

The Online Safety Act: Implementation and Impact Submission

This submission has been prepared by [Libby Barr](#), Policy Advisor (Technology, Justice & Cyber) at UCL's [Policy Impact Unit](#) on behalf of the Department of Computer Science and the Department of Security and Crime Science at University College London.

The Policy Impact Unit provides professional policy engagement expertise across the Faculty of Engineering, collaborating with leading researchers to help incorporate evidence into the policy decisions.

Consultation was undertaken with the following academics:

- Kyle Beadle, Dept of Computer Science
- Enrico Mariconti, Dept of Security and Crime Science
- Leonie Tanzer, Dept of Computer Science
- Marie Vasek, Dept of Computer Science

This submission is based on research findings presented in the paper, [“Toward inclusive online safety: evaluating the role of intersectionality on online abuse in the United Kingdom”](#), led by UCL researchers Kyle Beadle, Enrico Mariconti, Leonie Tanzcer and Marie Vasek.

Submitted 21 September 2026

Impact of the Online Safety Act

Has the OSA been effective in protecting groups disproportionately affected by harms within its scope, including women and girls, and those vulnerable to exploitation?

1. **The Online Safety Act has not been effective in protecting groups disproportionately affected by harms. This is due to the act prioritizing procedural compliance over addressing the social conditions that enable abuse.** As a result of the Act being largely reactive to how platforms currently exist, it continues to overlook and disproportionately expose users with marginalized identities to severe online harm. In addition to gaps in existing regulatory practices from Ofcom; current reporting methods do not account for intersectionality and overlook the scale and impact of online harms for all social identities.
2. **Understanding intersectionality is important to fully understand the impact of online harm, and therefore the effectiveness of the OSA.** Online harm is entangled with social hierarchies, such as how combinations of gender, race, and sexuality shape who is targeted and how protection is distributed. For example, Black, LGB+, women are more likely to experience online harms that are consistent with the cumulative impact of racism, homophobia, and sexism. If these experiences are viewed solely through the lens of sexism, then the nuance of the type of harms, and why those harms are occurring are lost for these women and girls.
3. **Our research evidences that the groups most likely to experience online harms are people with intersectional identities, who are often marginalised and disproportionately affected by harms.** The results of our study demonstrate that this abuse is amplified in online spaces. Since the OSA received Royal Assent in 2023:
 - Non-binary people experience increased prevalence and more types of harm (93%) than men (69.7%) and women (66.9%)

- Ethnic minority groups are more likely to experience online harm compared to white groups
- Online harm is more prevalent, and the types of harm are more varied among LGB+ people (81.6%) than straight people
- Age group 18-24 experienced the highest frequencies of harm at 82%
- Black non-binary people between the ages of 18-24 were predicted to experience most types of harm at 95%
- White LGB+ nonbinary people aged between 18-24 are predicted to be most likely to experience harm at 95%

Have the OSA's provisions been effective in driving services to a) identify relevant risks, b) take proportionate action that adequately mitigates those risks, and c) make meaningful changes to their systems, processes and design features?

4. **Our research highlights three main limitations with the OSA's approach of mandating compliance via a systems- and processes-based model** which requires platforms to implement risk assessments, mitigation measures, and provide mechanisms for reporting and redress.¹ These are,
 - I. **Platform risk assessments. The OSA does not require platforms to conduct intersectional vulnerability assessments in a sufficiently explicit or operationalised way.** This perpetuates harms for those who are most likely to experience the highest frequency and variation of harms. Stronger enforcement is

¹ Huber, A. R., & Ward, Z. (2025). Non-consensual intimate image distribution: Nature, removal, and implications for the online safety act. *European Journal of Criminology*.

Nash, V., & Felton, L. (2024) Treating the symptoms or the disease? analysing the UK online safety act's approach to digital regulation. *Policy & Internet*.

Smith & M., Nolan, M., & Gaffey, J. (2024). Online safety and social media regulation in Australia: eSafety commissioner v X corp. *Griffith Law Review*.

Walker, A. M., DeVito, M. A., Badillo-Urquiola, K., Bellini, R., Chancellor, S., Feuston, J. L. (2024). "What is safety?": Building bridges across approaches to digital risks and harms. In Companion publication of the 2024 conference on computer-supported cooperative work and social computing.

required from Ofcom to ensure intersectional vulnerability assessments are embedded into the platforms' risk assessment and mitigation processes.

- II. **Mitigation measures are currently too reactive and do not go far enough in proactively limiting the opportunity for harms to occur online.** For example, certain platform functionalities act as catalysts which amplify online harm to marginalised groups. Marginalised identities become visible through profile information, names, images, language, community affiliation, hashtags, or participation in identity-based communities.² These markers make users legible to hostile audiences and can shape the type of abuse they then experience. Additionally, digital infrastructures, by design through algorithmic ranking, then recommend, search, tag, and share accounts which can bring marginalized users to the attention of hostile audiences.³ The findings from our research conclude that the online landscape will not become less harmful until the OSA pursues structural, proactive transformations in platform design and governance.
- III. **Mechanisms for reporting and redress. The provisions place the onus on providers to make singular, reactive changes such as focusing on content take down efforts or making amendments to content moderation.** Yet, the majority of these actions can only be taken once harm has already occurred to the user, instead of preventing harmful activity in the first place. They do not adequately incentivise nor enforce platforms to change approach from addressing individual incidents and account for systemic changes needed to address digital harms. The OSA needs to

² Carrasco, M., & Kerne, A. (2018). Queer visibility: Supporting lgbt+ selective visibility on social media. In Proceedings of the 2018 chi conference on human factors in computing systems.

³ Heung, S., Jiang, L., Azenkot, S., & Vashistha, A. (2024). "Vulnerable, victimized, and objectified": Understanding ableist hate and harassment experienced by disabled content creators on social media. In Proceedings of the CHI conference on human factors in computing systems.

go further in mandating structural design changes to platforms that stop the possibility of illegal content, and illegal but harmful content, to occur.

5. **As a result of these challenges, the OSA emphasizes platform compliance over platform scrutiny.** Although the Act requires platforms to implement mitigations and undergo procedural reviews, it does not require them to assess whether their systems reproduce identity specific or intersectional patterns of harm.
6. **Ultimately, online safety requires structural transformation and community-led governance and cannot be achieved through procedural compliance alone.** To achieve this, the OSA must both greater incentivise platform scrutiny to explore the root cause of why harms are occurring and deploy methods to limit the ability to perpetrate these harms via online platforms.

Ofcom's implementation and enforcement

Has Ofcom's interpretation and implementation of the OSA to date given appropriate effect to its provisions?

7. **Ofcom has a mandate to keep all online users safe and set standards for how platform providers steward behaviour online. Yet, it operates on the assumption that online harms occur randomly as one-off events due to bad actors, or they are seen as unfortunate hazards users accept the risk of encountering to engage in online spaces.**
By prioritising individualized risk frameworks, Ofcom does not account for the structural inequalities that drive online behaviour at the root. This approach presents a significant gap in how the OSA is implemented to give effect to its provisions as our work provides evidence that the people experiencing the greatest harm are those who face the most structural discrimination, offline and online.
8. **To address this limitation, we recommend that online abuse is understood as a result of structural social inequalities which are amplified online,** not interpreted by Ofcom as something that is perpetrated by bad actors or an unavoidable hazard that users accept to participate online.

What changes, if any, are needed to improve the effectiveness of Ofcom's a) illegal content and protection of children codes of practice, b) guidance on violence against women and girls, or c) wider regulatory approach?

9. **To improve Ofcom's wider regulatory approach, Ofcom should change the way it considers risk to users.** Our study reveals that those who identify within overlapping marginalised groups are most at-risk, subject to the widest harm range, and experience the highest number of online harms. However, Ofcom's reporting metrics do not explicitly measure or report on intersectionality, instead focusing on individual identity groups, such

as gender or race or ethnicity or sexuality and, even within that uses very broad terminology. For example, Ofcom use categories such as “children”, “women and girls” or “ethnic minorities”, which limits the ability to derive richer data insights from the surveys to address inequalities. This undermines reporting of the prevalence of online harms and makes it challenging to derive meaningful data with which to hold online service providers to account. For platforms and regulators, broad harm categories can obscure which groups face the highest risk and from which types of harm.

10. **The continuation of this approach risks encouraging both intentional and accidental oversight of the root cause and drivers of online harm and fails to hold companies to account of their role in protecting the most vulnerable users.** To address this, regulation activities need to explicitly report on the impact of intersectionality on users’ experiences online to better capture the nuance of the harm experienced and effectiveness of policy interventions.
11. **To improve the quality of data collected and analysed Ofcom should sample intersectional identity** by investigating the specific experiences of different ethnicities or the intersectional experiences of respondents, report on how representative their sampling is with the current UK population and expand their harm categories to capture false reporting, account lockout or control, explicitly measure experiences such as “People sending unwanted/unsolicited sexual or nude images or videos, e.g. ‘cyber flashing.’”
12. **To improve how Ofcom data is analysed, advanced statistical methods when reviewing data should be used to derive deeper data-driven insights into the impact of these harms on users.**

Where duties are in force, how effective and timely have Ofcom’s supervision and enforcement been in driving improvements and securing compliance?

13. **Our research has found that no offline events, such as new regulation, social movements, or technical changes, have impacted people’s experiences of online**

harm in the UK since the Online Safety Act received royal assent. The OSA mandates providers of online services to assess risk and protect users from online harm. However, the mechanisms through which service providers manage and monitor risk (e.g reporting, setting moderation thresholds for illegal and harmful content, and platform risk assessments) treat patterned abuse as isolated incidents. Our research has identified that this is an inaccurate understanding of online harms and so demonstrates the ineffectiveness of Ofcom’s supervision of enforcement in driving compliance.

14. **Currently, the real risk picture of online harms is obscured.** Online harms, in particular the intersectionality of them, are inadequately measured and therefore the subsequent policy implications are not considering this issue in the round. If this continues, policies will fail to protect the marginalised groups they aim to support. Change is imperative not only to ensure impactful policy development but also because the collection and analysis of public data via surveys serve as a measurement tool for how successful technology companies are in complying with the Online Safety Act. These insights are then used as a benchmark to highlight where compliance must go further to protect users.

15. **To address this, we recommend that future surveys and reporting methods should consider intersectionality when analysing the data recorded to better monitor the effectiveness of online harm interventions designed by regulators, policymakers and technology companies.** To truly make the online world safer for all, online regulation must understand not only what is happening to users, but why it is happening and how it manifests. This approach would allow Ofcom to be more proactive in addressing online social conditions which enable harms, rather than prioritising reactive measures.