

Written evidence submitted by the Molly Rose Foundation (SMR0002)

Response to the Science, Innovation and Technology Committee's call for evidence on social media restrictions

Summary

Molly Rose Foundation opposes calls for an Australian-style social media ban. Bans risk offering parents a false sense of safety, with early indications from Australia suggesting that most children's social media accounts still remain active.

Social media bans are blunt and simplistic responses that could trigger a range of unintended consequences. They fail to tackle the business models and commercial incentives that underpin high-risk, engagement-based harms. Furthermore, a ban will likely stall regulatory progress across other parts of the stack – including on gaming sites, messaging platforms and AI products - where chronic and acute harms continue to proliferate.

Parents' and young people's attitudes towards a social media ban

Parents are profoundly concerned about children's online safety, and overwhelmingly support further action to address product safety risks on social media. Forthcoming Molly Rose Foundation research finds that 91% of UK adults are concerned about children's online safety, with 91% worried about teenagers being algorithmically recommended suicide and self-harm content.¹

However, in respect of further online safety measures, it appears that most parents remain focused on the ends rather than a particular set of means. Our polling shows that three quarters (73%) of UK adults support new legislation to strengthen regulation and better protect children and young people from online harm, with support for regulation stronger than recent polling in favour of Australian style social media ban.²

Among those who support new legislation, the overwhelming majority (84%) want to see legislation introduced as soon as possible, ideally this year.³

Parents want to see action taken quickly, but also want to have confidence that the measures being adopted can succeed. When deciding on its next steps on online safety, equal proportions of UK adults support the UK Government actively considering the views of parents (82%) and evidence around children's online safety (82%). A significant proportion (69%) also support the views of young people actively being taken into account.

Recent polling of 70,000 UK schoolchildren found that an overwhelming proportion of young people opposed a social media ban, including 89% of secondary school pupils.⁴ However, opposition to a

¹ Polling of 2,048 adults undertaken by Savanta for Molly Rose Foundation, undertaken in January 2026

² 66% of adults supported an Australian-style social media ban, according to polling undertaken by the Good Growth Foundation in January 2026

³ Savanta polling for Molly Rose Foundation

⁴ Results of in-school polling undertaken in PSHE lessons. Conducted by Votes for Schools in January 2026.

social media ban should not be taken as evidence that children do not support further action. Young people themselves recognise that the status quo is entirely unacceptable, with clear support for stronger protections that can protect and promote their safety and wellbeing online.

Why an Australian-style social media ban is the wrong approach

Early evidence from Australia suggests that social media bans risk providing parents with a false sense of safety.

Though well-intentioned, bans risk doing more harm than good - causing harm to migrate to other high-risk platforms, leaving older teenagers at the risk of a 'cliff edge' when they turn sixteen, and risking a range of unintended mental health and wellbeing outcomes.

An Australian-style social media ban:

1. Gives parents an illusory sense of safety and comfort

In Australia, the early indications are that most children's social media accounts still remain active. Instagram claims it has deactivated 330,000 accounts, the equivalent of one account for every eight Australians aged between 8 and 15.⁵ Snapchat has removed the equivalent of one account for every six Australians in this age group.⁶

Prior to the ban coming into force, data from the eSafety Commissioner found that 95% of teens aged 13-15 were using at least one major social media site.⁷ Largely driven by malign compliance, it therefore appears that a majority of under-16s social media accounts remain either operational and/or still in use.

Some children have suggested that age assurance measures being adopted have been readily circumventable, and in some cases have enabled them to open entirely new accounts, including on major platforms.⁸

2. Creates a damaging cliff-edge at age 16

If properly enforced, a ban would introduce a deeply damaging cliff-edge for older teens - and particularly girls - who are suddenly exposed to poorly-regulated social media sites on their sixteenth birthday. Given a social media ban would likely have a chilling effect on regulatory outcomes, it is unlikely there would be adequate safety-by-design measures or protective guardrails in place to support them.

Given the current structural design of the Online Safety Act, we can reasonably assume that the child safety duties would continue to apply until girls turn 18. However, in the absence of targeted measures to address the underpinning business models and the immediate impacts of engagement-based design choices, we should be realistic about the likely efficacy of the resulting protections.

⁵ Data issued by Meta. According to the Office of the eSafety Commissioner, there are 2.5 million Australians aged between 8 and 15.

⁶ Data published by Snapchat.

⁷ Office of the eSafety Commissioner (2025) eSafety report shows widespread underage use of social media and minimal measures to prevent kids signing up. Published 20/02/25.

⁸ ABC (2026) Two months since the social media ban and teens say it isn't working. Published on 4th February 2026.

Social media companies are likely to argue that it is economically disproportionate to continue to offer safety-by-design guardrails solely to protect 16- and 17-year-olds. Given the regulator has consistently erred on the side of regulated platforms when undertaking proportionality assessments about codes of practice, we should reasonably assume these arguments would gain traction, and in turn could substantially constrain Ofcom's ambition and approach.

3. Means taking policy decisions ahead of the available evidence

We currently lack strong scientific evidence on the impact of bans, with no high-quality systematic study that has tested the impact of reducing or wholly eliminating social media use among healthy under-16s (or that has systematically evaluated the consequences).

While some supporters of a ban have invoked the precautionary principle to justify their position, the reality is that we simply cannot determine what the harm-to-benefit balance of such a drastic policy solution, however well-intentioned, may be.

In 12 to 18 months, we will be much better informed about the long-term impacts of the ban in Australia.⁹ We will also have the results of a large-scale randomised controlled trial in Bradford that will test a range of phone- and social media-based interventions.¹⁰

We strongly support the Government's intention to consult and gather evidence on a range of policy solutions, rather than considering the harm-to-benefit ratio of a ban in isolation.

In reaching its conclusions, the Government should appropriately consider both the evidence around the harms caused by social media, as well as the wide plurality of policy solutions that may address them.

MRF strongly welcomes the renewed policy focus on online safety improvements. However, we are concerned that the recent framing of this debate – typically positioned as a false binary between a ban or the status quo – has proven unnecessarily reductive and unhelpful.

4. Risks online harm migrating to other sites

Restricting children's access to certain social media platforms will not necessarily improve their safety and well-being, but may instead result in harms migrating to those platforms that children - and bad actors - are still readily able to use.

In Australia, high-risk gaming platforms like Discord and Roblox have been excluded from the ban, despite already being at the leading edge of new and emerging threats, including sadistically motivated group-based offending (perpetrated by so-called Com groups) that target vulnerable children and coerce them into acts of self-harm, sexual abuse and even suicide.¹¹

Although YouTube is included in the Australian ban, the platform remains accessible without a sign-in.

⁹ More than a dozen academic evaluations are currently underway.

¹⁰ 'The IRL Trial' will be run jointly by the University of Cambridge's MRC Cognition and Brain Sciences Unit and the Born in Bradford Initiative.

¹¹ Resolver, in partnership with Molly Rose Foundation (2026) Weaponised loneliness: critical harm intelligence briefing

Counterintuitively, this means children may be able to more readily encounter certain forms of harmful content, and can continue to use the platform but with fewer account protections than they previously enjoyed.

It is important to note that we have not yet seen evidence of a substantial migration by young people to platforms they didn't already use. Neither does the evidence suggest a large-scale uptick in VPN usage to circumvent the ban.

However, this may well be primarily driven by the limited impact of the social media ban to date - i.e. if young people remain largely able to use their existing accounts, and the network effects of social media platforms remain intact, there are insufficient incentives to migrate to other products.

5. Risks of unintended consequences, including for teens with vulnerabilities

Many young people rely on social media for connection, identity exploration and support. For LGBTQ or neurodiverse children, being online can offer huge benefits around identity, connection, self-esteem and peer-support.

In Australia, the country's youth mental health agency has reported that 10% of new referrals are related to the ban – despite the relatively small proportion of accounts that have actually been removed.¹²

Kids Helpline, Australia's equivalent of Childline, has reported levels of distress among children who have been cut off from their support networks, including children experiencing suicide ideation, teens who used social media to control self-harming behaviours, and numerous other mental health crises.¹³

Clearly any social media ban would offer some immediate protective effects, including the potential to prevent young people from being exposed to the cumulative effects of harmful content on the social media platforms in scope.¹⁴ However, these potential benefits need to be assessed alongside the potential unintended consequences and iatrogenic risks that may be caused by a ban.

MRF would again encourage a focus on the range of policy solutions that could offer protective effects against the product safety issues in scope, rather than considering any measure in isolation.

MRF is particularly concerned that a social media ban could make it much more difficult for children to disclose abuse and other negative experiences online. For example, if a child remains able to use a social media site, whether through malign compliance or as a result of user circumvention, they may be markedly less likely to disclose grooming or abuse that takes place on it (or equally to disclose the negative impacts of harmful content), for fear that they will be in trouble for having continued to access a prohibited service.

While everyone would recognise that poorly regulated social media is a problematic space, we are particularly concerned that a ban may lead some teenagers to access help and peer support from even

¹² Wilson, C (2025) One in 10 these seeking mental health support from Headspace cite social media ban as an issue. Published on Crikey.com.au 16/01/26

¹³ The West Australian (2026) 'Distressed' teens turn to Kids Helpline following social media ban, saying they've lost support networks. Published 7/01/26

¹⁴ Molly Rose Foundation research conducted shortly before the Online Safety Act came into effect found that almost half of girls (49%) had been exposed to suicide, self-harm, depression and eating disorder content in the last week.

more damaging online spaces, including high-risk suicide, self-harm and eating disorder forums.

Our analysis suggests that some young people who use a suicide forum linked to at least 135 UK deaths have found the forum after feeling failed by youth mental health services, or having been unable to find adequate peer support elsewhere.¹⁵

While social media is currently a deeply imperfect vehicle for peer support, the unintended consequences of a social media ban for young people experiencing suicidality, self-harm ideation and poor mental health could be substantial.

In the absence of social media platforms being available to them, the user pathways that lead young people to exceptionally high-risk, malign fringe forums may be shortened.

A better alternative to bans – a new online safety settlement

As set out above, MRF believe that an Australian-style social media ban risks giving parents a false sense of safety and will fail to offer the protection that children and parents deserve.

However, there is another way – a new online settlement for children and families. MRF has proposed a five-step plan that could deliver quick, effective and decisive action to tackle inherently preventable online harm.

At its core is a commitment to fix and strengthen the Online Safety Act – recognising that online safety risks are primarily a function of product safety failures, but that the regime cannot be expected to bite until it creates the regulatory and legal incentives commensurate to the size and scope of the tech companies in scope.

Ministers must:

1. Fix and decisively strengthen the Online Safety Act

Regulation remains the most powerful tool available to tackle preventable harm and the underlying commercial incentives and business models that drive inaction. However, regulation is not working as intended.

Amendments could be passed within months to introduce robust, risk-based minimum age limits. This would mean platforms with high-risk functionalities - including livestreaming and AI chatbots - carrying the highest age ratings, with tech firms finally being incentivised to engineer safer and more age-appropriate products if they wish to be available to younger teens. Government should also table a package of amendments that can fix existing weaknesses in the current Act, including measures to make companies act on the full set of harms in their risk assessments.¹⁶

Ultimately, a new Act is required. This must make provisions for a systemic Duty of Care, reset regulatory incentives in favour of harm reduction, and deliver an outcomes- and conduct-based regulatory approach that learns the lessons from financial services regulation, ensuring that the Act is better targeted to the size and cash-rich position of the largest companies in the world.

¹⁵ Molly Rose Foundation (2025) Missed Chances, Lost Lives: a substance and suicide forum cost lives and the State missed countless chances to act.

¹⁶ A full set of amendments have been prepared by the Online Safety Act Network. Online Safety Act Network (2026) Strengthening the Online Safety Act: a ten-point plan for Government

2. Extend the Act to cover child wellbeing, making wellbeing-by-design the price of admission to the UK market

The Government should expand the scope and ambition of the Act – widening its objectives to not only necessitate harm reduction but to actively deliver wellbeing-by-design. Through new duties on platforms, and new strategic objectives of Ofcom, we can introduce a robust and demanding set of measures that mark an end to addictive and harmful design choices, with a new duty to ensure digital products are built to be age appropriate, high-quality and nourishing by design.

Algorithms should be fundamentally repurposed – if a platform intends to use them, they should not only be free of harmful toxic material but should promote and support child development. Feeds should be made to recommend high quality, age-appropriate content from a diverse range of trusted sources, including trusted mental health practitioners, education providers and public service broadcasters.

3. Require new levels of transparency and accountability and candour from Big Tech

Transparency is a powerful lever for change – and one of the most impactful tools available to us to shape digital markets that can work for children, families and society. We therefore recommend:

- New disclosure duties that require the largest Category 1 firms – and their senior decision-makers – to proactively disclose information about which a regulator would expect to be made aware, mirroring successful requirements in place for financial services;
- A new approach to transparency across the stack, including for advertisers who unwittingly monetise social media harm, and whose advertising spend – at the core of Big Tech’s business model – is attritionally lost to opaque advertising models;¹⁷
- A longer-term push to promote new accountancy standards that can draw on lessons from tackling climate change. This would require corporates to report on their exposure to online safety risks, including in their supply chain.

4. Adopt a ‘polluter pays’ and whole-stack approach to harm reduction

The Government should adopt a ‘polluter pays’ funding approach, with an extension of the industry levy that currently funds Ofcom to also support civil society and academic research, pump-priming the evidence base that the regulator needs to act.

A new Act should bring app stores into scope and set out clear and consistent minimum standards for parental controls. Google and Apple should be legally required to adopt technically feasible device-level measures that can prevent children from being coerced into sending self-generated images. This would disrupt the rapid growth of sadistically motivated group-based grooming by so-called Com groups, which primarily penetrate harm on gaming and messaging sites (rather than on social networks.)¹⁸

¹⁷ MRF research found that adverts were placed adjacent to one in 9.5 harmful posts on TikTok’s For You Page and one in ten Instagram Reels. Molly Rose Foundation (2025) Pervasive by Design: suicide, self-harm and intense depression content on TikTok and Instagram, and how their algorithms recommend it to teens

¹⁸ Resolver, in partnership with Molly Rose Foundation (2026) Weaponised loneliness: critical harm intelligence briefing

5. Prioritise education as inoculation, with a new strategic focus on critical digital and media literacy

MRF supports a bold reset of school-age education, recognising that critical digital and media literacy is a foundational life skill that not only inoculates children against the worst effects of online harm today, but equips them with the skills they need to flourish in the AI and digital economy of tomorrow.

Building on England's Curriculum Review, we should invest in foundational skills around platform and algorithmic literacy, giving young people a critical understanding of how digital environments work and shape our lives – including the role of algorithms, data and persuasive design. The Government must also support schools to robustly embed critical digital and media literacy across the entire curriculum and all age groups, while training teachers to take responsibility for building children's digital life skills. Together with stronger regulation, high-quality education can help to turn the tide on online harm, draining the power from algorithms that continue to expose young people to high-risk content.

This also provides an opportunity to meaningfully invest in the brain capital and economic potential of generation, giving young people the critical thinking, digital citizenship and practical skills they need to thrive in an AI economy, to prepare for a new voting age of 16, and deal with the increasing threats that a fractured information ecosystem poses to our democracy and national security.

24 February 2026