

Consultation response form

Please complete this form in full and return to OS-Section54@ofcom.org.uk

Consultation title	Consultation: A safer life online for women and girls: practical guidance for tech companies
Full name	Alexandra Krendel ; Veronika Koller ; Mark McGlashan ; Jessica Aiston
Contact phone number	
Representing (delete as appropriate)	Organisation
Organisation name	University of Southampton; Lancaster University; University of Liverpool; Queen Mary University of London
Email address	a.krendel@soton.ac.uk ; v.koller@lancaster.ac.uk ; m.mcglashan@liverpool.ac.uk ; j.aiston@qmul.ac.uk

Confidentiality

We ask for your contact details along with your response so that we can engage with you on this consultation. For further information about how Ofcom handles your personal information and your corresponding rights, see [Ofcom's General Privacy Statement](#).

Your details: We will keep your contact number and email address confidential. Is there anything else you want to keep confidential? Delete as appropriate.	We are a group of four academics in the field of linguistics who have been analysing the language of the manosphere since 2018. We have published widely on the topic (see an overview of our work here) and are keen to support Ofcom's work on addressing this issue. There is nothing we want to keep confidential in this response.
Your response: Please indicate how much of your response you want to keep confidential. Delete as appropriate.	None

For confidential responses, can Ofcom publish a reference to the contents of your response?

N/A

Your response

Question	Your response
Question 1: Do you have any comments on our proposed approach to 'content and activity' which 'disproportionately affects women and girls'?	<p>Confidential? – N</p> <ol style="list-style-type: none">1. We welcome the focus on online misogyny, which is especially important given the prevalence of manosphere influencers and the filtering of online misogynistic content from niche online forums into mainstream social media, as well as classrooms and everyday settings.2. The consultation document does not use the term 'manosphere,' although there are references to incels and misogynistic influencers. The manosphere, a loose network of online antifeminist communities, would certainly fall under the document's definition of online misogyny as 'the circulation of content that actively encourages or reinforces misogynistic ideas or behaviours, including content that incites hatred, abuse or threats toward women and girls. It also includes sexual or explicit content that normalises or encourages harmful sexual behaviour.' We hope that Ofcom includes the whole manosphere as part of their approach.3. Conversely, we express concern that incels have been singled out. Incels have the most links to offline violence against women and girls, though we argue that in seeking to address 'legal but harmful' content or content which 'normalises or encourages harmful sexual behaviour' then the rest of the manosphere must also be taken into account. This includes: pick-up artists, who encourage sexual harassment and share harmful 'seduction' strategies such as telling men to

Question	Your response
	<p>‘push through’ women’s non-consent to sexual activity; Men Going Their Own Way, who argue that heterosexual relationships are toxic and tell men to avoid women in many areas of their life; and men’s rights advocates, who downplay the harms of sexism and misogyny.</p> <p>4. Based on the proposed actions, it seems that stopping the spread of the manosphere is not fully addressed. In the rest of our response, we provide comments on the proposed actions and suggest areas where these actions could be more tailored to the manosphere specifically.</p>
<p>Question 2: Do you have any comments on the nine proposed actions? Please provide evidence to support your answer.</p>	<p>Confidential? – N</p> <p>5. The actions are split between foundational steps and good practice steps, which creates two tiers of action. This corresponds to the priority areas (harm to children and illegal content) and legal but harmful content, respectively. We argue that the good practice steps are key to preventing online harms such as those brought about by engaging with manosphere content. We advise that the guidance should be firmer in urging companies to implement good practice steps (e.g., “<u>in some cases</u>, providers <u>may</u> also <u>seek</u> to limit the circulation of such content” [2.66] -> “providers <u>should</u> make <u>every effort</u> to limit the circulation of such content”).</p> <p>Our detailed response to each of the individual actions are summarised in the table below.</p>

	Description	Action	Example of good practice steps	Our comments on the proposed actions
Taking responsibility	How providers can make decisions and conduct assessments that account for women and girls' experiences	Action 1: Ensure accountability processes address online women and girls' online safety	This could include having policies that are designed to tackle forms of online gender-based harms	<ul style="list-style-type: none"> ● Primary research on and by platforms is needed. This should involve study of content on the platform informed by prior research on (online) misogyny as well as through user and stakeholder surveys (incl. charities). ● Primary research should be carried out alongside independent, external assessors to identify, capture, and verify data on potential harms. ● Data on potential harms should include clear data points and behavioural indicators (e.g. trigger words/phrases, hashtags, visual tags, user profile and bio contents) ● Risk assessments should monitor the frequency of occurrences of indicators/data points related to potential harms. For example, does a particular hashtag related to the manosphere become more frequent across the platform and does it proliferate (i.e. become used by more users)? ● Risk assessments should report on frequency of occurrence of risk indicators, as well as periodic comparison to show quarterly/monthly variations in usage. Ofcom should then be made aware of any significant increases in use. ● Involve work with expert stakeholders (academics, charities, etc.) who understand the kinds of features used in misogynistic content to test systems
		Action 2: Conduct risk assessments that capture harms to women and girls	This could include conducting user surveys to better understand the experiences of different groups	
		Action 3: Be transparent about women and girls' online safety	This could include sharing information about the prevalence of different forms of online gender-based harms	

Preventing harm	How providers can prevent harm through the design of their services	Action 4: Conduct abusability evaluations and product testing	This could include using red teaming for abusability testing	
		Action 5: Set safer defaults	This could include setting strong and customisable defaults around user interaction and privacy	<ul style="list-style-type: none"> • Proactive blocklisting: Provide proactive measures for blocking and filtering. Providers typically allow users to limit posts from individual accounts. Could a prepopulated (and audited) blocklist of accounts and content associated with misogyny be made available for users to ‘switch on’? Our research on the manosphere, for example, has made possible the identification of manosphere terms that could be used to identify such accounts.
		Action 6: Reduce the circulation of online gender-based harms	This could include continuously improving automated content moderation	<ul style="list-style-type: none"> • Taking down manosphere content: While taking down manosphere content is ideal, we are aware that not all manosphere content may meet the criteria for removal, and so we welcome the action “for legal content, in some cases, providers may also seek to limit the circulation of such content through persuasion, removal and reduction” (p.30), as this is important for avoiding harm to children or addressing violations of a site’s terms of service (TOS). However, it should be noted that services do not necessarily follow their ToS at present, so firmer guidance may be required. • We note that some of the recommendations for Action 2 could be used to address these harms: “use external assessors for monitoring the threat landscape”, “de-monetise sites that promote online-gender based harm”, and “send high risk and highly contextual user reports of gender-based harms for

				<p>review by specifically trained moderators”. We particularly welcome bringing in experts (e.g. academic experts, relevant NGOs, those with lived experience) to help address this harm. This also ties in with a recommendation for Action 1 which outlines consulting with subject matter experts on gender-based harms.</p> <ul style="list-style-type: none"> ● Nudges: active feedback to users that might be in the process of posting misogynistic content towards information that might educate on the use of problematic terms or content. This should be informed by work done with regards to ACTIONS 1-3
Supporting women and girls	What providers can do to support women and girls when harms happen on services.	Action 7: Give users better control over their experiences	This could include allowing users to signal what content they do not want to see, and what content they want to see more of	<ul style="list-style-type: none"> ● See response to ACTION 5
		Action 8: Enable users who experience online gender-based harm to make reports	This could include allowing users to track and manage their reports	<ul style="list-style-type: none"> ● This should also feed into primary research and theme of ‘Taking responsibility’ ● User reports should enable specific claims of VAWG; gender-based harassment, violence, etc.; and misogyny ● Evaluation of contents of texts flagged by users should form part of testing, evaluation, and audit of platforms’ support for women and girls

		Action 9: Take appropriate action when online gender-based harm occurs	This could include taking action against users who continuously violate a service's Terms of Service	
--	--	---	--	--

<https://doi.org/10.5258/SOTON/PP0126>

Question	Your response
<p>Question 3: Do you have any comments about the effectiveness, applicability or risks of the good practice steps or associated case studies we have highlighted in Chapter 3, 4 and 5? Are there any additional examples of good practices we should consider? Please provide evidence to support your comment.</p>	<p>Confidential? – N</p> <ol style="list-style-type: none"> <li data-bbox="730 477 1374 1249">6. Action 1 (“Ensure accountability processes address online women and girls’ online safety”) is formulated in a way that does not make clear who is accountable and who they are accountable to. Our recommendation is to involve subject matter experts such as groups dedicated to preventing violence against women and girls. They should give guidance who should hold tech companies accountable. It is also worth clarifying to what degree tech companies will be held legally accountable for the distribution and consumption of content that leads to misogynistic activity. It is positive to see case study 3 for Action 1 discuss external oversight of content moderation, as this is certainly relevant to the manosphere as it manifests on the content aggregation site Reddit. Ideally, such oversight should be mentioned in the foundational section. <li data-bbox="730 1317 1374 1921">7. We also note that case study 11 obliquely mentions the manosphere on the foundational level as “[a] growing community of misogynistic influencers (sometimes referred to as ‘misogyny influencers’) can have considerable influence over the propagation of misogynistic content.” The case study further says that “recommender systems reward influencers creating misogynistic content with greater reach” and suggests content recommendation algorithms should be trained to be gender-sensitive. We strongly advise that rather than changing “gender-sensitive recommender system algorithms”, legal but harmful content should be removed in the first place.

Question	Your response
	<p>8. However, if the recommendation is to audit and evaluate recommender algorithms based on “a new training dataset put together by a diverse group that includes humans with a high level of sensitivity and training on gender-based harms”, we agree that human annotators need sufficient time and training in order to evaluate this content and that the full context must be considered. Research shows¹ that more implicit hateful speech is usually not removed from sites like Facebook despite contravening the site’s content policies. Some forms of hateful speech may require sophisticated language knowledge (e.g., decoding puns or emojis) or specific socio-cultural, political or historical knowledge in order for the user or moderator to work out the meaning. These kinds of hateful speech may be even more difficult to recognise in a context where social media moderation is outsourced to Asia and Africa and when workers are not given sufficient time to evaluate the content.² Content moderation is also known to cause great psychological distress and workers are not always given the mental health and wellbeing support that they need.³</p> <p>9. Automated moderation systems (such as systems which automatically flag for “toxicity”) can be a good starting point and minimise some of the burden on human moderators. However, they are better at identifying profanities and explicit aggression than misogynistic ideologies.⁴ Members of the manosphere also employ novel words and acronyms which convey misogynistic ideas (e.g., <i>foids</i> meaning female humanoids)</p>

¹ <https://firstmonday.org/ojs/index.php/fm/article/view/10288/8327>

² <https://www.theguardian.com/commentisfree/2025/feb/12/moderator-facebook-real-cost-outsourcing-digital-labour>

³ <https://www.theguardian.com/global-development/2023/sep/11/i-log-into-a-torture-chamber-each-day-strain-of-moderating-social-media-india>

⁴ <https://journals.sagepub.com/doi/10.1177/13548565221111075>

Question	Your response
	<p>and such language is not usually recognised by automated systems.</p> <p>10. Therefore, we agree with Ofcom that human annotators need sufficient time and training. To this, we would add that human annotators need better and more humane working conditions, including access to appropriate psychological support. Automated systems can help to take some of the burden off human moderators, although it cannot replace humans entirely. We would also specify that when they are used, algorithms should be trained on manosphere datasets so that manosphere-specific terminology can be learned.</p> <p>11. Case study 16 states that “[s]ignposting to supportive information that is clear and accessible can increase users’ awareness of the user control tools available to them and encourage users to consider their safety online”. However, such signposting is limited to users reporting harmful content, rather than providing supportive information to <i>all</i> users to encourage them to report such content.</p>
<p>Question 4: Do you have any feedback on our approach to encouraging providers to follow this guidance, including our proposal to publishing an assessment of how providers are addressing women and girls’ safety? Do you have any examples or suggestions of other ways we could encourage providers to take up the ‘good practice’ recommendations?</p>	<p>Confidential? – N</p> <p>12. As stated above, the ‘good practice’ recommendations are key to reducing the harms associated with the spread of the manosphere. The guidance on this should be firmer., as suggested in our response to Question 2.</p>
<p>Question 5: Do you have any comments on our impact assessment, rights assessment, or equality impact</p>	<p>Confidential? – N</p> <p>N/A</p>

Question	Your response
assessment? Please provide any information or evidence in support of your views.	
Question 6: Do you agree that our draft Guidance is likely to have positive effects on opportunities to use Welsh and treating Welsh no less favourably than English? If you disagree, please explain why, including how you consider the draft Guidance could be revised to have positive effects or more positive effects, or no adverse effects or fewer adverse effects on opportunities to use Welsh and treating Welsh no less favourably than English.	Confidential? – N N/A

Please complete this form in full and return to OS-Section54@ofcom.org.uk.

<https://doi.org/10.5258/SOTON/PP0126>