

Your response

Question	Your response
The incidence of content that is harmful to children on regulated services We invite stakeholders to provide evidence relating to the incidence of PPC, PC or NDC on regulated user-to-user, search or combined services. Please provide as much detail as you can. This could include but is not limited to evidence about:	
The quantity, prevalence or presence of specific kind(s) of PPC, PC or NDC on regulated services and the frequency with which children are encountering this content.	Confidential – N In May 2025, the Mental Health Foundation reported that harmful online content had become widespread among young people and represents a major emerging threat to their mental health. A UK survey of 2,157 people aged 16–21 found that 68% had encountered disturbing online material, including suicide or self-harm content (35%), pro-eating-disorder content (28%), racism or discrimination (42%), bullying (40%), and misogynistic content (32%). Many young people reported feeling unsafe on major social media platforms. The Foundation warned that exposure to such content can seriously harm wellbeing and calls for stronger online safety education in schools and improved protections from harmful digital environments. While the PPC, PC or NDC categories are language-neutral in law, their effective operation in practice depends on robust language coverage within platforms’ moderation systems, risk assessments and safeguards. If Welsh language content is less likely to be flagged, assessed or acted upon by platform safeguards, this could create a relative risk that children who use Welsh may be exposed to harmful content that is not being detected or mitigated to the same extent as equivalent English language content. At present, there is limited published evidence on the prevalence of such material in Welsh, which is why we consider it important for Ofcom to test

Question	Your response
	and publish evidence on Welsh language-specific detection and moderation.
How demographic characteristics such as age, gender, race, sexual orientation and disability influence the likelihood that children will encounter specific kind(s) of PPC, PC or NDC.	Confidential – N An intensive approach to data collection by online services can facilitate the creation of highly personalised content feeds leading to the amplification of harmful content especially for vulnerable users. An investigation in 2023 by Amnesty International into TikTok’s content targeting found that its hyper-personalised ‘For You’ feed could quickly pull children and young people who show even minimal interest in mental health topics into “rabbit holes” of harmful content, including videos that romanticise depression, self-harm, and suicide. The platform’s design—optimised for engagement and rapid recommendation—exposed vulnerable users to escalating amounts of distressing material, worsening risks of anxiety, depression, and self-harm. The report argued that TikTok’s algorithmic systems and business model actively amplified this dangerous content, placing young users’ mental and physical wellbeing at significant risk. It also published a companion report which examined how TikTok’s business model and surveillance-driven data practices put children and young people at risk. It highlighted that TikTok’s recommender system relies on intensive data collection, creating highly personalised content feeds that can amplify depressive, self-harm, and suicidal content, especially for vulnerable users. The report argued that TikTok’s design prioritised engagement and data extraction over user safety, applying protections unevenly across regions. Combined with the platform’s manipulative recommendation algorithm, these surveillance practices expose young people to serious mental-health harms, deepen risks for those already struggling, and underscore the need for stronger regulation and platform accountability.

Question	Your response
The services or types of service where children are most likely to encounter specific kind(s) of PPC, PC or NDC.	Confidential – N Children who use Welsh are also active users of main-stream social media platforms, messaging services, search engines and, increasingly, AI-enabled tools and conversational agents. It follows that they may encounter content harmful to children in Welsh in the same ways as English-speaking children encounter such content in English. Therefore, we would like Ofcom to re-search and publish evidence that harmful content in Welsh is being detected, classified, moderated and mitigated with the same degree of accuracy, consistency and speed as is the case with English language content. Beyond this Call for Evidence, we would also ask Ofcom to consider whether the current evidence base sufficiently captures harms experienced by Welsh-speaking children, or whether additional data collection, testing or stakeholder engagement is required. Without such assurance, there is a risk that harmful content in Welsh remains under-reported and under-regulated, contrary to the Act’s intention that all children in the UK should benefit from equivalent levels of protection online.
The severity of harm that UK children suffer, or may suffer, as a result of content that is harmful to children We invite stakeholders to provide any evidence regarding the severity of harm that UK children suffer, or may suffer, as a result of encountering PPC, PC or NDC. Please provide as much detail as you can. This could include but is not limited to evidence about:	
The impact that specific kind(s) of PPC, PC or NDC can have on a child’s wellbeing, behaviours, or other physical or psychological outcomes. This	Confidential – N Several recent reports/reviews have highlighted the impact of harmful content on young people.

Question	Your response
might include evidence concerning cumulative exposure to specific kind(s) or combinations of PPC, PC or NDC.	A systematic review published in the Journal of Child Psychology and Psychiatry in 2023 examined 15 empirical studies on the impact of viewing self-harm images online. All studies identified harmful effects, including escalation of self-harm, reinforcement of self-harm behaviours (such as sharing or commenting), increased social comparison, development of a self-harm identity, and emotional or physiological responses that can trigger urges or acts of self-harm. However, nine studies also reported protective effects, such as reduced urges, increased social connection, support-seeking, and encouragement of recovery. No studies established causality, and mechanisms were often unexplored. Overall, harmful effects predominated, but impacts were complex and varied across individuals. A 2023 update of a review published by the National Suicide Research Foundation, Ireland examined evidence on the harmful impact of online suicide and self-harm content. It identified how search engines, social media, imagery, videos, forums, pro-suicide sites, suicide “games,” the Darknet, and livestreamed suicides can normalise, encourage, or facilitate self-harm. Exposure is linked to contagion effects, method learning, increased distress, and reinforcement of harmful behaviours. Vulnerable individuals, especially young people, face heightened risks. A report published by NHS Greater Glasgow and Clyde Health Board in November 2024 reviewed research on social media use among young people, focusing on mental health, self-harm, and suicidality. It outlined how young people use social media for escapism, connection, and peer support, while highlighting risks such as exposure to harmful content and negative social comparison. Overall, it emphasised the complex, mixed impact of social media, calling for clearer research approaches and better support for young people online.
Any causal links between a child encountering specific kind(s) of PPC, PC or NDC and subsequently suffering physical or psychological harm.	

Question	Your response
How demographic characteristics such as age, gender, race, sexual orientation and disability can influence the severity of harm that UK children suffer, or may suffer, as a result of encountering specific kind(s) of PPC, PC or NDC.	Confidential – N While the initial focus has been on the potential harms generated by online services, this should be extended to genAI and chatbot technologies. Evidence from civil society and academic research suggests that many large language models and content moderation systems perform less reliably in minority and lesser-resourced languages. AI tools can generate fluent Welsh-language responses and content. However, where safety guardrails, refusal mechanisms and harm-prevention policies are primarily trained or tested in English, there is a risk that harmful outputs in Welsh may not be adequately prevented or intercepted. Taken together, these limitations in language coverage raise wider questions about how gaps in detection and moderation may shape harmful behaviour over time, particularly where perpetrators adapt their tactics to avoid platform safeguards. An additional emerging risk is that uneven or weaker flagging of harmful content in minority languages, including Welsh, could create an incentive for malicious actors to exploit those languages as a means of avoiding detection when engaging in bullying, abuse or other harmful behaviours. At present, the extent to which this is occurring, or may occur in future, is unclear, which further underlines the need for Ofcom to test and publish evidence on language-specific safeguards. Therefore, we consider it essential that the statutory report addresses language coverage explicitly and encourage Ofcom to provide evidence on: • how it is assuring itself that regulated services are identifying and mitigating PPC, PC and NDC in Welsh to a standard equivalent to English. • what engagement Ofcom has undertaken with social media companies and AI developers regarding Welsh-language moderation, reporting tools and safety guardrails. • whether children’s risk assessments and content-classification processes used by services explicitly consider minority languages, including Welsh.

Question	Your response
	• how Ofcom intends to monitor, test and enforce compliance in relation to Welsh-language content.
Whether it is appropriate to make changes to the kinds of PPC and PC in the Act We invite stakeholders to provide evidence that suggests any kind(s) of PPC and/or PC in the Act are inappropriate or incorrectly categorised. This could include but is not limited to evidence indicating that:	
Any kind of PPC or PC in the Act should cease to be categorised as content harmful to children.	
Any kind of PPC would be more appropriately categorised as PC.	
Any kind of PC would be more appropriately categorised as PPC.	
We invite stakeholders to provide evidence in support of additional kind(s) of content being added to the PPC or PC categories in the Act. This should include, but is not limited to:	
Evidence that the kind(s) of content pose a material risk to an appreciable number of children.	

Question	Your response
Evidence that the kind(s) of content pose a risk of significant harm to children.	
Evidence to support each kind of content being categorised as either PPC or PC.	

Please tell us how you came across about this call for evidence.

- ☒ Email from Ofcom
- ☐ Saw it on social media
- ☐ Found it on Ofcom's website
- ☐ Found it on another website
- ☐ Heard about it on TV or radio
- ☐ Read about it in a newspaper or magazine
- ☐ Heard about it at an event
- ☐ Somebody told me or shared it with me
- ☐ Other (please specify)

Please complete this form in full and return to ContentHarmfultoChildren@ofcom.org.uk.