

Molly Rose Foundation response: Ofcom call for evidence on content harmful to children

Question 1: The incidence of content that is harmful to children on regulated services

Molly Rose Foundation (MRF) research conducted immediately before the Protection of Children Codes came into effect in July 2025 identified that:

- a) Content that is harmful to children is present in overwhelming quantities on major social media platforms, with engagement-based features actively pushing this to children
- b) Significant numbers of children are then encountering this content

The availability and reach of harmful content on TikTok and Instagram

Our research identified that both **Instagram and TikTok were algorithmically recommending overwhelming volumes of harmful suicide, self-harm and depression content to teenage accounts.**¹

Using a well-established safety testing methodology (including an ‘avatar’ account set up as a 15-year-old girl), after a short period of engagement with suicide, self-harm and depression material, 96% of algorithmically recommended videos on TikTok’s For You Page and 97% of short-form videos of Instagram ‘Reels’ were found to be potentially harmful. On Reels, this included 8% of recommended content promoting or glorifying suicide or self-harm, and 20% referencing suicide methods.

MRF’s broader analysis of suicide and self-harm material also identified that **suicide and self-harm content was reaching vast audiences**. Harmful posts that were algorithmically recommended on TikTok’s For You Page, for example, had been liked by an average of over 303,000 accounts, with one in ten (9%) liked at least a million times. On Instagram, harmful Reels attracted an average of 226,000 likes.

Much of this material also had a high degree of interaction. Harmful Instagram reels, for example, receiving an average of 3,340 comments, and 55,227 private shares. On TikTok’s For You Page, one in five (20%) harmful posts had received more than 5,000 comments.

¹ Molly Rose Foundation (2025) Pervasive by Design

Children's exposure to harmful suicide, self-harm, depression and eating disorder content on six major social media platforms

In Spring 2025, Molly Rose Foundation conducted a large-scale survey to understand teens' exposure to potentially harmful content on six major social media platforms² across four themes: suicide, self-harm, depression and eating disorders. 1,897 children aged 13-17 from across the UK took part, covering a range of backgrounds and characteristics.³

Children were asked not only whether they had encountered content that they should be protected or prevented from encountering under the Online Safety Act (i.e. suicide, self-harm and eating disorder content that likely to be classed as Primary Priority Content (PPC), and depression content likely to be classed as Non-designated content that is harmful to children(NDC)), but also about broader content types that could contribute to cumulative harm when reviewed repeatedly or in combination with the highest risk content.

Our analysis identified that – in the previous week - **over a third (37%) of 13-17 year olds had seen at least one kind of suicide, self-harm, depression or eating disorder content likely to be classed as either PPC or NDC**. This included the following:

- **5%** of 13-17 year olds who had seen *content that encourages or promotes suicide*
- **9%** of 13-17 year olds who had seen *content that makes suicidal thoughts seem normal, appealing or cool*
- **6%** of 13-17 year olds who had seen *content that encourages or promotes self-harm*
- **8%** of 13-17 year olds who had seen *content that makes self-harm seem normal appealing or cool*
- **15%** of 13-17 year olds who had seen *content that shows self-harm*
- **17%** of 13-17 year olds who had seen *content that makes low mood or feeling depressed seem normal, appealing, or cool*
- **22%** of 13-17 year olds who had seen *content about feeling hopeless, despairing or suicidal*
- **12%** of 13-17 year olds who had seen *content that encourages or promotes eating disorders*

The frequency with which children are encountering PPC and NDC

Our analysis⁴ also identified that many 13-17 year olds were encountering suicide, self-harm and eating disorder content likely to be classed as PPC and depression content likely to be classed as NDC repeatedly, increasing their risk of experiencing cumulative harm.

² TikTok, Instagram, X, YouTube, Pinterest and Snapchat

³ Molly Rose Foundation (2025) Children's exposure to suicide, self-harm, depression and eating disorder content on social media

⁴ Ibid.

Of those who had seen each kind of PPC and NDC listed above, between 13% and 27% had seen this 10 times or more on at least one platform in the last week.⁵

This included 27% of those who had seen *content that encourages or promotes suicide*; 20% of those who had seen *content that encourages or promotes self-harm*; 14% of those who had seen *content that makes low mood or feeling depressed seem normal, appealing, or cool*, and; 13% who had seen *content that encourages or promotes eating disorders*.

How demographic characteristics influence the likelihood that children will encounter specific kind(s) of PPC and NDC

Our analysis⁶ also identified that groups of children known to be more vulnerable to online harm were at even higher risk.

Around half (49%) of the girls reported seeing at least one type of PPC or NDC in the last week, compared with a quarter (25%) of boys.

This included one in five (21%) of social media-using girls who had seen *content that shows self-harm* (compared to 9% of boys), one in five (20%) who had seen *content that encourages or promotes eating disorders* (compared to 6% of boys), and around a third (32%) who had seen *content about feeling hopeless, despairing or suicidal* (compared to 14% of boys).

Overall, girls were significantly more likely to have seen all eight of the content types likely to qualify as PPC or NDC, compared to boys.⁷

Over two thirds (68%) of 13–17 year olds categorised as having low wellbeing⁸ reported seeing at least one type of PPC or NDC in the last week, compared to one in five (20%) with high wellbeing.

This included around a quarter (23%) of those who used social media and had low wellbeing who had seen *content that encourages or promotes suicide*, compared to only 3% of those with high wellbeing.

Again, children with low wellbeing were statistically significantly more likely to have seen all eight of the content types likely to qualify as PPC or NDC, compared to boys.

⁵ This is likely to be underestimate of the overall quantity of harmful content that children were encountering, as it does not capture those who saw content types 10 times or more across multiple platforms, and it also includes some children who saw content types 10+ times across multiple platforms.

⁶ Molly Rose Foundation (2025) Children's exposure to suicide, self-harm, depression and eating disorder content on social media

⁷ By significant, we mean a statistically significant difference at the 95% confidence interval for this group compared to their peers.

⁸ Children's wellbeing was measured using the Short Warwick-Edinburgh Mental Wellbeing Scale (SWEMWBS)

Children with SEND were also at higher risk than their peers. Two in five (43%) of with SEND reported that they had seen high risk content, compared to 35% of those without.

This included 14% of those who used social media and had SEND who had seen *content that encourages or promotes suicide* (compared to 3% of those without SEND), and 10% of who had seen *content that encourages or promotes self-harm* (compared to 5% without SEND).

Overall, Children with SEND were significantly more likely to have seen five out of the eight types of content likely to qualify as PPC or NDC.

The services and surfaces where children were most likely to encounter PPC and NDC

Children were considerably more likely to be exposed to violative content on two major platforms, TikTok and X.

Comparing those who spent a similar amount of time on each platform, **children using TikTok and X were more than twice as likely to have encountered the high risk content compared to users of other platforms.** Of those spending 30–120 minutes on each platform, 31% of TikTok users and 29% of X users reported seeing at least one type of high risk content, compared to 13% of Instagram users, 13% of Pinterest users and 11% of YouTube users.

When asked where they had encountered potentially harmful content,⁹ **over half of children reported encountering this on recommender feeds**, such as Instagram’s Explore and Reels pages, or TikTok’s For You page.

This finding underscores the central role that these functionalities play in exposing children to harmful or age inappropriate content, and aligns closely with the above findings that demonstrate how aggressive algorithms push high volumes of harmful content to children in the interest of increased user engagement.

However, there were also range of **other high risk product surfaces** not currently addressed by targeted measures in Ofcom’s codes. This included 11–14% of those who had seen potentially harmful content seeing this on ‘Story’ features, and 9–13% seeing this on ‘Someone’s profile’.

Children’s exposure to harmful content in combination with compounding material, and the implications for cumulative harm

⁹ Children were asked where they had seen potentially harmful content at the theme level (i.e. suicide, self-harm, eating disorders or depression content), rather than in relation to specific content type, i.e. *content that encourages or promotes suicide*.

Though not explicitly in scope of this call for evidence's focus on PPC, PC, and NDC, MRF's large-scale survey of 13-17 year olds¹⁰ identified extremely concerning patterns of exposure that suggest many young people are at higher risk of cumulative harm than may previously have been understood.

Our methodology enabled us to determine that **not only were many children encountering high risk content repeatedly, many were seeing high risk content in combination with other forms of related material likely to increase their risk of experiencing cumulative harm.**

For example, of those who had seen content encouraging suicide or self-harm in the last week, over a third (34%) had also seen *content about feeling sad, down or lonely* 10 times or more on at least one platform.

Similarly, two in five (42%) of those who seen *content that encourages or promotes eating disorders* had also seen 'content that makes being slim look cool or desirable' 10 times or more on at least one platform.

These findings – to our knowledge the first of their kind – are deeply disturbing and demand a regulatory regime that takes a more nuanced approach to understanding and monitoring online risks, focusing not just on PPC, PC or NDC, but on the role of other 'lower risk' forms of content in compounding the impact of this material.

It is reasonably likely that children experiencing from poor mental health or low wellbeing are likely to be adversely impacted from being exposed to multiple categories of content that express and reinforce negative feelings and responses, and that may reasonable result in increased rumination on thoughts about poor wellbeing and health.

¹⁰ Molly Rose Foundation (2025) Children's exposure to suicide, self-harm, depression, and eating disorder content on social media

Question 2: The severity of harm that UK children suffer, or may suffer, as a result of content that is harmful to children

Evidence for the harmful effects of exposure to suicide, self-harm and depression content

Molly Rose Foundation has summarised recent evidence for the impact that suicide, self-harm and depression content may have on young people in ‘Pervasive by Design’ (2025). This evidence is briefly summarised here.

There is growing evidence of the relationship between exposure to harmful online content and resulting suicide and self-harm risks, with problematic Internet usage an observed antecedent in youth suicide cases.¹¹

Suicide-related Internet use has been reported in 24% of deaths by suicide among young people aged 10 to 19, equivalent to a young life being lost every single week.¹² Suicide related Internet use is recorded more frequently in the death by suicide of girls, and in cases affecting adolescents identified as LGBTQ+.¹³

While further research is needed to determine the strength of any causal relationship, and suicide and self-harm material has been found to have both harmful and protective effects, there is evidence that harmful effects currently predominate.¹⁴ The harmful effects of engaging with suicide, self-harm and intense depression content may include:

- Increases in the frequency and/or severity of suicide and self-harm ideation and behaviours. Arendt et al (2019) found that one third of participants in their study carried out the same or similar types of self-harm after observing it on Instagram.¹⁵
- Engagement behaviours such as sharing, liking or commenting on suicide and self-harm content may reinforce the creation and sharing of self-harm images, and in turn encourage further harmful behaviours;¹⁶ while engaging with material may result in a range of emotional, cognitive and physiological impacts, for example

¹¹ Susi, K et al (2023) Research Review: Viewing self-harm images on the Internet and social media platforms: systematic review of the impact and associated psychological mechanisms. *Journal of Child Psychology and Psychiatry*, 64(8). pp1115–1139

¹² MRF analysis, applying Rodway et al’s modelling to Office for National Statistics on death registrations in England and Wales

¹³ Rodway, C et al (2022) Online harms? Suicide related online experience: a UK wide case series study of young people who die by suicide. *Psychological Medicine*, 53 (10), pp1–12

¹⁴ Susi, K et al (2023) Research Review: Viewing self-harm images on the Internet and social media platforms: systematic review of the impact and associated psychological mechanisms. *Journal of Child Psychology and Psychiatry*, 64(8). Pp1115–1139

¹⁵ Arendt, F et al (2019) Effects of exposure to self-harm on social media: evidence from a two-way panel study among young adults. *New Media and Society*, 21, pp2422–2442

¹⁶ Ibid.

increased rumination, which may go on to trigger or exacerbate themes of despair, hopelessness and in some cases suicide and self-harm ideation;

- Algorithms and other engagement-based design features may result in an ‘assortative relating’ effect, with young people experiencing suicide ideation or thoughts of self-harm being more likely to identify and build relationships with other users experiencing similar actions and thoughts.¹⁷
- Online communities of practice may form, which in some cases may offer protective effects but that can also result in young people experiencing an exaggerated perception of suicide, self-harm and intense depression behaviours; a risk that self-harm is seen as normalised or beneficial coping mechanism; and a risk that these online cultures may inadvertently preclude offline or expert oriented forms of help seeking (cementing a sense that those who do not self-harm “cannot understand”).¹⁸

Question 3: Whether it is appropriate to make changes to the kinds of PPC and PC in the Act

Addressing the risk posed by content that may exacerbate the risk of cumulative harm

As it stands, the OSA and codes of practice do little to address the risks posed by seeing high quantities of ‘lower risk’ content that does not qualify as either PPC, PC, or NDC but that our research suggests is often being recommended alongside violative material.

As outlined above, a significant proportion of children are being exposed to PPC and NDC in combination with related but lesser forms of material. When consumed together, it is reasonably foreseeable that this will deepen the risks of and impacts associated with cumulative harm.¹⁹ For example, of the 13-17 year olds who had seen content encouraging suicide or self-harm in the last week, over a third (34%) had also seen *content about feeling sad, down or lonely* at least 10 times.

In the context of our findings, we would strongly urge the regulator **to explore opportunities to broaden the scope of its work and better target the risks posed by ‘lower risk’ content that can contribute to cumulative harm.**

¹⁷ Ibid.

¹⁸ See for example Lavis, A et al (2020) #Online harms or benefits? The graphic analysis of the positives and negatives of peer support around self-harm on social media. *Journal of Child Psychology and Psychiatry*, 61, pp842–854

¹⁹ Molly Rose Foundation (2025) Children’s exposure to suicide, self-harm, depression and eating disorder content on social media

In particular, measures targeted at high-risk algorithmic feeds and other engagement-based functionalities that serve content to children should not only require the removal or reduced prominence of PPC, PC, and NDC, but also **require platforms to identify and mitigate patterns of exposure involving repeated exposure to ‘lower risk’ content, or ‘lower risk’ content in combination with PPC, PC and NDC.**

Evidence of the widespread risk posed by ‘lower risk’ content also necessitates **a shift in how the regulator understands and monitors algorithmically driven risks.** In this context, we would encourage the regulator to explore whether tools like the Children’s Online Safety Tracker could encompass broader forms of content that the regulator itself recognises may contribute to cumulative harm.

Depression content and dangerous stunts and challenges content – upgrading the regime’s response

As it stands, Ofcom recommends that depression content (NDC) and content relating to dangerous stunts and challenges (PC) need only be ‘reduced in prominence’ on recommender feeds for certain age groups.²⁰ Proposals set out in the scope of the consultation on Additional Safety Measures also fail to include either of these content types in scope of potential proactive technology measures.²¹

MRF would strongly urge the regulator to reconsider the strength of measures targeted at these two content types, due to the elevated level of risk they pose to children.

As this response sets out, many children are being repeatedly exposed to large quantities of harmful depression content, which evidence shows can have devastating consequences. In Molly’s case, for example, she saw over 2000 items of harmful suicide, self-harm *and* depression content on Instagram in the months before her death. There is also a clear case for a stronger measures targeted at depression content given the close-interplay between this content and suicide and self-harm content, with it often being difficult to distinguish between their boundaries.

Content relating to dangerous stunts and challenges also poses an elevated level of risk to children. Ofcom’s Register of Risks, for example, acknowledges that such content can engender risks including ‘substantial injury, permanent harm, and in extreme cases death.’ The evidence cited also suggests that children do not necessarily need to see overwhelming quantities of this content in order to participate in a challenge.²²

In this context, we would urge the regulator to explore options to ‘level the playing field’ - either reconsidering the classification of depression content and content relating to

²⁰ Ofcom (2025) Protection of Children Code of Practice for user-to-user services

²¹ Ofcom (2025) Additional Safety Measures

²² Ortega-Barón et al (2022) Viral internet challenges scale in preadolescents: An exploratory study, *Current Psychology*, 42, pp.12530-12540

dangerous stunts and challenges, or to ensure measures in its codes of practice are better targeted to address the risks these content types pose.

Assessing codes of practice against their intended impact

This call for evidence spotlights the vital importance of measuring the incidence and impact of different kinds of harmful content in order to understand whether regulation is working.

In this context, we would strongly urge the regulator to clearly set out and then report against the **expected impact** of measures targeted at children on their exposure to harmful content, in order to better assess the efficacy of the online safety regime and where action needs to be strengthened.