the amount of time spent online partially explains age differences in self-reported exposure to online harms, with participants in younger age brackets spending larger amounts of time online than those in the older age brackets. Similarly, participants in younger age brackets typically use a greater number of social media accounts than those in older age brackets. For internet use frequency and number of social media accounts held broken down by demographics, see Supplementary Information.

Members of ethnic minority groups reported greater exposure to online harms than White participants. 76% of Black, African, Caribbean or Black British participants and 70% of Asian or Asian British participants reported having witnessed harmful content online compared to 65% of White participants, and 39% of Black,

African, Caribbean or Black British participants and 32% of Asian or Asian British participants reported having directly received harmful content compared to 28% of White participants. Figure 1, below, shows levels of self-reported exposure to online harms across categories for sex, age, ethnicity and education level.

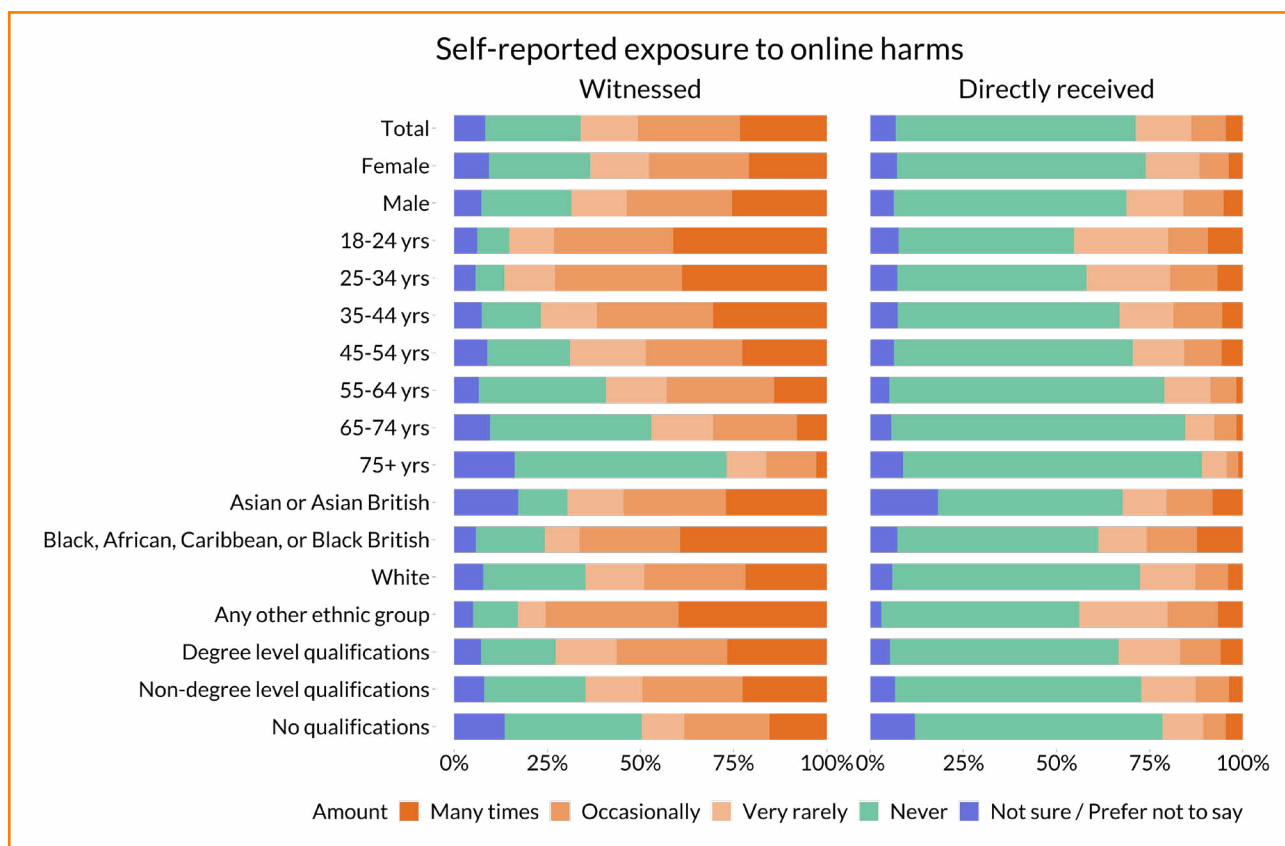


Figure 1: Self-reported exposure to online harms amongst the British public broken down by categories for sex, age, ethnicity and education level. The first panel shows the extent to which people had witnessed harmful content. The second panel shows the extent to which people had directly received harmful content.

What should social media platforms do to tackle online harms?

To understand attitudes towards online safety interventions implemented by social media platforms, we asked participants what, if anything, they thought social media platforms should do to tackle content which poses a risk of harm. Participants were given a choice of

several actions that social media platforms might take, and could select all that applied.

The response choices were:

1. *Remove harmful content*
2. *Ban or suspend users who create harmful content*
3. *Add warning labels to harmful content*

4. *Make it easier for people to report harmful content to the platform*
5. *Make harmful content difficult to find through search terms*
6. *Stop paid promotion of harmful content*
7. *Give users greater control over what they see on social media platforms, for example by allowing them to choose how their social media feeds are organised*
8. *Something else (free text)*
9. *Nothing - social media platforms should not do anything to tackle harmful content*
10. *None of the above*
11. *Don't know*

Choices 9-11 were exclusive such that participants could not choose these along with any others.

On the whole, participants were in favour of social media platforms taking action to tackle the harmful content they host, with less than 1% choosing 'Nothing - platforms should not do anything to tackle harmful content'. The most commonly chosen action for social media platforms was banning or suspending users who create harmful content, with 79% of participants in favour. 73% of participants thought social media platforms should remove harmful content, 69% thought that platforms should make it easier for people to report harmful content and 68% favoured platforms stopping paid promotion of harmful content. Figure 2, below, shows overall percentages of participants favouring each action.

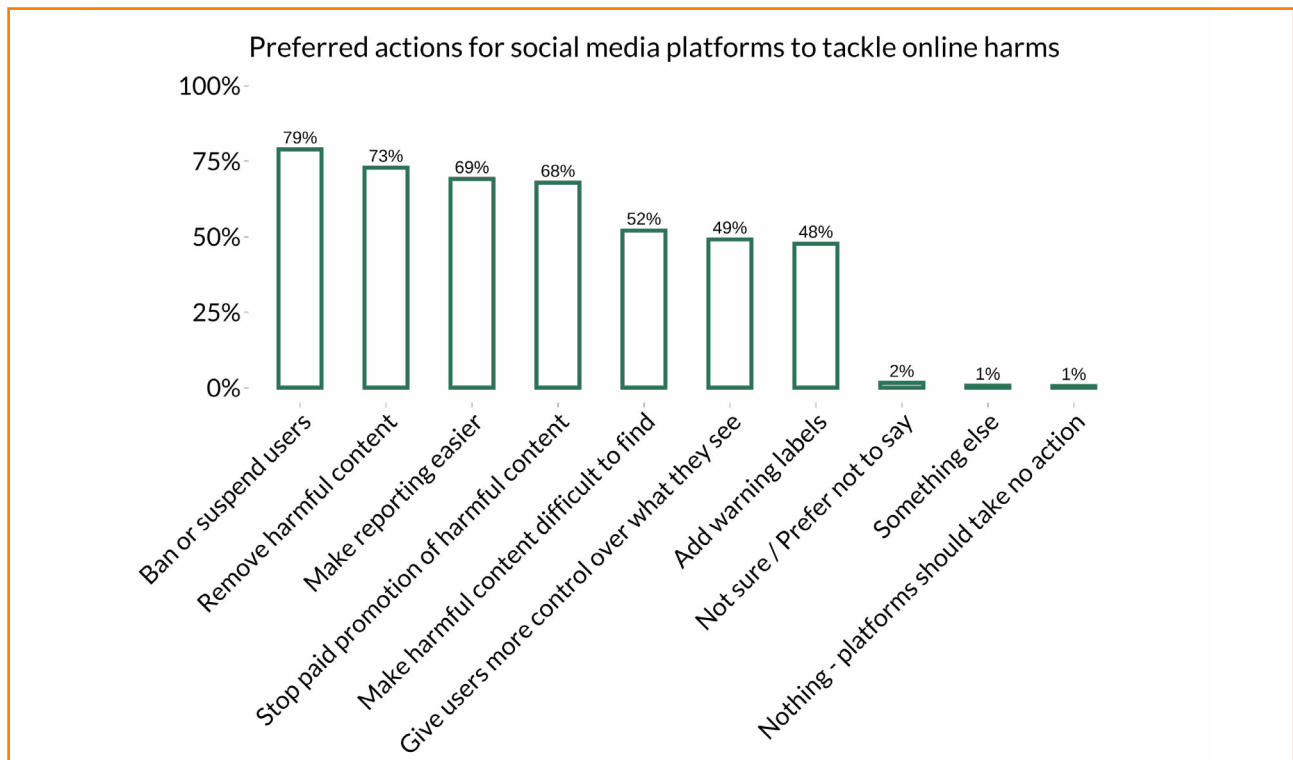


Figure 2: Preferred actions by social media platforms for tackling online harms. The bars show overall percentages of participants in favour of each action in order of frequency of choice. People most commonly indicated that they thought platforms should ban or suspend users and remove harmful content. The full descriptions of actions that were presented to participants are outlined in bullets 1-11, above.

Attitudes towards online safety measures implemented by social media platforms were fairly consistent across demographic groups. Banning or suspending users and removing harmful content were the two most popular choices for both male and female participants, however a higher percentage of female participants were in favour of social media actions on the whole, with 82% in favour of banning or suspending users compared to 76% of male participants, and 77% of females in favour of removing harmful content compared to 69% of males.

The overall pattern of responses was also similar across age groups, with banning or suspending users and removing harmful content again the two most popular choices across all age groups. However, older participants were in general more in favour of social media platforms taking action than younger participants - 69% of 18-24 year olds and 75% of 25-34 year olds thought that social media platforms should ban or suspend users, while 82% of 55-64 year olds and 88% of 65-74 year olds indicated the same. Similarly, 61% of 18-24 year olds and 66% of 25-34 year olds were in favour of social media platforms removing harmful content, while this rose to 80% of 55-64 year olds and 79% of 65-74 year olds.

Preferences for which actions social media actions should take to tackle online harms were consistent across political party affiliation, with banning or suspending users and removing harmful content the two most popular choices across Conservative, Labour and Liberal Democrat voters.⁷ 83% of Conservative voters

and 81% of Labour voters indicated that they thought social media platforms should ban or suspend users that create harmful content, while 78% of Conservative voters and 74% of Labour voters indicated that they thought social media platforms should remove harmful content. Figure S1 in Supplementary Information shows preferred actions for social media platforms by demographic categories for sex, age, ethnicity, education level and political vote choice.

What should the government do to tackle online harms?

To understand attitudes towards online safety interventions implemented by the government, we asked participants what, if anything, they thought the government should do to tackle content that poses a risk of harm on social media platforms. Participants were given a choice of several actions that the government might take, and could select all that applied.

The response choices were:

1. *Implement large fines for platforms that fail to deal with harmful content*
2. *Force platforms to systematically report how much harmful content they host and how they are attempting to combat this harmful content*
3. *Launch legal proceedings against platforms that fail to deal with harmful content*
4. *Publicly name and shame platforms that fail to deal with harmful content*
5. *Prevent platforms that fail to deal with*

⁷ According to how they voted in the 2019 general election. Only voters of these three parties were included in this comparison because of small sample sizes for voters of other parties.

harmful content from operating

6. *Force platforms to introduce strict age verification procedures,*
7. *Something else (free text)*
8. *Nothing - there should be no consequences for platforms that fail to deal with harmful content*
9. *None of the above*
10. *Don't know*

Choices 8-10 were exclusive such that participants could not choose these along with any others.

Participants were again strongly in favour of action against online harms, with less than 1% choosing '*Nothing - there should be no consequences for platforms that fail to deal with harmful content*'. The most commonly chosen action for government was implementing large fines for platforms that fail to deal with harmful content, with 72% of all respondents in favour. Additionally, 66% of respondents thought that the government should be able to launch legal proceedings against platforms that fail to deal with harmful content online. Figure 3, below, shows overall percentages of participants favouring each government action.

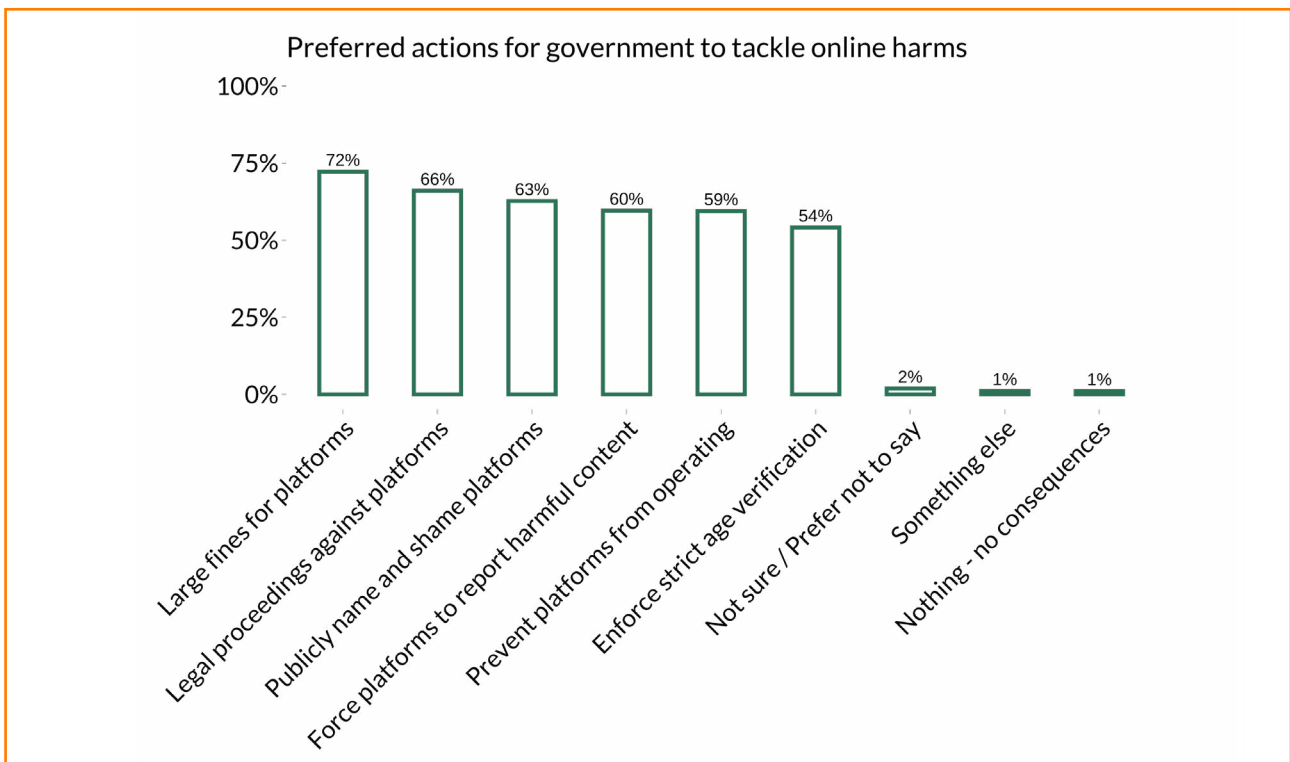


Figure 3: Preferred government actions for tackling online harms. The bars show overall percentages of participants in favour of each action in order of frequency of choice. People were most commonly in favour of large fines for and legal proceedings against platforms that fail to deal with harmful content online. The full descriptions of actions that were presented to participants are outlined in bullets 1-10, above.

Attitudes towards online safety measures implemented by social media platforms were fairly consistent across demographic groups.

Banning or suspending users and removing harmful content were the two most popular choices for both male and female participants,

however a higher percentage of female participants were in favour of social media actions on the whole, with 82% in favour of banning or suspending users compared to 76% of male participants, and 77% of females in favour of removing harmful content compared to 69% of males.

The overall pattern of responses was also similar across age groups, with banning or suspending users and removing harmful content again the two most popular choices across all age groups. However, older participants were in general more in favour of social media platforms taking action than younger participants - 69% of 18-24 year olds and 75% of 25-34 year olds thought that social media platforms should ban or suspend users, while 82% of 55-64 year olds and 88% of 65-74 year olds indicated the same. Similarly, 61% of 18-24 year olds and 66% of 25-34 year olds were in favour of social media platforms

removing harmful content, while this rose to 80% of 55-64 year olds and 79% of 65-74 year olds.

Preferences for which actions social media actions should take to tackle online harms were consistent across political party affiliation, with banning or suspending users and removing harmful content the two most popular choices across Conservative, Labour and Liberal Democrat voters. 83% of Conservative voters and 81% of Labour voters indicated that they thought social media platforms should ban or suspend users that create harmful content, while 78% of Conservative voters and 74% of Labour voters indicated that they thought social media platforms should remove harmful content. Figure S1 in Supplementary Information shows preferred actions for social media platforms by demographic categories for sex, age, ethnicity, education level and political vote choice.

Discussion and conclusion

To understand experiences of online harms and attitudes towards online safety interventions amongst the British public, we asked a nationally representative sample of over four thousand people the extent to which they had witnessed and received harmful content online (such as hate speech, false information, fake images, bullying, violence and self-harm promotion), and which actions they thought social media platforms and the government should take to tackle online harms.

Self-reported exposure to harmful content online was strikingly high. Two thirds of all participants had witnessed harmful content online before, and almost one in four had witnessed harmful content many times. More than one in four participants reported that they had directly received harmful content before (for example in a direct message or in a post explicitly naming them or their username), with one in twenty stating that they had directly received harmful content many times.

Participants in the youngest age brackets reported the greatest exposure to online harm. Almost nine in ten 18-34 year olds reported that they had previously witnessed harmful content online, and well over a third reported that they had witnessed such content many times. Almost half of the 18-24 year olds in the sample had directly received harmful content before, with almost one in ten having directly received such content many times. It is possible that some of the reported age differences are partly explained by the greater

amount of time that younger adults spend online and the greater number of social media accounts they hold compared to older adults.⁸ However, it is also possible that younger people's experiences online differ qualitatively as well as quantitatively and additional work will benefit from understanding the key drivers of demographic differences in exposure to online harms more fully.

Results also suggested that members of ethnic minority groups may be exposed to online harms to a greater extent than White participants. While this result corroborates findings from Ofcom's Online Nation report (p.67), conclusions are limited by relatively small sample sizes representing each ethnic group. Our sample was designed to be nationally representative of the Great British population, but with a target sample of four thousand, numbers in some demographic categories are still too small to confidently draw inferences about differences in attitudes and experiences. Important work elsewhere focuses on understanding how ethnic minority groups are disproportionately affected by certain kinds of online harms in greater detail.⁹

In addition to reporting high levels of exposure to online harms, participants were overwhelmingly in support of action from both government and social media platforms to tackle online harms. Almost four in five participants were in favour of social media platforms banning or suspending users who create harmful content and more than seven

⁸ Ofcom. (2022). *Online Nation: 2022 report*. (p.10).

⁹ E.g., Ofcom. (2023). *Qualitative research into the impact of online hate*. https://www.ofcom.org.uk/_data/assets/pdf_file/0020/252740/qual-research-impact-of-online-hate.pdf

in ten participants believed that social media platforms should remove harmful content. A similar proportion of participants believed that the government should be able to implement large fines for social media platforms that fail to deal with harmful content, and two thirds thought that the government should be able to launch legal proceedings against platforms that fail to deal with harmful content. On the whole, people were even more in favour of social media actions than of government actions, though support for both was high.

Our findings suggest exposure to online harms amongst the British public is widespread. However, we note that we asked about experience with online harms defined broadly, inclusive of content such as hate speech, false information, fake images, bullying, violence and self-harm promotion. Because of this, we cannot separate out which specific types of online harms people are most exposed to, or understand whether particular demographic groups are more at risk of certain harms. For example, work elsewhere suggests that while men and women report similar levels of exposure to online harassment, women targeted in online harassment are more than twice as likely as men to say most recent incident was very or extremely upsetting¹⁰ and are more likely than targeted men to become more cautious in expressing opinions online.¹¹ Additional work will benefit from a deeper understanding of the specific kinds of online harms that different groups of people in society

are most at risk of experiencing, along with the wider impact of these harms on individuals.

We also note that our measures of people's exposure to online harms are subjective. People were asked the extent to which they had witnessed or received content *which they considered to be harmful in the past*. It is possible that there are individual or group-level differences in people's perceptions of what is harmful. Nonetheless, people's subjective experiences of harm are likely to be important indicators of how they are impacted. Similarly, we do not put concrete numbers on the extent to which participants experienced online harms (response options were: 'Many times', 'Occasionally', 'Rarely' and 'Never'). Endeavours to track the prevalence of harmful content more concretely are important contributions for building a full picture of the landscape of online harms.¹²

Taken together, our findings show the high prevalence of exposure to online harms amongst the British public and demonstrate that people strongly welcome action both from government and social media platforms to tackle such content. At a time when conversations around the Online Safety Bill are heightening, our data highlights the crucial importance of efforts from researchers, practitioners and policy-makers in working towards a safer online environment for all members of society.

¹⁰ Vogels, E. A. (2021). *The State of Online Harassment*. Pew Research Center. <https://www.pewresearch.org/internet/2021/01/13/the-state-of-online-harassment/>

¹¹ Nadim, M., & Fladmoe, A. (2021). Silencing women? Gender and online harassment. *Social Science Computer Review*, 39(2), 245-258.

¹² E.g., Vidgen, B. et al. (2022). *Tracking abuse on Twitter against football players in the 2021 – 22 Premier League Season*. Ofcom. <https://www.ofcom.org.uk/research-and-data/online-research/twitter-abuse-against-premier-league-players>

Acknowledgements

We are grateful to Aidan Peppin, Octavia Reeve and Andrew Strait at The Ada Lovelace Institute and Professor Patrick Sturgis at The London School of Economics and Political Science (LSE) for their collaboration on the wider survey from which these findings were drawn. We thank Dr. Cosmina Dorobantu at

The Alan Turing Institute for valuable feedback on question design and data interpretation for the findings discussed in this report. We also thank Joudy Bourghli at The Alan Turing Institute for her help designing the presentation of this report.

¹³ <https://www.kantar.com/uki/expertise/policy-society/public-evidence/public-voice>

Supplementary Information

Methods

Participants

Participants were drawn through Kantar's public voice panel¹³ and data was collected in November and December 2022. A total of 4012 participants who completed the survey passed standard checks for data quality and were included in the final sample. The sample was designed to be nationally representative of the population of Great Britain across demographic variables of age, sex and ethnicity.

Of the 4012 participants, 3759 responded by web (94%) and 253 by telephone (6%). Participants in the sample were aged between 18 and 94, with a mean age of 48 (SD = 17.2). A total of 1912 (48%) participants were male and

2096 (52%) were female (note that this reflects sex at birth). 3546 (88%) of respondents were White, 261 (7%) were Asian or Asian British, 90 (2%) were Black, African, Caribbean, or Black British, and 103 (3%) were mixed, multiple or other ethnicities. While participants indicated more specific ethnic identities at the time of recruitment to the Public Voice panel, we combine them into these broader categories in our analyses for more meaningful sample sizes. 1562 (39%) had degree-level qualifications, 2156 (54%) had non-degree level qualifications and 284 (7%) had no qualifications. Data was weighted to match the demographic profile of the population. Table S1, below, shows sample demographics including both unweighted and weighted totals.

Table S1: Weighted and unweighted sample sizes for each demographic category

	Demographic group	Unweighted N	Unweighted %	Weighted N	Weighted %
Sex	Male	1912	47.7%	1964.4	48.9%
	Female	2096	52.2%	2036.6	50.7%
	NA	4	0.1%	12.4	0.3%
Ethnicity (reduced)	White	3546	88.4%	3487.6	86.9%
	Asian or Asian British	261	6.5%	295.8	7.4%
	Black, African, Caribbean, or Black British	90	2.2%	103.0	2.6%
	Any other ethnic group	103	2.6%	115.5	2.9%
	NA	12	0.3%	11.6	0.3%
Age group	18-24 yrs	343	8.5%	419.0	10.4%
	25-34 yrs	709	17.7%	682.1	17.0%
	35-44 yrs	741	18.5%	653.5	16.3%
	45-54 yrs	692	17.2%	665.5	16.6%
	55-64 yrs	696	17.3%	645.4	16.1%
	65-74 yrs	513	12.8%	517.2	12.9%
	75+ yrs	318	7.9%	430.7	10.7%
Education level	Degree level qualifications	1562	38.9%	1407.1	35.1%
	Non-degree level qualifications	2156	53.7%	2146.6	53.5%
	No qualifications	284	7.1%	446.5	11.1%
	NA	10	0.2%	13.2	0.3%
Political vote in 2019 GE	Labour	977	24.4%	893.6	22.3%
	Conservatives	1195	29.8%	1211.6	30.2%
	Liberal Democrats	372	9.3%	326.3	8.1%
	SNP (S)	111	2.8%	116.7	2.9%
	Plaid Cymru (W)	21	0.5%	25.8	0.6%
	Green Party	118	2.9%	108.7	2.7%
	The Brexit Party	34	0.8%	40.7	1.0%
	Other party or individual	20	0.5%	17.2	0.4%
	Did not vote	871	21.7%	977.7	24.4%
	NA	293	7.3%	295.2	7.4%

Survey materials and procedure

These results are part of a larger survey of public attitudes to AI and data-driven technologies conducted by The Alan Turing Institute and The Ada Lovelace Institute. Further findings from the survey will be published in a forthcoming report.

After completing the main body of the full survey, participants answered the three questions relating to online harms described in this report. The entire survey was designed to take each respondent approximately 25 minutes to complete and participants were compensated at a standard rate for their time.

Demographic information such as age, sex, ethnicity and education level was taken at the time of participant recruitment to the public

voice panel. Therefore, some variables that are not fixed, such as education level, may have changed between participant sign-up to the panel and completion of this particular survey. Political affiliation was determined at the time of sign-up by vote in the 2019 general election.

To understand possible drivers of differences in self-reported levels of harm exposure between demographic groups such as age groups, we provide supplementary analyses (below) examining time spent online and number of social media accounts held across the demographic groups of interest. These additional variables were also taken at the time of participants sign-up to the panel, not as part of this particular survey and so are also only estimates of participants' current internet and social media use.

Supplementary figures

Preferred social media platform and government actions for tackling online harms by demographic groups

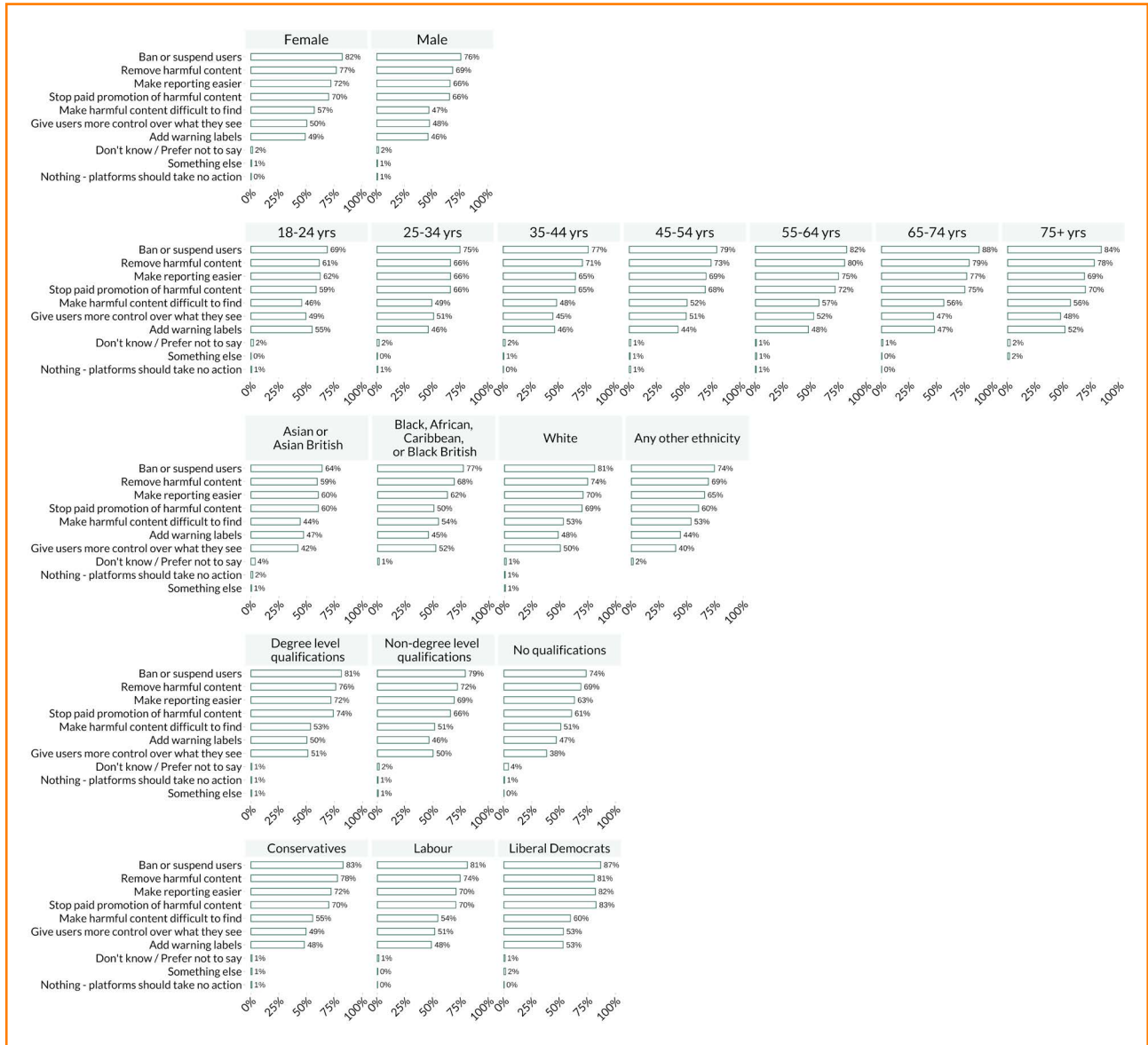


Figure S1 (above): Preferred social media platform actions for tackling online harms by demographic group. The bars show percentages of participants in favour of each action within each category for sex, age, ethnicity and education level.

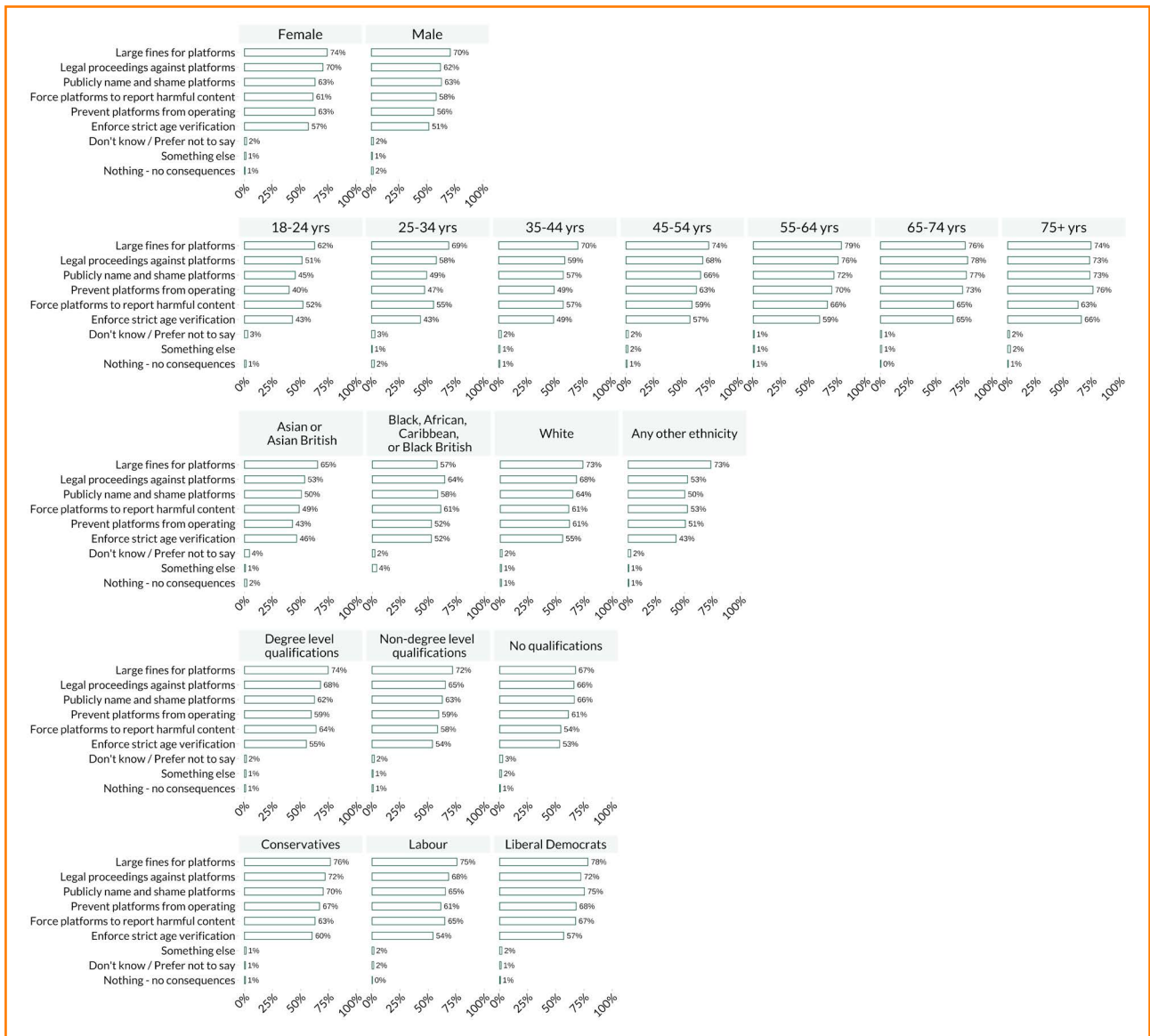
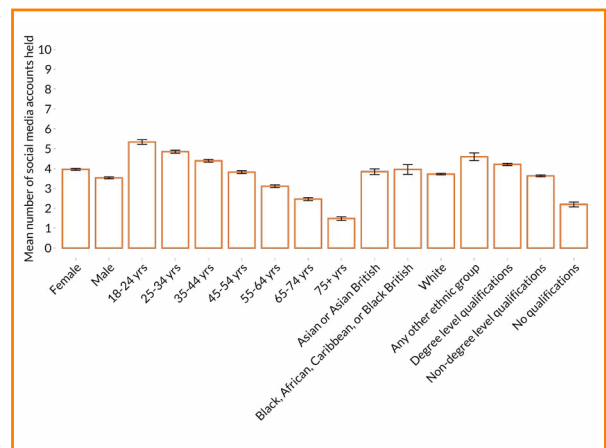


Figure S2 (above): Preferred government actions for tackling online harms by demographic group. The bars show percentages of participants in favour of each action within each category for sex, age, ethnicity and education level.

Number of social media accounts held and internet use frequency by demographic groups

Figure S3: Mean number of social media accounts held by participants within each demographic category for sex, age, ethnicity and education. Error bars represent standard errors of the mean.



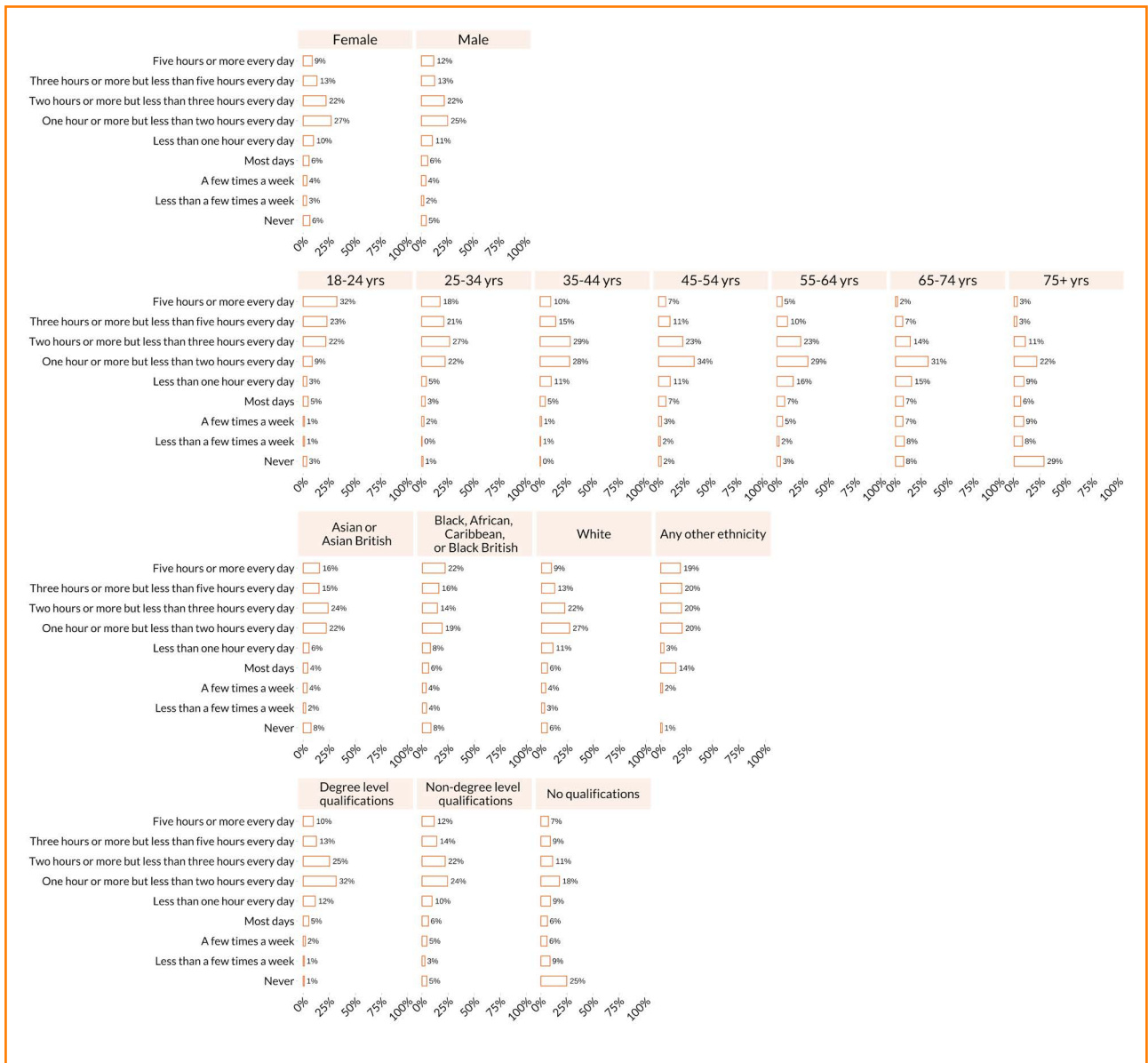


Figure S4 : Internet use frequency for this sample by demographic categories for sex, age, ethnicity and education. The bars show percentages of participants choosing each level of use within each demographic.



**turing.ac.uk | @turinginst
onlinesafety@turing.ac.uk**