

WARNING: This consultation response contains language and/or material that may be distressing

Your response

Question	Your response

WARNING: This consultation response contains language and/or material that may be distressing

Question 1: Do you have any comments on our proposed approach to 'content and activity' which 'disproportionately affects women and girls'?

Yes. I welcome the focus areas and the intention to improve safety for women and girls under the Online Safety Act. The emphasis on governance, abusability testing, and safer default settings is especially important. However, the framework could be strengthened further by addressing:

- **Age-appropriate versions of guidance and safety tools** - Ofcom should ensure that the final guidance is accessible and understandable for different age groups, in line with the UN Convention on the Rights of the Child (UNCRC). Just as the UNCRC publishes child-friendly versions of complex legal documents, this guidance could include materials specifically tailored to teenage girls or younger users.
- **Clear information on what happens when something is reported.** - A common reason young people avoid reporting abuse or harm is fear of retaliation, not being believed, or unclear outcomes. There should be a strong requirement for platforms to outline, in plain language, what happens after a report is submitted, whether it is anonymous, and what actions are taken.
- **Shadowbanning concerns.** - Especially on platforms like Instagram or TikTok, users who report abuse often fear being silenced. Platforms must ensure reporting does not result in penalties for the victim.
- **Intersectionality** - recognising how race, disability, sexuality, age, and class intersect with gender to compound harm. - The guidance could do more to recognise how race, disability, sexuality, gender identity, age, and class intersect with gender, compounding risk and harm. Women and girls who experience multiple forms of marginalisation often face more targeted and severe abuse, and this needs to be explicitly addressed in company risk assessments and mitigation strategies.
- **Algorithmic bias:** ensuring safety-by-design applies not only to moderation tools but also to content recommendation systems. Platforms often amplify harmful content through engagement-based ranking systems. Safety-by-design must extend to algorithms themselves—not just moderation tools or user settings.
- **Stronger checks for proof of age** - Currently, many

Confidential?
- N

WARNING: This consultation response contains language and/or material that may be distressing

online platforms only require users to enter a date of birth to create an account. This system is easily bypassed by entering a false date, meaning that children—especially younger girls—can access content, interactions, and environments that are inappropriate or unsafe for their age group. The guidance should require platforms to implement much stronger, privacy-conscious methods for age assurance. This might include AI-driven age estimation, verification via educational institutions or guardians (with consent), or two-step account setup for under-18s with locked safety settings by default. Platforms must also ensure that age checks are re-validated over time. A single age check at sign-up is not sufficient to protect minors as they continue using the platform and encounter different types of content.

Question 2: Do you have any comments on the nine proposed actions? Please provide evidence to support your answer.

Yes, I agree with the proposed actions in principle, but several key areas need strengthening to ensure the safety of women and girls online. Below are specific recommendations, each supported by evidence:

1. Transparency and Independent Auditing

Recommendation: Self-reporting is insufficient. Platforms should be required to undergo regular, independent safety audits to ensure truthful reporting and accountability.

Evidence:

- The 5Rights Foundation and the UN Special Rapporteur on freedom of expression warn that self-assessments often downplay systemic risks tied to commercial interests.

https://5rightsfoundation.com/wp-content/uploads/2025/04/Joint-submission_-recurrent-and-prominent-systemic-risks-in-the-EU-and-on-measures-for-their-mitigation-2.pdf

- The “Facebook Papers” leak (2021) revealed

WARNING: This consultation response contains language and/or material that may be distressing

serious gaps between internal safety data and public disclosures, including harms to teen girls.

<https://www.wsj.com/articles/the-facebook-files-11631713039>

- The Ada Lovelace Institute recommends regulatory oversight of algorithmic and moderation practices, citing major transparency gaps.

https://www.adalovelaceinstitute.org/wp-content/uploads/2021/12/ADA_Technical-methods-regulatory-inspection_report.pdf

2. Stronger Protections for Under-18s

Recommendation: Platforms must enforce stricter default protections for under-18s, especially young girls who face disproportionate harm.

Evidence:

- Ofcom’s 2024 report shows that children under 13 continue to use social media platforms, despite stated age limits.

https://www.ofcom.org.uk/_data/assets/pdf_file/0026/274593/Adults-media-use-and-attitudes-report-2024.pdf

- The Children’s Code (ICO) mandates higher privacy and safety standards for under-18s, but enforcement across global platforms remains limited.

<https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/childrens-information/childrens-code-guidance-and-resources/age-appropriate-design-a-code-of-practice-for-online-services/>

- The UK Children’s Commissioner reported that 60% of teenage girls had seen harmful content online, including self-harm and eating disorders.

<https://assets.childrenscommissioner.gov.uk/wpuploads/2024/10/Ive-seen-horrible-things-Online-Safety-Report-October-2024.pdf>

3. Support for Reporters

Recommendation: Platforms must improve signposting to support services for victims, especially those experiencing stalking, image-based abuse, or gendered harassment.

Evidence:

- The Revenge Porn Helpline found 74% of users didn't know where to go for support after an incident.
<https://swgfl.org.uk/research/revenge-porn-helpline-2022-report/>
- The EAWW Coalition highlights that survivors are often retraumatized by platform inaction or confusing processes.
<https://www.endviolenceagainstwomen.org.uk/wp-content/uploads/EAWW-OSB-Second-Reading-Briefing-April-2022.pdf>
- A Girlguiding UK survey found 48% of girls aged 13–17 saw upsetting content but didn't know how to report it.
<https://www.girlguiding.org.uk/globalassets/docs-and-resources/research-and-campaigns/girls-attitudes-survey-2024.pdf>

4. Clarity on Reportable Content

Recommendation: Young users must be educated about what constitutes harmful or illegal content, with clearer in-app guidance.

Evidence:

- Ofsted's review found that sexual harassment is often normalised by young people who see it online and do not recognise it as abuse.
<https://www.gov.uk/government/publications/review-of-sexual-abuse-in-schools-and-colleges/review-of-sexual-abuse-in-schools-and-colleges>
- The UK Safer Internet Centre reports that more

Confidential?
– N

WARNING: This consultation response contains language and/or material that may be distressing

than 50% of young people don't report harmful content because they're unsure if it breaks rules.

<https://saferinternet.org.uk/blog/uk-safer-internet-centre-welcomes-the-enforcement-of-ofcoms-illegal-harms-codes-but-there-are-still-significant-gaps-to-address>

5. Bystander Reporting

Recommendation: Platforms should support third-party reporting and ensure privacy laws are not misused to block it.

Evidence:

- The Law Commission (2021) recommends legal clarity to support witnesses and bystanders who report harm.

<https://assets.publishing.service.gov.uk/media/61ba022ad3bf7f05539de6f5/Modernising-Communications-Offences-2021-Law-Com-No-399.pdf>

- The NSPCC highlights that in grooming cases, bystanders often identify harm before the victim is aware.

<https://www.nspcc.org.uk/about-us/news-opinion/2024/online-grooming-crimes-increase/>

- Brook and CEOP found teenagers are more likely to report concerns about peers than themselves, but doubt the platform will act.

<https://www.ceopeducation.co.uk/professionals/guidance/social-media-and-young-people/>

6. Moderation Must Apply Equally Regardless of Verification

Recommendation: Platforms must apply the same moderation standards to verified accounts as non-verified users.

Evidence:

WARNING: This consultation response contains language and/or material that may be distressing

- Reset.Tech found that some verified accounts spread harmful content without being held to moderation standards.

<https://www.reset.tech/resources/reset-tech-research-note-doppelganger-revamped-network-of-verified-accounts-spreads-multilingual-propaganda-on-x.pdf>

- EAW and Amnesty document the systemic dismissal of abuse reports made against verified users or public figures.

<https://www.endviolenceagainstwomen.org.uk/wp-content/uploads/EAW-OSB-Second-Reading-Briefing-April-2022.pdf>

- Amnesty International’s research shows platforms failed to act on abuse by influential users, particularly during global crises.

<https://www.amnesty.org/en/latest/research/2018/03/online-violence-against-women-chapter-1-1/>

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.

The good practice steps outlined in Chapters 3, 4, and 5 are welcome. However, to ensure they are meaningful in practice, especially for women and girls, Ofcom must ensure these steps are underpinned by enforceable standards, inclusive design processes and stronger obligations on platforms with known exposure to gender-based online abuse.

1. Youth-Accessible Transparency Reports and Public Dashboards

Platforms must provide transparent and accessible information about how they are tackling abuse that disproportionately affects girls and young women.

Evidence:

- **The Web Foundation’s** report “*The Tech We Want*” (2021) highlights the need for gender-disaggregated transparency reporting to better track and respond to

WARNING: This consultation response contains language and/or material that may be distressing

online abuse targeting women and girls - <https://webfoundation.org/research/the-tech-we-want>

- **5Rights Foundation** advocates for youth-accessible digital transparency, including how safety policies are communicated to young people. <https://5rightsfoundation.com/our-work/youth-engagement/>

2. In-App Safety Walkthroughs Tailored to Girls' Needs

Girls face unique pressures online—including image-based abuse, unsolicited messages, and exposure to harmful beauty standards. Safety walkthroughs must be age-appropriate and account for gendered experiences.

Evidence:

- **Plan International UK's "The State of Girls' Rights in the UK"** (2020) shows that girls often don't know how to report abuse or are reluctant to do so due to fear of backlash or shame. Safety tools must speak directly to their experiences. <https://plan-uk.org/state-of-girls-rights>

- **Girlguiding's Girls' Attitudes Survey** (2022) found that 71% of girls aged 7–21 had experienced some form of online harm. Accessible, empowering safety features must be embedded into platforms. <https://www.girlguiding.org.uk/girls-making-change/girls-attitudes-survey/>

3. Co-Design with Young Women and Marginalised Girls

Co-designing safety tools with young women—particularly those from marginalised backgrounds—is crucial to effectiveness and trust.

Evidence:

- **The Social Switch Project** highlights the value of peer-led digital safety interventions, especially in preventing gang-related grooming and exploitation online. <https://www.thesocialswitchproject.org.uk>
- **Youth Online Safety** projects advocate for safety

Confidential?

– N

WARNING: This consultation response contains language and/or material that may be distressing

tech designed *with* young people rather than *for* them, to better reflect real risks and effective interventions.
<https://youthonlinesafety.org>

4. Stricter Obligations on Platforms with Under-18 Users

Platforms must do more to limit young girls' exposure to sexualised content, harassment, and algorithm-driven harm.

Evidence:

- **Tech Transparency Project & Reset's "Thinstagram"** – investigation into how Instagram's algorithms promote content related to eating disorders, disproportionately impacting young girls by reinforcing harmful body image ideals.

<https://www.techtransparencyproject.org/articles/thinstagram-instagram-algorithm-fuels-eating-disorder-epidemic>

- **Ofcom's own research (2024)** on children's exposure to harmful content underlines that current protections are inadequate for platforms with large underage user bases.

<https://www.ofcom.org.uk/media-use-and-attitudes/media-habits-children/children-and-parents-media-use-and-attitudes-report-2024>

5. Additional UK-Based Good Practice Examples

- **Glitch UK** runs training and digital self-defence programmes specifically for Black women and women of colour facing online abuse. They provide a model for targeted, survivor-informed safety design.

<https://glitchcharity.co.uk>

- **#EndOnlineAbuse** campaign by EAW (End Violence Against Women Coalition) demonstrates how platforms and policy can work better with women's organisations to tackle systemic online harms.

<https://www.endviolenceagainstwomen.org.uk/campaign/online-abuse/>

WARNING: This consultation response contains language and/or material that may be distressing

Summary

To maximise the **effectiveness and equity** of good practice steps:

- Platforms must **publicly demonstrate progress** via youth-friendly reports;
- Tools and guides must be **gender-sensitive, age-appropriate**, and **co-designed** with diverse girls and young women;
- Special attention must be given to **AI, algorithms, and nudity moderation**, which often reinforce gendered harm.

Without centring the lived realities of women and girls—particularly those most marginalised—the good practice steps risk being well-meaning but ineffective.

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?

The proposal to assess and publish how providers are addressing women and girls’ safety is a welcome and necessary step. Public accountability is a strong motivator, especially when combined with clear performance benchmarks and reputation or commercial incentives.

To **encourage compliance and proactive adoption of good practice**, Ofcom could strengthen its approach in the following ways:

1. Public Safety Performance Dashboards

Recommendation: Publish a regular, user-friendly online safety dashboard for each platform, with metrics specifically related to women’s and girls’ safety.

- **Why it works:** Transparency builds public trust

WARNING: This consultation response contains language and/or material that may be distressing

and pressures providers to improve. Metrics such as reporting rates, time to action, prevalence of gendered abuse, and use of safety features can all be disclosed.

- **Evidence:** The Web Foundation emphasizes the importance of gender-disaggregated transparency reports as essential to understanding and reducing online harms against women and girls. Their report on online gender-based violence and abuse advocates for such data to hold platforms accountable and develop effective solutions Tech Policy Design Lab: Online Gender-Based Violence and Abuse

- **Evidence:** The Information Commissioner’s Office (ICO) has encouraged providers to adopt more transparent practices, including child-impact assessments and reports on data handling, through the implementation of the Children’s Code Children’s code guidance and resources | ICO

2. Sector-wide Safety Accreditation or Kitemark

Recommendation: Introduce a voluntary but high-profile “**Online Safety for Women and Girls**” accreditation (e.g. a digital safety kitemark) for platforms that meet or exceed Ofcom’s safety expectations.

- **Why it works:** Positive incentives can complement enforcement. Accreditation recognises compliance and encourages platforms to compete on safety.

- **Evidence:** The UK’s Cyber Essentials certification and the Fair Tax Mark show how third-party accreditation schemes can improve organisational standards through public visibility and trust <https://www.ncsc.gov.uk/cyberessentials/overview>

3. Incorporate Women and Girls’ Safety into Service-

WARNING: This consultation response contains language and/or material that may be distressing

Level Agreements and Procurement

Recommendation: UK public sector procurement rules could include online safety commitments as part of digital service contracts, favouring vendors that follow Ofcom's good practice guidance.

- **Why it works:** Economic incentives can drive change faster than regulation alone, especially for mid-size platforms.

- **Evidence:** The UK Government's Procurement Policy Note (PPN) 002 outlines the Social Value Model, which incorporates social value into procurement criteria. Expanding this to include digital safety for women and girls would align public funds with public protection PPN 002: Taking account of social value in the award of central government contracts

Confidential?
– N

4. Co-Design with Lived Experience Experts

Recommendation: Require or incentivise platforms to work with **women's rights organisations, young people, and survivors** to co-design safety features and risk assessments.

- **Why it works:** Participatory design ensures that solutions reflect the real experiences of harm and build more trust among users.

- **Evidence:** Coventry University's framework for co-designing digital safety tools with young women is a leading example in the UK context <https://www.coventry.ac.uk/research/about-us/research-news/2024/researchers-develop-framework-digital-toolkit-global-online-safety-women-girls/>

5. Time-bound Benchmarks with Tiered Accountability

Recommendation: Set clear **implementation timelines**

WARNING: This consultation response contains language and/or material that may be distressing

and benchmarks for each good practice, with a tiered system for encouraging adoption (e.g. traffic-light rating, naming top/bottom performers, public or parliamentary reporting).

- **Why it works:** Timelines create urgency and signal consequences. Tiered systems show progress and spotlight leaders and laggards.
 - **Evidence:** Ofsted and the EHRC use similar public grading systems to drive compliance in education and equality contexts.
-

6. Promote Peer Learning and Sector Standards

Recommendation: Create a **cross-sector safety learning forum** facilitated by Ofcom, where providers can share innovations, challenges, and case studies related to women and girls' online safety.

- **Why it works:** Smaller platforms in particular benefit from shared knowledge. This also raises the baseline for the sector.
 - **Evidence:** The Web Foundation's Tech Policy Design Lab uses a collaborative model to build solutions across companies and civil society <https://www.girlguiding.org.uk/girls-making-change/girls-attitudes-survey/>
-

7. Encourage Youth and Public Engagement in Oversight

Recommendation: Allow for public or youth input into platform safety reviews — through citizen panels, surveys, or consultative roles — with attention to under-25s and girls in particular.

- **Why it works:** Including affected users ensures relevance and strengthens the social contract between platforms and the public.
- **Evidence:** Girlguiding UK's "Girls' Attitudes

WARNING: This consultation response contains language and/or material that may be distressing

Survey” shows how engaging girls in policy decisions can yield targeted, responsive recommendations <https://www.girlguiding.org.uk/girls-making-change/girls-attitudes-survey/>

Summary

To encourage providers to follow the guidance, Ofcom should go beyond assessments by combining public transparency, economic incentives, timelines, accreditation schemes, and participatory design. Clear visibility of outcomes, and recognising both good and poor practice, will drive a culture shift toward safer digital environments for women and girls.

Question 5: Do you have any comments on our impact assessment, rights assessment, or equality impact assessment? Please provide any information or evidence in support of your views.

The Equality Impact Assessment is a positive step; however, to ensure comprehensive protection, it should delve deeper into **intersectional inequalities**. Women and girls who experience overlapping marginalisations—such as race, disability, or LGBTQ+ identity—face distinct and compounded online harms. Additionally, the assessments should explicitly address the **risks associated with over-reliance on automated systems**, particularly concerning cultural or linguistic biases that can lead to disproportionate impacts on marginalised groups.

1. Intersectional Risk and Overlapping Harms

Recommendation: Enhance the Equality Impact Assessment to explicitly consider how intersecting identities (e.g., race, disability, sexual orientation) amplify online harms for women and girls.

- **Evidence:** Amnesty International’s research

WARNING: This consultation response contains language and/or material that may be distressing

indicates that Black and minority women in the UK experience higher volumes of online abuse. For instance, Black women MPs received 35% more abusive tweets than their white counterparts, highlighting the compounded nature of racial and gender-based online harassment. https://www.amnesty.org.uk/online-violence-women-mps?utm_source=chatgpt.com

- **Evidence:** Glitch UK’s report, “The Ripple Effect,” found that Black women are 84% more likely to experience online abuse than white women. The report also notes that LGBTQ+ women and disabled women face uniquely patterned harms, emphasising the need for intersectional approaches to online safety. https://committees.parliament.uk/writtenevidence/39245/pdf/?utm_source=chatgpt.com

- **Evidence:** UN Women’s “Intersectionality Resource Guide and Toolkit” underscores the necessity of integrating intersectional design into safety standards to ensure effective protection for all women and girls, particularly those with disabilities. https://gendercoordinationandmainstreaming.unwomen.org/index.php/resource/intersectionality-resource-guide-and-toolkit?utm_source=chatgpt.com

- **Evidence:** The UK Government’s “Adult Online Hate, Harassment and Abuse” rapid evidence assessment highlights patterns of abuse based on age, sexual orientation, race, and religion, underscoring the importance of considering intersectionality in addressing online harms. https://www.gov.uk/government/publications/adult-online-hate-harassment-and-abuse-a-rapid-evidence-assessment?utm_source=chatgpt.com

Confidential?
– N

2. Risks of Over-Reliance on Automated Systems

Recommendation: Acknowledge and address the potential for automated content moderation systems to inadvertently harm marginalised communities due to cultural or linguistic biases.

WARNING: This consultation response contains language and/or material that may be distressing

<ul style="list-style-type: none">• Evidence: Research by MIT and Stanford University revealed that facial recognition AI systems exhibit significant biases, underperforming for individuals with darker skin tones and women. This underperformance raises concerns about the reliability of automated systems in accurately identifying and moderating content related to marginalised groups. https://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf• Evidence: The Mozilla Foundation reports that AI-driven content moderation tools often have high error rates when applied to cultural, dialectal, or regional variations. For example, Instagram’s content moderation system erroneously flagged and removed the hashtag #Sikh, demonstrating how automated systems can misinterpret culturally significant content. https://www.mozillafoundation.org/en/• Evidence: AlgorithmWatch’s 2022 Annual Report highlights that automated tools can inadvertently penalise marginalised voices, leading to false positives that silence advocacy and disproportionately affect those already at risk of online harms. https://algorithmwatch.org/en/	
<p>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.</p> <p>Yes. It is positive that the guidance supports treating Welsh equally. This commitment should extend to moderation services, automated systems, and support tools, ensuring that Welsh-speaking users have equal access to safety mechanisms and content reporting features. It also demonstrates a commitment to diversity and inclusion in other areas.</p>	<p>Confidential? – N</p>

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