

# Should the UK Ban Social Media for Under-16s?

## A CCOD Briefing on Evidence, Risks and Better Alternatives

### Executive summary

Online threats like disinformation are real and rising. The CCOD welcomes the UK government's commitment to regulating big tech and acknowledges the growing consensus that action is overdue. But a blanket social media ban for under-16s is a blunt tool that mistakes visible action for effective policy. The evidence does not support it. Such a policy risks serious unintended consequences as it shifts responsibility away from platform providers onto families, while failing to address the root causes of the issue, including the need to improve social media platform design. More crucially, the future is digital and excluding young people from social media does not protect them and instead leaves them ill-equipped to navigate a digital world they will inevitably inherit. If a ban is introduced, it needs to be accompanied by reforms that target the root causes of online harms: The UK should invest in national digital education, invest in safer platform design and domestic players providing solutions, and reserve targeted bans for non-compliant companies, which can be lifted if platforms comply.

### 1. The problem: serious online threats, but complex causes

Parents, teachers and young people themselves are rightly worried about what teenagers encounter online. Evidence links certain patterns of social media use to poorer mental health, reduced self-control and riskier behaviour in teenagers, but effects are often inconsistent.<sup>1,2,3,4</sup> What seems to matter most is **how frequently social media is used, what it is used for**, and the wider context of a young person's life.<sup>5,6</sup> These nuances are often left out in propositions for a blanket ban.

Social media can be an important tool for civic engagement, identity exploration and social connections, with teenagers benefiting from online communities and support that they cannot easily find offline.<sup>7,8</sup> Crucially, **social media has become a pivotal news source for teenagers**: In 2022, the social media platforms such as Instagram (29%) and TikTok (28%), for the first time, overtook BBC One and BBC Two (24%) as top news sources for 12-15 year-olds.<sup>9</sup> In 2024, **82% of 16-24 year-olds were using social media for news. Yet they were less likely to rate social media news as trustworthy (37%), accurate (37%) or**

---

<sup>1</sup> Burnell, K., Meter, D.J., Andrade, F.C., Slocum, A.N. and George, M.J. (2025) *SSM Mental Health*, 7, p.100459.

<sup>2</sup> Ansari, S., Iqbal, N., Azeem, A. and Danyal, K. (2024) *Cyberpsychology, Behavior, and Social Networking*, 27, pp.753–770.

<sup>3</sup> van der Linden, S. (2026) 'Social media bans for teens lack evidence', *Nature Health*. Available at: <https://doi.org/10.1038/s44360-026-00083-4>.

<sup>4</sup> Puukko, K., Hietajärvi, L., Maksniemi, E., Alho, K. and Salmela-Aro, K. (2020) 'Social media use and depressive symptoms: a longitudinal study from early to late adolescence', *International Journal of Environmental Research and Public Health*, 17(16), p.5921.

<sup>5</sup> Nuffield Foundation (2025) *Growing Up Online*. Available at: [https://www.nuffieldfoundation.org/wp-content/uploads/2025/10/DC3-Digital-lives\\_7.10.pdf](https://www.nuffieldfoundation.org/wp-content/uploads/2025/10/DC3-Digital-lives_7.10.pdf)

<sup>6</sup> OsloMet (2023) *Screen Quality Matters More Than Screen Time*. Available at: <https://www.oslomet.no/en/research/featured-research/screen-quality-matters-more-than-screen-time>

<sup>7</sup> Nuffield Foundation (2025) *Growing Up Online*.

<sup>8</sup> OsloMet (2023) *Screen Quality Matters More Than Screen Time*.

<sup>9</sup> Ofcom (2022) *News Consumption in the UK 2022 Report*. Available at: <https://www.ofcom.org.uk/siteassets/resources/documents/research-and-data/tv-radio-and-on-demand-research/tv-research/news/news-consumption-2022/news-consumption-in-the-uk-2022-report.pdf>

**impartial** (42%) than older age groups, demonstrating that heavy reliance on social media does not translate into greater trust in the information encountered on these platforms.<sup>10</sup> A study by Ipsos and JOE Media further showed that despite social media becoming the main news source for British 16-34 year-olds, this age group is in fact **more concerned about misinformation than the general public** (79% vs 72%), particularly compared to those aged 55 to 75 (67%).<sup>11</sup>

This points to the need for information integrity and digital skills. Rather than punishing teenagers and young adults for finding a form of news consumption that appeals to them, the CCOD believes that regulators should **target the root cause of the problem by removing incentives for hosting false content, especially if perpetrated by bot farms and hostile state actors, and penalise platforms if they keep hosting content that is false and misleading.**

## 2. Why a blanket ban is the wrong tool

A complete, age-based ban is neither the best, nor the safest response to addressing the rising threat of online disinformation.

### Limited evidence on blanket-ban effectiveness

We simply do not know whether a blanket ban would solve issues surrounding false information and online harms on social media. Australia, the first country to implement a social media ban for under-16-year-olds, has announced a formal evaluation that will collect data on 4,000 Australian teens over the next two years.<sup>12</sup>

Early signals already show that the ban might not have had the desired effects. Molly Rose Foundation polling has found 61% of Australian 12–15-year-olds retained social media access. Platforms barely tried to enforce the ban: 60–64% of continuing users said no action was taken on their accounts, and only 4% said they had to use VPNs. The same data shows 14% of Australian children felt less safe after the ban, with 51% saying it has made no difference.<sup>13</sup> In the UK, the result of the ban could be similar.

### Trade-offs between effectiveness, privacy, and technical feasibility

Implementing a blanket ban would require platforms to introduce age verification for both new and existing users. There's no clean way of doing this. **Every technical approach forces trade-offs between three things:** how **effective** it is at actually keeping under-16s off the platform, how much **sensitive user data** needs to be collected and stored, and **how easy it is to implement for both platforms and users.**<sup>14</sup> Improving one typically makes the others worse.

---

<sup>10</sup> Ofcom (2024) *News Consumption in the UK 2024 Report*. Available at:

<https://www.ofcom.org.uk/siteassets/resources/documents/research-and-data/tv-radio-and-on-demand-research/tv-research/news/news-consumption-2024/news-consumption-in-the-uk-2024-report.pdf>

<sup>11</sup> Ipsos (2025) *Modern Masculinities: Media Consumption*. Available at:

[https://www.ipsos.com/sites/default/files/ct/news/documents/2025-10/Modern%20Masculinity%20-%20Media%20Consumption\\_Charts\\_2.pdf](https://www.ipsos.com/sites/default/files/ct/news/documents/2025-10/Modern%20Masculinity%20-%20Media%20Consumption_Charts_2.pdf)

<sup>12</sup> The Guardian (2026) 'Three months into Australia's world-first social media ban for under-16s, has it been a success?'. Available at: <https://www.theguardian.com/australia-news/2026/mar/04/australia-social-media-ban-under-16s-three-month-review>

<sup>13</sup> Molly Rose Foundation (2026) *Australia's Social Media Ban: Is It Working?* Available at: [https://mollyrosefoundation.org/wp-content/uploads/2026/04/MRF\\_Australia-Social-Media-Ban-Research\\_Briefing-April-26.pdf](https://mollyrosefoundation.org/wp-content/uploads/2026/04/MRF_Australia-Social-Media-Ban-Research_Briefing-April-26.pdf)

<sup>14</sup> Leibniz Institute for Media Research (2026) *Why Banning Social Media Is Not the Solution*. Available at: <https://leibniz-hbi.de/en/hbi-news/news/why-banning-social-media-is-not-the-solution/>

A large-scale assessment on safeguarding of children policies in Australia found there is “no single reliable solution”, which is why different approaches have emerged in practice following the ban.<sup>15</sup> One option is ID-based verification through third-party providers, similar to opening a bank account. This can create new privacy risks, as data protection responsibility gets pushed onto external providers, many of whom operate across jurisdictions and do not store data where the platform itself is based. Storing sensitive data has also made these providers a target for cybercriminals, as seen in 2025, when hackers targeted Discord’s identity verification provider and 70,000 user IDs were leaked.<sup>16</sup> The alternative is to integrate AI and biometric solutions directly into the platform, where AI scans users’ faces and determines their age accordingly. This is usually cheaper and smoother to implement, but raises a different set of privacy problems and opens up questions about acceptable error margins and data bias.

Any policy approach needs to confront these trade-offs upfront across feasibility, privacy, and fairness.

### **A ban does not address the problem, it postpones it**

Introducing an age-based ban **does not address the “root incentive structure”** that drives harm through social media in the first place.<sup>17</sup> A blanket ban creates no incentives to improve the quality and safety of social media platforms, nor does it introduce safeguards for information integrity. The **volume of harmful content on social media platforms remains unchanged**. This also means that once children reach the permitted age threshold of 16, they will encounter the same if not worse risks the ban sought to mitigate, since online risks are constantly changing in nature and effectiveness due to AI. The only difference will be that once these teenagers will be allowed back on the platforms, they will do so without having developed the skills necessary to navigate its risks.

### **A ban shifts responsibility of youth digital engagement to parents and teachers**

Proponents argue that even an imperfect ban sends a valuable cultural signal that platforms are something to be cautious about. However, with the strong cultural signal also comes a **shift of responsibility** away from platforms to parents, teachers, and other adults with oversight functions. The problem is no age group is immune to the online threats of social media platforms.<sup>18</sup>

**Adults are also vulnerable to online harms**, subject to the same algorithmic manipulation and dopamine-driven engagement loops that the research consistently shows override individual willpower regardless of age or media literacy. Since the enforcement of the social media ban inevitably relies on parental oversight, this becomes a problem, since adults are also ill-equipped to serve as guides. A safety-by-design approach that restructures platforms for *all* users is therefore a far more structurally honest response.

### **A democratic “cliff edge” at 16**

A blanket social media ban up to the age of 16 risks creating a democratic “cliff edge” at precisely the moment young people are expected to assume civic responsibility. With teenagers already relying on

---

<sup>15</sup> Age Assurance Technology Trial (2026) *Age Assurance Technology Trial: Report*. Available at: <https://ageassurance.com.au/report/>

<sup>16</sup> BBC (2026), ‘ID Photos of 70,000 users may have been leaked, Discord says’. Available at: <https://www.bbc.com/news/articles/c8jmzd972leo>

<sup>17</sup> Stevens, M.W.R., Radunz, M., Király, O., Griffiths, M.D., Delfabbro, P., Demetrovics, Z., Czakó, A. and King, D.L. (2026) ‘Beyond the debate: toward pragmatic evaluation of Australia’s social media age restrictions’, *Lancet Regional Health – Western Pacific*, 68, p.101818.

<sup>18</sup> Yao, Y. (2025) ‘Regulating addictive algorithms and designs: protecting older adults from digital exploitation beyond a youth-centric approach’, *Frontiers in Psychology*, 16, p.1579604.

social media as their primary source of news and political information, removing this channel risks unintended and negative consequences for how informed and prepared they become as citizens.

The UK is moving towards lowering the voting age to 16. In a context where political trust and youth voter turnout are already under strain, a blanket ban would mean that young people encounter the full, unfiltered social media information environment for the first time at the exact moment they are expected to participate in democratic decision-making. Unlike generations before who developed digital instincts over time, young people subject to a ban would arrive at 16 with no prior experience of differentiating coordinated disinformation from credible content, leaving them less prepared, less media-literate and more vulnerable to misinformation and manipulation.

The project is exacerbated by structural inequalities in access to information. For young people who are not in education, employment, or training (NEET) — of which there are almost a million, with the figure projected to rise<sup>19</sup> — social media may constitute one of the few accessible channels for staying informed about current affairs. Restricting access without providing viable alternatives risks deepening informational exclusion among already marginalised groups.

Put together, these dynamics suggest that such a “cliff edge” approach may lead either to poorly informed participation or to disengagement altogether.

### **Likely to be circumvented and unequal in practice**

Teenagers can reasonably be expected to find ways around a ban, whether through VPNs, older friends' accounts or less regulated platforms. Beyond bypassing the traditional routes, evidence suggests that an age-based ban can trigger psychological reactance, driving more covert and risky behaviour.<sup>20</sup> A complete ban may also create a false sense of security among parents who assume their children are no longer on social media, when in reality their activity has just become harder to see and monitor. Meanwhile those with more resources will find it easier to circumvent restrictions, while more disadvantaged young people risk losing access to positive online spaces and communities. A blanket ban, therefore, risks deepening existing inequalities rather than closing them.

In parallel, proposed bans usually cover only certain social media platforms. They do not cover other technologies teenagers are already using, such as encrypted messaging apps, online games and AI chatbots, some of which have been linked to serious harms and threats to information integrity. It is neither realistic nor desirable to attempt to ban every new digital technology that emerges.

## **3. A better path: prepare, don't prohibit**

CCOD believes that the most effective and proportionate way to protect young people from online harm is not prohibition but reform. We propose a three-pillar strategy built on education, safer and accountable platform design and conditional enforcement.

### **Pillar 1: National digital education**

Critical reasoning and the ability to make sense of information are skills that develop through childhood, and children as young as four are already developing judgment of what is true and false. . Research shows

---

<sup>19</sup> BBC (2026) 'UK at risk of "lost generation" of young people without work, report warns'. Available at: <https://www.bbc.com/news/articles/cy026x9jpd0o>

<sup>20</sup> Köhler-Dauner, F., Peter, L., Sitarski, E., Chauviré-Geib, K., Haag, A.C. and Fegert, J.M. (2025) 'Digital child protection in social networks: age verification and age-tiered regulation in Europe', *Child and Adolescent Psychiatry and Mental Health*, 19(1), p.143.

that digital media education in schools can improve young people's ability to spot misinformation and navigate online spaces more safely.<sup>21</sup> And the earlier it begins, the more effective it is.

Measures could be funded through a levy on the UK-based revenues of major technology platforms. Given the significant role these companies play in shaping young people's online environments, part of their revenues should be reinvested into education-based initiatives that equip users with the skills needed to navigate digital services safely and critically. The levy could be modelled on the voluntary levy introduced for gambling companies, which was later made statutory.

The CCOD recommends:

- **A national, evidence-based digital literacy curriculum beginning in early primary school and continuing through secondary education.** The curriculum should cover privacy and data management, recognising disinformation and understanding the commercial and political intent behind it, and educating young people on their rights and how to exercise them online. Finland offers a strong case study: its media literacy programme is embedded in the curriculum starting from primary school across all subjects and languages, integrated into the curriculum rather than treated as a standalone module. This is believed to be a contributing factor to Finland ranking high in the EU on resistance to disinformation.<sup>22</sup> Based on this, the CCOD believes that the delivery of this programme should bring in external providers rather than placing additional demands on inadequately skilled and resource-short teachers, and media literacy should be integrated into existing subjects rather than treated as a cosmetic add-on to the school day.
- **A gradual, supervised exposure to digital technologies,** instead of a complete blackout to help develop relevant skills and resilience. There is little evidence on how best to teach children to use social media safely and responsibly. Closing this gap should itself be a policy priority. In practice, a staggered model of digital exposure offers a satisfactory framework. In early primary years, it might begin with supervised and timed use of tablets/screens for educational purposes. In later years, digital aids and tools can be introduced to support reading and writing. By secondary school, young students could be introduced to a structured, school-based social media environment, where pupils can chat with friends, learn to communicate online while being taught the implications of privacy, impact of posting, and the real-world consequences of sharing and commenting. The aim would be to instil a sense of responsibility and accountability: *what impact does it have when I like and share content?*

Enabling young people to navigate the digital world takes them seriously as future citizens and equips them to navigate both the challenges and the advantages of the digital world. Enabling them to problem-solve will more sustainably prepare them for new and changing threats in the digital space.

## ***Pillar 2: Shift responsibility to platforms and reward tech solutions that add value***

Placing the burden of youth digital engagement on parents and teachers is neither fair nor an effective way to protect teenagers from online harm. Instead, social media companies can be incentivised to design and operate systems that are optimised for the well-being and safety of all users, while not threatening revenue

---

<sup>21</sup> Li, C., Zhang, M. and Xie, K. (2025) *Information, Communication and Society*. Available at: <https://doi.org/10.1080/1369118X.2025.2584142>

<sup>22</sup> World Economic Forum (2025) *Rethinking Media Literacy: A New Ecosystem Model*. Available at: <https://www.weforum.org/publications/rethinking-media-literacy-a-new-ecosystem-model-for-information-integrity/>

models or stalling digital innovation. The CCOD recommends that government and regulators incentivise platforms to adopt the following measures to secure a healthier digital information ecosystem:

- **Mandated user empowerment features.** Amendments to the OSA and Communications Act should be introduced to place a duty on social media platforms to create simple, user-friendly controls that allow users to adjust how particular types of content appear in their social media feeds (e.g. give users the option to filter out AI-generated content or filter for factual accuracy of posts). These ‘user empowerment’ toggles in filtered feeds would encourage and teach *all* users, not just teens, to take greater control over their social media diets.
- **Tackling fake accounts and bots.** Fake social media accounts play an outsized role in disinformation online. Since bot accounts breach the terms of service of platforms, they need to be put under pressure to comply with their own rules and continuously monitor for bot-driven traffic and material.
- **Tackling disinformation and fostering information integrity.** This can be tackled on three interrelated fronts:
  - (1) Introducing enforceable **transparency rules for AI-generated content**. This could take the form of **mandatory watermarking labels** that allow both users and platforms to identify synthetic content through machine-readable markers. The UK has not yet adopted a comprehensive, economy-wide AI content labelling practice. Existing technical approaches, such as Adobe’s C2PA standard, demonstrate how such systems could function in practice, although C2PA itself is neither the only possible framework nor a legal standard. Policy approaches should align the UK more closely with the EU and California, both of which have moved toward disclosure and provenance requirements for certain synthetic content, though not for every form of AI-generated output.
  - (2) **Supporting innovation and promoting home-grown tech solutions** that target harmful online content, including AI-assisted systems for identifying fake news and coordinated disinformation. Such tools would have to be independently tested and carefully governed.
  - (3) Mandating regulators to demand that platforms operating in the UK meet **standards for safety and information integrity**, while penalising those who fail to follow.
- **Algorithmic design and news prominence.** Platforms should review the way their algorithms amplify content during a public order crisis to implement safeguards that diminish the virulence and speed at which false content spreads, potentially avoiding a situation like Southport or Epsom. One way to achieve this would be to develop an additional toggle or filter system that would give prominence to Public Service Broadcasters (PSBs) or other sources already governed by the Ofcom Broadcasting Code. In a crisis scenario, this setting could automatically prioritise PSBs within users’ algorithms and social media feeds, effectively slowing down the spread of potentially dangerous disinformation. In this way, young people’s interest in online news consumption is not harmed.

### ***Pillar 3: Conditional, targeted use of bans***

Bans should remain available as a policy tool, but as a last resort and targeted at non-compliant services, not at an entire age group. The CCOD suggests:

- Defining clear, evidence-based **tests that platforms must meet on safety and information integrity**. The ban would be listed if they meet the tests and implement the above recommendations.
- Non-compliant platforms should face **temporary or partial access restrictions until measurable safety standards are met**. An incentive-based framework should run in parallel, allowing platforms to earn reputational recognition and competitive advantage by submitting to independent stress tests of their safety measures. Restriction should be lifted as a direct reward for demonstrated compliance.

This approach turns bans into an enforcement mechanism to change company behaviour, rather than a blunt restriction on young people's autonomy.

#### 4. Recommendations for the government

In light of the current debate and the national consultation on a social media ban for under-16s, CCOD recommends that the UK government:

1. **Reject a blanket under-16 social media ban** as currently proposed, on the grounds of inconclusive evidence of effectiveness, significant unintended risks, and inconsistency with children's rights to information and democratic participation.
2. **Should the government proceed regardless, the ban must be accompanied by the following measures:**
  - Fund a national digital literacy programme that builds critical thinking and digital resilience from primary school onwards, delivered by experts in the topic in an integrated and gradual manner.
  - Place a legal duty on platforms to design safer products, including user empowerment tools, bot and disinformation safeguards, mandatory AI content transparency, and crisis-response protocols.
  - Invest in domestic technology solutions that detect and address harmful content at scale.
  - Reserve bans as a last-resort enforcement tool against platforms that persistently fail to meet safety standards.

#### Closing statement

The UK faces a real and urgent challenge in protecting young people from online harms and threats to information integrity. But protecting teenagers should not mean excluding them from the digital public sphere. A blanket ban is an easy option that would make visible noise and gain political attention; however, at the core, it is insufficient and is likely to do more harm than good. The harder option, but the one likely to work, includes fixing the platforms, designing and funding an integrated curriculum, and supporting technologies built with care and accountability. A strategy that prepares, educates and empowers young people, while holding technology companies to account, is more effective, more proportionate and more consistent with the values we hope to instil in the next generation: that evidence matters, rights matter, and their voices matter.

## About us

The CCOD is a mission-led organisation fighting against online disinformation. We have two primary goals that define who we are, what we do, and what we stand for:

1) We aim to influence the policy and regulatory environment to ensure that the UK government effectively tackles the growing issue of online misinformation and disinformation. CCOD has helped convene a cross-party group of parliamentarians and works directly with policymakers to improve understanding of disinformation threats. We educate policymakers on disinformation challenges and work with them on workable policy, regulatory and legislative initiatives to protect online information integrity in the UK.

2) Our objective is to bring together the key actors working across the information integrity ecosystem, ranging from technology developers and researchers to civil society organisations and media specialists, and ensure that those at the forefront of developing innovative solutions to online disinformation are given a prominent voice in shaping public policy outcomes. In parallel, we support organisations in protecting themselves and their workforces from disinformation risks with our advisory services, helping to safeguard markets, economic stability, and public trust.

Our collaborators and partners have included RUSI, the Aspen Institute, the Alan Turing Institute, Adobe, Egregious, Refute, Ghost Shift, and academics from the University of Cambridge, UCL, and LSE. Our work has been featured on Times Radio, Politico and PoliticsHome.

Our research and policy thinking is supported by a world-class advisory board, which includes, among others, John Penrose (former MP and Government Minister); Professor Sander van der Linden (Professor of Social Psychology at the University of Cambridge); Matthew Wicker (Assistant Professor at Imperial College and Research Associate at the Alan Turing Institute); Joseph Brandifino (Leads Policy Strategy at Scale AI); Naomi Krieger Carmy (VP Portfolio Management at the DeepTech VC fund Earth & Beyond Ventures); Clare Rewcastle Brown (award-winning investigative journalist); and a former Chair of Ofcom.

© 2026 Council for Countering Online Disinformation

Authors: Barbara Waldner

Editors: Sakshee Singh, Giovanna Custódia, Thomas Barton

Website: <https://ccod.co.uk/>

