5RIGHTS FOUNDATION

Growing up in the Online World: a national consultation

An overview of 5Rights' response

5Rights Foundation is an international NGO working with and for children to build a digital world that honours their rights as set out in the UN Convention on the Rights of the Child. 5Rights conducts groundbreaking research through our joint centre with the London School of Economics (Digital Futures for Children Centre), spearheads age-appropriate design and child online safety legislation across the world and translates advocacy work into practical service design reforms, by partnering with professional associations. To date, our work has inspired scores of concrete design changes, making services safer for millions of children around the globe. For more information visit www.5rightsfoundation.com.

Chapter 1 – Understanding how Children use technology

1. A rights-based approach to understanding the benefits and risks of social media for children

The question of whether social media is "beneficial" or "harmful" for children is too narrow and ultimately the wrong framing. It imposes a false binary on what is, in reality, a complex and integrated part of children's lives. Social media is not an optional add-on to childhood today; it is embedded in how children socialise, communicate, and participate in the world around them. Children themselves are clear on this.[1]

These online spaces are neither wholly positive nor wholly negative. They are simply environments where life happens. This is reflected in the concept of "platform paradoxes" identified in Steve Wood's recent work for the Digital Futures for Children research centre, and in findings from Internet Matters' 2025 reporting on children's digital lives. Children's online experiences frequently involve connection, creativity, entertainment and participation alongside exposure to pressure, exclusion, harmful and age-inappropriate content, and exploitative design practices.[2]

[1] The Children's Society, *The Good Childhood Report 2025; Internet Matters, Children's Wellbeing in a Digital World Index Report*, 2025; Internet Matters (2025) *Children's wellbeing in a digital world index report 2025*.

[2] Wood, S. (2026) *Impact of regulation on children's digital lives: Phase 2*. Digital Futures for Children, LSE; Internet Matters (2025) *Children's wellbeing in a digital world index report 2025*.

Crucially, children do not meaningfully distinguish between "online" and "offline" life. For them, it is a single, continuous experience. Digital spaces must therefore be understood as real environments, not secondary or separate ones.[3]

It is also important not to overstate social media as the root cause of challenges in children's lives. Evidence consistently shows that children's primary concerns relate to education, future employment, financial security, and wider global issues such as climate change and conflict, as highlighted by the Children's Society in "*The Good Childhood Report 2025*".[4] Digital experiences are one factor among many influencing wellbeing, and overly simplistic causal narratives risk obscuring this broader context.

A more productive approach is to centre children's rights. Children do not have a right to access any specific digital service or product. However, they do have rights to participate in social and cultural life (Article 31, UN Convention on the Rights of the Child (UNCRC)), which today includes digital spaces.[5] They also have rights to access information (Article 17), express themselves (Article 13), and be heard in decisions that affect them (Article 12), as recognised under the UN Convention on the Rights of the Child (UNCRC) and General Comment No. 25 on children's rights in relation to the digital environment.[6]

Listening to children is essential. They are experts in their own experiences and provide critical insight into how digital services function in practice, what supports them, and what undermines their wellbeing.

At the same time, participation must be balanced with the responsibility of adults and institutions to act in children's best interests. Popularity or engagement alone cannot determine what is appropriate. This requires careful consideration of inherent trade-offs, including between privacy and safety, and autonomy and protection. These decisions cannot be left to default service design or commercial incentives; they require deliberate, principled decision-making.

It is also important to recognise that children are not a homogeneous group. A rights-based approach requires acknowledging that children experience digital environments differently depending on age, maturity, vulnerability and context, consistent with General Comment No. 25 and 5Rights' *Digital Childhood* report

[3] 5Rights Foundation (2023) *Digital childhood: addressing childhood development milestones in the digital environment*.

[4] The Children's Society (2025) *The Good Childhood Report 2025.*

[5] UN Convention on the Rights of the Child.

[6] UN Convention on the Rights of the Child; UN Committee on the Rights of the Child (2021) *General comment No. 25 (2021) on children's rights in relation to the digital environment; UN Convention on the Rights of the Child.*

(2023).[7] Regulatory and policy approaches must therefore move beyond narrow "harms" framings and align more closely with children's developmental needs and evolving capacities.

This has direct implications for responsibility and design.

The principle of safety by design is already well established across other sectors: toys, food, medicines, vehicles and consumer products are expected to meet safety standards before reaching children. Responsibility sits with those designing, deploying and profiting from products to identify and mitigate foreseeable risks from the outset.

The same expectation should apply to digital services. Children should not be left to manage risks created by service design, nor should responsibility rest primarily on parents.

Therefore, the focus should shift from restricting access to improving environments.

A child rights-by-design approach provides a practical framework for achieving this, ensuring that safety, privacy, transparency and agency are built into services from the outset. Safety-by-design is a core component of this approach, ensuring foreseeable harms are prevented through design rather than addressed only after they occur.

This includes high privacy settings by default, age-appropriate design and clear terms, minimising data collection, reducing profiling and engagement optimisation practices that are not in children's best interests, meaningful reporting and redress mechanisms, robust risk assessments, and designing for wellbeing rather than maximising engagement.[8]

To support effective implementation of safety-by-design, 5Rights Foundation, the Online Safety Act Network and civil society partners launched the Safety by Design Code of Practice in May 2026.[9] The Code sets out a framework for embedding safety, rights and wellbeing into digital services from the earliest stages of product and system design, including governance structures, recommender systems, data practices, engagement features and accountability mechanisms. It demonstrates that safety-by-design is not an abstract principle but an operational and achievable approach that can be embedded across digital services in practice.

---

[7] UN Committee on the Rights of the Child (2021) *General comment No. 25 (2021) on children's rights in relation to the digital environment*; *5Rights Foundation (2023) Digital childhood: addressing childhood development milestones in the digital environment*.

[8] 5Rights Foundation (2019) *Designing for children: best practice in the Age Appropriate Design Code*; *Online Safety Act 2023, s.1(3)*.

[9] Online Safety Act Network, 5Rights Foundation and others (2026) Safety By Design Code of Practice.

Ultimately, social media presents both benefits and risks for children, which often occur simultaneously within the same platforms. It is not meaningful to treat these as opposing outcomes to be weighed against one another in a binary way.

The more important question is how digital environments are designed, governed and regulated to ensure that children's rights are respected, risks are addressed at source, and beneficial aspects of participation are supported in practice.

Chapter 2 – Interventions for safer, more positive experiences

Restricting social media services by age

1. Minimum age of access requirements

5Rights supports a legal requirement for social media services to have a minimum age of access, however, we disagree that "social media services should have a minimum age of access of at least 16 and should not be accessible to any children under that age". We would instead support a legal requirement for social media services to have a minimum age of access of 13.

The rationale for setting a minimum age of 13 is grounded in a precautionary, rights-based approach that recognises childhood as a period of evolving capacity and vulnerability. It does not assume that harm begins or ends at a single threshold, nor that digital services suddenly become "safe" at a particular age. Rather, the purpose of establishing a minimum age is to create a clear and enforceable baseline that protects younger children from the highest risk features of the digital environment, particularly those driven by data collection, profiling, behavioural targeting, and algorithmic systems. Certain forms of risk are simply not developmentally appropriate for younger children, and children's ability to navigate those risks develops progressively and unevenly over time.

Children's rights apply equally in digital environments. These include the right to safety and protection, the right to privacy and data protection, the right to understand how their data is used, and the right to participate in social and cultural life in ways appropriate to their age and development.[10] Where children are least able to exercise these rights independently, systems must provide stronger protections by design.

[10] UN Committee on the Rights of the Child (2021) *General comment No. 25 (2021) on children's rights in relation to the digital environment*; UN Convention on the Rights of the Child.

Evidence from 5Rights' Digital Childhood research is clear that childhood is a process of development, not a fixed state.[11] Children's cognitive, emotional, and social capacities evolve gradually. Importantly, not only are children not adults, but children of different ages have different levels of maturity, understanding, and capacity.[12]

For children under 13, this has specific implications. They are less able to understand how digital services shape their behaviour, less able to assess longer-term consequences, and less likely to critically interpret how their data is collected and used.[13] They are also more likely to accept content and interaction at face value, and more susceptible to persuasive design features and systems optimised to capture and retain attention.[14]

At the same time, these systems are not neutral. Many digital services are intentionally designed to maximise engagement through profiling, recommendation systems, behavioural nudging, and attention-optimised design.[15] This creates a structural imbalance between younger children and the digital environments they are expected to navigate. One that cannot realistically be addressed through individual choice or parental oversight alone.

The use of 13 as a minimum age reflects this developmental reality, as well as established regulatory and industry practice. It provides a clear and widely understood baseline. However, it is not sufficient in isolation.

Turning 13 does not remove risk; rather, it marks a transition point within childhood. Evidence from the *Impact of Regulation on Children's Digital Lives: Phase 2* (Steve Wood, 2026) highlights that regulatory interventions interact differently with children's experiences across age groups, underlining the importance of differentiated approaches throughout childhood and adolescence.[16] Risks persist even where protections increase, and outcomes vary depending on implementation, design choices, and a child's developmental stage. Adolescence therefore remains a period of significant vulnerability, particularly in relation to identity formation, social comparison, autonomy, peer influence, and engagement-driven platform dynamics.[17]

From this perspective, two policy positions follow:

[11] 5Rights Foundation (2023) *Digital childhood: addressing childhood development milestones in the digital environment*.

[12] 5Rights Foundation (2023) *Digital childhood: addressing childhood development milestones in the digital environment*.

[13] 5Rights Foundation (2021) *Pathways: how digital design puts children at risk*.

[14] 5Rights Foundation (2021) *Pathways: how digital design puts children at risk*.

[15] 5Rights Foundation (2023) *Disrupted childhood: the cost of persuasive design*.

[16] Wood, S. (2026) *Impact of regulation on children's digital lives: Phase 2*. Digital Futures for Children, LSE.

[17] Wood, S. (2026) *Impact of regulation on children's digital lives: Phase 2*. Digital Futures for Children, LSE.

First, there should be a strict prohibition on personalised, data-driven services for children under 13. This includes behavioural profiling, targeted content, and algorithmic recommendation systems. These features are not compatible with younger children's developmental stage or their rights, given the structural imbalance they create.[18]

Second, a pre-certification regime should be introduced for services accessed by 13–18-year-olds. Services should be required to demonstrate compliance with clear standards for safety, privacy, transparency, and age-appropriate design before being permitted to serve this age group. Certification should operate across tiered age bands (e.g. age 13-15 and 16 – 18), reflecting the reality that children and young people of different ages have different capacities, sensitivities, and levels of resilience.

This approach places responsibility where it belongs. Children have the right to participate in digital life. It is the responsibility of those who design, operate, and regulate these environments to ensure that such participation is safe, fair, and consistent with their rights from the outset.

2. The impacts of having a minimum age requirement higher than 13 for social media services

Our position is to establish 13 as a minimum baseline, combined with a robust system of age assurance and service certification up to 18. This creates a coherent framework that treats safety not as a question of access alone, but as a question of how digital services are designed, operated, and governed for children and young people.

The key issue is that raising the minimum age beyond 13 risks misdirecting attention towards entry thresholds rather than the conditions that actually shape children's experiences online. Age limits in isolation would also likely shift responsibility away from technology companies and onto parents, carers, and children themselves. In no other context would children be expected to manage risks created by pervasive and systemic commercial design.

Safety should not depend on a child's capacity to navigate complex digital systems, nor on families independently managing risks embedded in platform architecture, particularly given significant variation in digital literacy, time, and resources across households, which already contributes to unequal ability to manage online risk.

Focusing primarily on age thresholds risks encouraging highly visible but simplified interventions such as standalone bans or strict age limits. While these approaches can appear decisive, they risk displacing attention from

[18] *Data Protection Act 2018; Information Commissioner's Office (ICO) (2020) Age appropriate design: a code of practice for online services.*

the underlying commercial incentives and design choices that generate harm, creating the appearance of resolution without addressing system architecture.

Crucially, age thresholds alone do not change how services operate. They are also imperfectly implementable, meaning children may still access services but without the protections that should accompany them. This creates a false sense of safety while leaving the underlying drivers of harm unchanged.

It is important to note that those underlying harms are not incidental. They are produced through design. Features such as algorithmic recommendation systems, infinite scroll, engagement loops, profiling, and targeted content actively structure children's online experiences and can amplify exposure to harm.[19] The key determinant of risk is therefore not simply whether a child is present on a service, but how that service is designed and optimised.

As also reflected in Steve Wood's *Impact of Regulation on Children's Digital Lives: Phase 2* (2026), threshold-based approaches alone are insufficient because risks are embedded in service design and business models.[20] Recommendation systems, engagement optimisation, and profiling operate independently of age thresholds, meaning structural harms persist unless the underlying systems are addressed.

For this reason, the Government should prioritise a safety-by-design approach and make it a condition of services being available to child users. This directly addresses the system architecture that produces risk, rather than relying on exclusion through age thresholds. It embeds protections within product design itself, rather than depending primarily on avoidance, parental oversight, or after the fact mitigation. This also means that older teenagers who have a right to safety under the UNCRC until they are 18 are afforded the benefit of safer environments.[21]

Furthermore, a design-led approach is also more effective in improving privacy and data protection outcomes. It reduces unnecessary data collection and limits profiling and behavioural targeting, whereas raising age thresholds alone does not alter underlying business models or engagement incentives.[22]

This design-led approach is reflected in the Safety by Design Code of Practice developed by 5Rights Foundation alongside the Online Safety Act Network and civil society partners, which sets out practical expectations for embedding

[19] 5Rights Foundation (2021) *Pathways: how digital design puts children at risk*.

[20] Wood, S. (2026) *Impact of regulation on children's digital lives: Phase 2*. Digital Futures for Children, LSE.

[21] UN Convention on the Rights of the Child; UN Committee on the Rights of the Child (2021) *General comment No. 25 (2021) on children's rights in relation to the digital environment*.

[22] *Data Protection Act 2018*; Information Commissioner's Office (ICO) (2020) *Age appropriate design: a code of practice for online services*.

safety, privacy, accountability, and children's rights across product design, data practices, recommender systems, and platform governance.[23] This provides an example of the kind of clear regulatory framework needed to require services to demonstrate safety for children as a condition of access to the UK market.

Such an approach should be understood not as a barrier to innovation, but as a baseline expectation for responsible operation. It would incentivise the development of privacy-preserving, age-appropriate services while reducing ambiguity and the need for reactive redesign later in the product lifecycle.

The priority is a precautionary, system-level framework in which safety is not assessed after children are already exposed to services but instead embedded into the conditions of access from the outset. Within this, services would be required to demonstrate safety before reaching children, and new features would default to higher protections unless they can be shown to be safe.

Raising the minimum age beyond 13 does not, in itself, deliver safer outcomes. A more effective approach is to maintain a clear baseline while placing responsibility on companies to design safe services, ensuring children's safety and rights are built into systems as a precondition for access rather than treated as an afterthought.

Age of digital consent

1. Raising the age of digital consent in the UK for information society services

Raising the digital age of consent will not be effective or workable in isolation. Its impact depends entirely on whether it is embedded within a broader, enforceable system of age assurance, design standards, and regulatory oversight. Without this, it risks being largely symbolic.

Consent alone is not a reliable safeguard for children in digital environments. Children are not in a position to meaningfully understand the full extent of data processing across services, and parents are often unable to effectively oversee or verify consent at scale across multiple services and complex data practices. As a result, consent operates more as a formal compliance mechanism than an effective tool for control or safeguarding.

This means that raising the digital age of consent does not, by itself, address how children's data is actually processed or how risk is created. It risks focusing policy attention on access and permission rather than the underlying design of services and the systems that drive data collection and targeting.

[23] Online Safety Act Network, 5Rights Foundation and others (2026) Safety By Design Code of Practice.

The practical effect of increasing the digital age of consent is also limited. The GDPR provides multiple lawful bases for processing personal data, and organisations are not required to rely solely on consent.[24] The ICO also confirms that consent is only one of six lawful bases, while legitimate interests is widely used for a range of commercial and operational purposes, subject to a balancing test.[25]

In practice, services often rely on multiple lawful bases across different types of processing, meaning that changing consent thresholds alone does not necessarily reduce the scale or nature of data use. For example, Meta's Privacy Policy relies on "legitimate interests" under Article 6(1)(f) GDPR for certain processing, including profiling and personalised experiences across its products, such as understanding user interests and delivering personalised content and advertising.[26] Without structural change to underlying data practices, these activities can continue under different legal bases.

For these reasons, effectiveness depends far more on system-level measures than on the age threshold itself. A workable approach requires robust, privacy-preserving age assurance across services, applied consistently and proportionately to risk, and designed to avoid fragmentation or loopholes. This ensures protections are applied in practice, rather than relying on self-declared age or consent-based mechanisms.

Workability also depends on strong enforcement. The Age Appropriate Design Code (AADC)[27] provides an existing framework for requiring services to design with children in mind, but its effectiveness relies on regulators having sufficient capacity to monitor compliance in practice, audit how services operate, and take action where standards are not met. Without adequate regulatory resourcing and operability, even well-designed rules will not translate into consistent real-world outcomes.

We recommend placing the AADC on a statutory footing so that its requirements become legally binding rather than advisory. This would embed its expectations directly into enforceable data protection law, ensuring services likely to be accessed by children are required to design in their best interests as a primary consideration. It would also strengthen consistency across industry and give regulators clearer powers to assess compliance and take action where standards are not met.

In summary, raising the digital age of consent can only be effective if it sits within a wider enforceable system that prioritises age assurance,

[24] *Art. 6 GDPR* – Lawfulness of processing.

[25] Information Commissioner's Office (ICO) *Lawful basis for processing*; Information Commissioner's Office (ICO) *Legitimate interests*.

[26] Meta Platforms Inc. *Privacy Policy; Art. 6 GDPR* – Lawfulness of processing.

[27] Information Commissioner's Office (ICO) (2020) *Age appropriate design: a code of practice for online services*.

strengthens design-based regulation through the AADC, and ensures regulators have the resources and powers to secure compliance in practice.

On its own, it does not deliver meaningful protection for children.

2. Changing the digital age of consent for some services but not others

We do not agree that there is a case for changing the digital age of consent for some online services but not others.

The risks to children are not confined to a narrow category of providers and services, but arise across a wide range of digital services, including social media, online gaming, Education Technology (EdTech) platforms, and AI chatbots. These services increasingly share core characteristics such as large-scale data collection, profiling, algorithmic recommendation or influence, and behavioural design that shapes engagement. Because the underlying drivers of risk are consistent across services, there is no coherent basis for applying different age thresholds to different sectors.

A differentiated approach based on online services would also undermine regulatory coherence and international alignment. Children's online safety and data protection increasingly depend on shared expectations across borders.

Fragmenting the digital age of consent by service type would risk creating a patchwork of standards that is harder to understand, harder to enforce, and easier to circumvent. We support nationally and internationally aligned standards and collaboration on children's rights and online safety, rather than divergent thresholds that weaken consistency and reduce global effectiveness.

In practice, changing the digital age of consent for some services but not others is also unlikely to have a significant impact on children's exposure to services or risks. As mentioned in our previous response, UK GDPR provides multiple lawful bases for processing personal data, and organisations are not required to rely solely on consent.[28]

As the ICO sets out, consent is only one of six lawful bases, and legitimate interests is a widely used and flexible alternative that may be applied to a range of commercial and operational purposes, subject to a balancing test.[29] As mentioned previously, in practice, digital services often rely on a combination of lawful bases depending on the purpose of processing, including legitimate interests for aspects of analytics, service improvement and commercial activity, alongside consent where required.

Any changes to the age of consent will have limited impact without effective enforcement. Experience to date demonstrates that regulatory powers are not

[28] Article 6 GDPR.

[29] Information Commissioner's Office (ICO) *Lawful basis for processing*; Information Commissioner's Office (ICO) *Legitimate interests*.

actively and consistently used by the ICO. There has been very limited enforcement in this area, despite clear standards already being in place. This indicates a gap between the existence of regulatory requirements and their consistent application in practice. As a result, simply adjusting the age of consent will not deliver the intended outcomes in terms of children's safety or data protection.

To have meaningful impact, government should look closely at how data protection law is enforced in practice, including how existing provisions are applied and where gaps remain, and recognise that weak or inconsistent enforcement undermines even well-designed regulatory frameworks.

For these reasons, the priority should not be varying the digital age of consent across services but strengthening the operability and enforcement of existing protections.

Again, placing the AADC on a statutory footing in primary legislation would improve legal certainty, reinforce regulatory authority, and support more consistent and proactive enforcement, both nationally and in alignment with international best practice.

Restricting services based on 'risky' functionalities

1. Restricting specific functionalities based on minimum age thresholds

Minimum age thresholds can form *part* of a regulatory framework which meets children's rights, particularly where features present heightened risks to children's safety, wellbeing, or privacy. However, we do not find it appropriate for decisions about children's access to specific digital functionalities to be determined through broad consultation questions of this kind alone.

Questions about functionalities such as live-streaming, disappearing content, location sharing and talking to strangers, involve complex and evolving considerations relating to child development, children's rights, technical design, data practices, and evidence of risk and effectiveness. These are not static questions, and the risks associated with a feature can vary significantly depending on how it is designed, deployed, moderated, and monetised. We are concerned that approaching the issue through questions of this nature risks oversimplifying the discussion.

Crucially, it is not appropriate for the ability for children to send nude images or videos to be framed as a "functionality" within scope in the same way as other features. This item is illegal and inherently harmful, and any framing that normalises it as a design option risks undermining child protection principles rather than strengthening them.

Digital services and functionalities evolve rapidly, with new engagement mechanisms and AI-driven features being introduced continuously. At the

same time, the risks associated with features can vary significantly depending on how they are designed, deployed, moderated, and monetised. A framework based solely on predefined lists of functionalities and fixed age thresholds risks becoming quickly outdated and insufficiently responsive to technological and market developments.

We recommend establishing a statutory framework to deliver age-appropriate online experiences for children and young people, built around three core elements: a clear minimum baseline for under-13s, a mandatory pre-certification regime for services accessed by under-18s, and a strengthened enforcement framework.

As a minimum baseline, personalised and algorithmically driven services, including behavioural profiling and recommendation systems, should not be accessible to children under 13, with all in-scope services required to implement highly effective age assurance measures to enforce this restriction.

For users aged 13–18, access to digital services should be conditional on compliance with a mandatory pre-certification regime based on graduated age-appropriate design and safety standards. Services should be required to demonstrate compliance before being made available to children and young people, with differentiated standards applied across age bands such as 13+ and 15/16+.

A precautionary principle should underpin the regime, meaning that services, features, and significant product changes are treated as unavailable to under-18s unless and until they are independently assessed and certified as compliant.

The framework should be supported by an independent multidisciplinary expert panel responsible for developing and maintaining certification criteria, determining appropriate age thresholds for services and functionalities, and reviewing emerging technologies, AI-driven features, and risks on an ongoing basis.

Highly effective age assurance should form a core part of the regime and should be risk-based and proportionate, privacy-preserving and secure, effective against circumvention, accessible and inclusive, and transparent and accountable.[30]

Naturally, the framework must also be backed by strong enforcement. Regulators, including Ofcom and the Information Commissioner's Office, should work alongside the independent expert panel and be empowered and resourced to require evidence of compliance, investigate breaches, and impose meaningful sanctions, including business disruption measures, injunctive relief, and executive accountability where appropriate.

[30] Ofcom (2025) *Highly effective age assurance: Guidance on children's access and protection duties under the Online Safety Act.*

2. Persuasive design features and minimum age restrictions

We believe that persuasive design features such as infinite scrolling, autoplay, affirmation features, alerts and push notifications and content recommendation algorithms should be age restricted.

These features are, however, already within scope of UK data protection law and children's design regulation, including the Data Protection Act 2018, UK GDPR, and the Age Appropriate Design Code (AADC).[31]

The AADC establishes a coherent set of requirements that place the best interests of the child as a primary consideration, restrict the use of children's data up to the age of 18 in ways that are detrimental to their wellbeing, require high privacy settings by default, and ensure that design choices properly account for risks arising from profiling, engagement optimisation, and behavioural influence.

These provisions already govern how engagement-driven and personalised systems should be designed and deployed.

In practice, features such as infinite scroll, autoplay, social feedback mechanisms, push notifications, and algorithmically personalised alerts operate as core components of behavioural and engagement design. They are not neutral interface choices. Infinite scrolling and autoplay remove natural stopping cues and extend time spent on platform, while likes, comments, and other feedback loops embed social validation into children's online experience and shape identity formation.

Similarly, personalised notifications and alerts are designed to drive repeated re-engagement and can displace sleep, attention, and offline activity. As set out in *Disrupted Childhood (2023)*, children are operating within an attention economy in which services are deliberately designed to maximise sustained engagement.[32]

Recommendation systems sit at the centre of this ecosystem. They rely on large-scale profiling, inference, and personalisation and should be understood as core engagement infrastructure rather than neutral content delivery tools. Evidence from 5Rights Foundation's *Pathways: How digital design puts children at risk (2021)* shows how recommender systems, combined with frictionless design, can rapidly channel children towards harmful content pathways and reinforce prolonged engagement patterns. [33]

These risks are further amplified where profiling, personalisation, and engagement optimisation are embedded by default, as highlighted in *Designing*

---

[31] *Data Protection Act 2018*; *UK GDPR*; *Information Commissioner's Office (ICO) (2020) Age appropriate design: a code of practice for online services*.

[32] 5Rights Foundation (2023) *Disrupted childhood: the cost of persuasive design*.

[33] 5Rights Foundation (2021) *Pathways: how digital design puts children at risk*.

for Children: Best Practice in the Age Appropriate Design Code (5Rights Foundation, 2019).[34]

Persuasive design features are not novel or unregulated practices, but established forms of data-driven design that fall squarely within the AADC and wider data protection framework.

The central issue is not regulatory absence, but implementation and enforcement.

In practice, requirements are not consistently or fully applied, and many services continue to deploy these systems as core tenets of services in ways that prioritise engagement over children's best interests.

This is consistent with evidence from Steve Wood's analysis of the impact of regulation on children's digital lives, produced with the Digital Futures Commission across two phases.[35] Phase 1 found that the AADC had already driven meaningful changes in design, particularly through stronger default settings and a shift towards safety-by-design practices.[36] Phase 2 shows that while UK and EU frameworks continue to influence design and governance, there is increasing reliance on end-user tools rather than embedded protections, alongside uneven implementation and signs of reduced regulatory momentum in some areas.[37]

The priority should be robust and consistent enforcement of the AADC and wider data protection framework, ensuring children's data is used only in ways that demonstrably support their rights, wellbeing, and development. As noted previously, the AADC should be placed on a statutory footing in primary legislation. This would strengthen its legal authority, improve consistency of interpretation and application, and ensure its standards are more reliably translated into real-world design outcomes across services and regulators.

3. Measures to restrict specific features and functionalities

Measures such as age or feature restrictions, when based on the level of risk posed to children, can form an important part of a broader safety-by-design strategy. However, focusing on age gating features in isolation risks misdiagnosing the nature of harm in digital environments. It frames harm primarily as a matter of children's behaviour or self-control, rather than recognising that many risks stem from deliberately engineered, engagement-driven design choices embedded within digital products and services.

---

[34] 5Rights Foundation (2019) *Designing for children: best practice in the Age Appropriate Design Code*.

[35] Wood, S. (2024) *Impact of regulation on children's digital lives: Phase 1*. Digital Futures for Children, LSE; Wood, S. (2026) *Impact of regulation on children's digital lives: Phase 2*. Digital Futures for Children, LSE.

[36] Wood, S. (2024) *Impact of regulation on children's digital lives: Phase 1*. Digital Futures for Children, LSE.

[37] Wood, S. (2026) *Impact of regulation on children's digital lives: Phase 2*. Digital Futures for Children, LSE.

Research by 5Rights Foundation, particularly *Pathways: How Digital Design Puts Children at Risk* (2021), demonstrates how commercial incentives directly shape design decisions.[38] Through avatar-based research and interviews with designers and children, the research shows how recommender systems, autoplay, infinite feeds, frictionless interaction, notifications and engagement optimisation can rapidly direct children towards harmful content pathways and reinforce persistent patterns of use.[39] The report highlights that these systems are intentionally designed to maximise attention, retention and engagement, rather than operating as neutral or passive technologies.[40]

This is consistent with wider evidence on systemic online harms. Livingstone and Stoilova's *The 4Cs: Classifying Online Risk to Children* (2021) identifies how harms increasingly arise through interconnected systems of content, contact, conduct and commercial risks, particularly where services rely on profiling, behavioural targeting, algorithmic curation and large-scale data processing.[41] The report also identifies privacy, health and fair treatment as cross-cutting risks shaped by the broader design and operation of digital services.[42]

Emerging legal and policy scrutiny reinforces these findings, with investigations and regulatory evidence increasingly identifying recommender systems, infinite scroll, and engagement optimisation as structural drivers of harm rather than isolated product features.[43]

Approaches focused only on restricting features or imposing age-based limits should not be seen as alternatives to safety-by-design, but as complementary tools within it. Used alone, they risk placing responsibility on children and families to manage harms that are structurally shaped by service design, while diverting attention from the role of companies in creating and amplifying those risks through recommender systems, persuasive design, profiling and engagement optimisation.

A more effective approach is to combine proportionate age or feature restrictions with stronger upstream design obligations that require services to minimise risk at source. This is consistent with obligations under the Data Protection Act 2018, UK GDPR, and the ICO's Age Appropriate Design Code, which require children's data to be processed in their best interests and risks

[38] 5Rights Foundation (2021) *Pathways: how digital design puts children at risk*.

[39] 5Rights Foundation (2021) *Pathways: how digital design puts children at risk*.

[40] 5Rights Foundation (2021) *Pathways: how digital design puts children at risk*.

[41] Livingstone, S., & Stoilova, M. (2021). *The 4Cs: Classifying Online Risk to Children*. (CO:RE Short Report Series on Key Topics). Hamburg: Leibniz-Institut für Medienforschung | Hans-Bredow-Institut (HBI); CO:RE - Children Online: Research and Evidence.

[42] Livingstone, S., & Stoilova, M. (2021). *The 4Cs: Classifying Online Risk to Children*. (CO:RE Short Report Series on Key Topics). Hamburg: Leibniz-Institut für Medienforschung | Hans-Bredow-Institut (HBI); CO:RE - Children Online: Research and Evidence.

[43] Jamali, L. (2026) Instagram and Youtube owners built 'addiction machines', trial hears. *BBC News*

from profiling and engagement-driven systems to be mitigated at the design stage wherever possible.[44] It also aligns with the Safety by Design Code of Practice, which has been developed by civil society professionals as a practical framework for embedding safety into system architecture from the outset.[45]

Building on this, we recommend that Government require persuasive design features to be off by default for all under-18s, prohibit content recommendation algorithms from optimising for commercial engagement metrics where children are the user, mandate standardised disengagement opportunities including save functions and time-awareness prompts, and require transparency about which persuasive features are operating.

These measures should be implemented through adoption of the Safety by Design Code of Practice[46] as a statutory framework, ensuring these requirements are embedded consistently across services rather than addressed in isolation or on a feature-by-feature basis.

Which services should age restrictions apply to?

1. Factors to consider when determining which apps, sites or services to apply minimum age of access restrictions to

The question of whether minimum age of access restrictions should apply to an app, site or service should not be reduced to a checklist of features or a static classification of product types. Harm does not sit neatly within product categories. It emerges from how services are designed, how data is used, and how systems shape children's behaviour over time. The assessment of which apps, sites or services to apply minimum age of access restrictions to must therefore be grounded in a structured, evidence-based understanding of children and the unique risks they face online.

A starting point, as mentioned previously, should be the CO:RE 4Cs classification of online risk, which provides an evidence-based framework for understanding how harm manifests in children's digital lives.[47] Content risks, contact risks, conduct risks and commercial risks capture the core categories through which children experience harm online.[48] These include exposure to harmful or age-inappropriate material, risks arising from user-to-user

[44] *Data Protection Act 2018*; *UK GDPR*; *Information Commissioner's Office (ICO) (2020) Age appropriate design: a code of practice for online services*.

[45] Online Safety Act Network, 5Rights Foundation and others (2026) Safety By Design Code of Practice.

[46] Online Safety Act Network, 5Rights Foundation and others (2026) Safety By Design Code of Practice.

[47] Livingstone, S., & Stoilova, M. (2021). *The 4Cs: Classifying Online Risk to Children*. (CO:RE Short Report Series on Key Topics). Hamburg: Leibniz-Institut für Medienforschung | Hans-Bredow-Institut (HBI); CO:RE - Children Online: Research and Evidence.

[48] Livingstone, S., & Stoilova, M. (2021). *The 4Cs: Classifying Online Risk to Children*. (CO:RE Short Report Series on Key Topics). Hamburg: Leibniz-Institut für Medienforschung | Hans-Bredow-Institut (HBI); CO:RE - Children Online: Research and Evidence.

interaction, peer-to-peer behaviours, and commercial practices that exploit attention, data or vulnerability.[49]

However, focusing only on discrete categories is no longer sufficient. Risks increasingly arise from systemic design features that operate across all four Cs at once.[50] These include compulsive and addictive design patterns, recommender systems that shape what children see and how they move through services, persuasive engagement mechanics, and commercial models built on continuous data extraction and behavioural optimisation.[51] The impact is cumulative, affecting sleep, attention, learning and emotional development over time rather than through isolated incidents.[52]

A critical dimension of this is the rise of highly personalised, data-intensive services. These systems are defined by extensive data collection, profiling, inference and behavioural targeting, combined with continuous algorithmic curation of content and interactions. This is not a separate category of harm but an accelerant across all four Cs, increasing exposure to harmful content, shaping contact pathways, influencing peer dynamics and conduct, and enabling more sophisticated forms of commercial manipulation.[53]

In practice, relevant considerations for age-based restrictions should therefore include the degree of user-to-user interaction, the ability to upload, stream or generate content including AI-generated media, the use of recommender systems and algorithmic ranking, and most importantly, data usage including the extent of profiling and behavioural personalisation. The presence of persuasive or compulsive design features, the intensity of data collection, the likely age profile of users, and the scale and reach of the service are also central to understanding risk.

Crucially, this assessment should be grounded in a child rights impact assessment that is integrated into the wider risk framework rather than treated as a standalone procedural step. A child rights impact assessment should translate the CO:RE four Cs and system-level risk factors into a focused evaluation of how a service affects children's rights in practice, including safety, privacy, development, participation and protection from exploitation.

---

[49] Livingstone, S., & Stoilova, M. (2021). *The 4Cs: Classifying Online Risk to Children*. (CO:RE Short Report Series on Key Topics). Hamburg: Leibniz-Institut für Medienforschung | Hans-Bredow-Institut (HBI); CO:RE - Children Online: Research and Evidence.

[50] Livingstone, S., & Stoilova, M. (2021). *The 4Cs: Classifying Online Risk to Children*. (CO:RE Short Report Series on Key Topics). Hamburg: Leibniz-Institut für Medienforschung | Hans-Bredow-Institut (HBI); CO:RE - Children Online: Research and Evidence.

[51] Livingstone, S., & Stoilova, M. (2021). *The 4Cs: Classifying Online Risk to Children*. (CO:RE Short Report Series on Key Topics). Hamburg: Leibniz-Institut für Medienforschung | Hans-Bredow-Institut (HBI); CO:RE - Children Online: Research and Evidence.

[52] Livingstone, S., & Stoilova, M. (2021). *The 4Cs: Classifying Online Risk to Children*. (CO:RE Short Report Series on Key Topics). Hamburg: Leibniz-Institut für Medienforschung | Hans-Bredow-Institut (HBI); CO:RE - Children Online: Research and Evidence.

[53] Livingstone, S., & Stoilova, M. (2021). *The 4Cs: Classifying Online Risk to Children*. (CO:RE Short Report Series on Key Topics). Hamburg: Leibniz-Institut für Medienforschung | Hans-Bredow-Institut (HBI); CO:RE - Children Online: Research and Evidence.

It should go beyond identifying risks to assessing how design and system-level dynamics can produce cumulative harm over time, and whether mitigations are effective in practice and embedded by default. This ensures age-restriction decisions are both risk-informed and grounded in children's best interests as they are experienced in real-world use.

This assessment must also ask whether harms are being effectively mitigated through safety-by-design approaches. This includes whether protections are built in by default, rather than relying on user action, and whether a service can demonstrate that it is operating in children's best interests in line with existing regulatory expectations.

This reflects the approach referenced in section 1(3) of the Online Safety Act 2023, as well as the Department for Science, Innovation and Technology's Statement of Strategic Priorities for online safety regulation, which calls for services to embed safety-by-design to deliver safe online experiences for all users, particularly children.[54] The Safety by Design Code of Practice, as referenced in our previous responses, provides a clear framework on how approach should be implemented in practice.[55]

Such a framework ensures that decisions about age restrictions are rooted in how services actually function, not how they are labelled. It recognises that harm is not a feature of individual tools but of system design. It also avoids the risk of regulatory lag, where static lists of features quickly fall behind rapidly evolving technologies.

2. Apps, sites or services that should be captured by minimum age of access restrictions

We would urge caution against an approach that focuses primarily on naming specific types of apps, sites or services for minimum age of access restrictions.

New services can easily emerge to supplant apps, platforms services or sites with similar functions thus immediately circumventing rules aimed at particular categories.

From a children's rights perspective, the focus should therefore be on risks, functionalities, design features and business practices, rather than on labels. A service should not be able to avoid its child protection duties simply by describing itself as a gaming service, messaging service, social media platform, AI chatbot, educational tool or entertainment service.

[54] *Online Safety Act 2023, s.1(3)*; Department for Science, Innovation and Technology (2025) *Final Statement of Strategic Priorities for Online Safety*.

[55] Online Safety Act Network, 5Rights Foundation and others (2026) Safety By Design Code of Practice.

Children should be able to access the benefits of the digital world, including education, creativity, play, information, communication and participation. The policy aim should not be to exclude children from online spaces wholesale, but to require services likely to be accessed by children to be safe, age-appropriate and rights-respecting by design and default.

Minimum age restrictions are certainly an appropriate response for services or features that are inherently unsuitable for children, or where the core purpose of the service carries a significant risk of serious harm. However, for the wider digital environment, a more effective approach would be to regulate harmful design and functionality across services.

This should include, for example, recommender systems that amplify harmful content, compulsive or addictive design features, infinite scroll, autoplay, streaks, manipulative notifications, dark patterns, excessive data collection, commercial profiling, unsolicited contact from adults, and features that encourage overuse or expose children to harmful interactions.

Moreover, some services and content are already subject to age-related restrictions or age-assurance requirements in the UK, and the UK already has an Age-Appropriate Design Code.[56] The issue is therefore not only whether additional services should be age-restricted, but whether existing duties are being effectively enforced and whether harmful design features are being appropriately and sufficiently addressed.

3. Exempt services

We do not consider it appropriate to introduce additional category-based exemptions from existing regulatory requirements where children are reasonably likely to access or use a service.

Exemptions based on service type or stated purpose, for example for “educational” services, “teen” versions of products, or services offering parental controls, risk creating significant regulatory gaps that fail to reflect the realities of children’s digital lives.

Digital services are increasingly convergent in design, combining educational, social, commercial, and AI-enabled functionalities within the same environment. The relevant question is therefore not what a service calls itself, but whether it exposes children to risk in practice.

As discussed previously, a risk-based approach, focused on the 4Cs (content, contact, conduct and commercial risks), alongside systemic risks such as recommender systems, profiling, compulsive design, and generative AI, is more

[56] Information Commissioner’s Office (ICO) (2020) *Age appropriate design: a code of practice for online services*.

appropriate than category-based exemptions.[57] Alongside this, there is a need for the adoption of a robust safety-by-design code of practice, such as the Code developed by 5Rights and partners, setting out the concrete design standards and mitigations services should be expected to implement where children are likely to be affected.[58]

EdTech services provide a particularly clear example of this convergence. Emerging findings from the Digital Futures for Children centre's ongoing *Better EdTech futures for children* research shows that these services now routinely combine functions that extend far beyond teaching and learning.[59] They include user-to-user interaction, content sharing tools, behavioural tracking, real-time learning analytics, algorithmic personalisation and AI-enabled systems.[60] These features are not peripheral. They actively shape how children learn, behave and are assessed through data-driven and engagement-oriented design logics that increasingly mirror wider commercial platform environments.[61]

This is reinforced by *A child rights audit of GenAI in EdTech: Learning from five UK case studies* (Atabey, Sylwander and Livingstone, 2025), which identifies opaque data practices, commercial tracking, profiling, and unreliable or unsafe AI outputs across widely used educational tools.[62] It also finds that children's perspectives are rarely meaningfully incorporated into the design, deployment, or governance of these systems.[63]

For example, the research examined MagicSchool AI, a widely used platform presented as supporting teachers by reducing workload, assisting with lesson planning, and enabling personalised learning.[64] When the system was prompted by a child expressing suicidal thoughts, it initially provided only United States emergency contacts and did not respond appropriately to the UK context until prompted multiple times by the child, after which UK-specific numbers were eventually provided.[65] In addition, despite stated privacy commitments, the

[57] Livingstone, S., & Stoilova, M. (2021). *The 4Cs: Classifying Online Risk to Children*. (CO:RE Short Report Series on Key Topics). Hamburg: Leibniz-Institut für Medienforschung | Hans-Bredow-Institut (HBI); CO:RE - Children Online: Research and Evidence.

[58] Online Safety Act Network, 5Rights Foundation and others (2026) Safety By Design Code of Practice.

[59] Digital Futures for Children centre and 5Rights Foundation (2025) *Better EdTech futures for children*.

[60] Digital Futures for Children centre and 5Rights Foundation (2025) *Better EdTech futures for children*.

[61] Digital Futures for Children centre and 5Rights Foundation (2025) *Better EdTech futures for children*.

[62] Atabey, A., Sylwander, K. R., & Livingstone, S. (2025). *A child rights audit of GenAI in EdTech: Learning from five UK case studies.* Digital Futures for Children Centre & 5Rights Foundation.

[63] Atabey, A., Sylwander, K. R., & Livingstone, S. (2025). *A child rights audit of GenAI in EdTech: Learning from five UK case studies.* Digital Futures for Children Centre & 5Rights Foundation.

[64] Atabey, A., Sylwander, K. R., & Livingstone, S. (2025). *A child rights audit of GenAI in EdTech: Learning from five UK case studies.* Digital Futures for Children Centre & 5Rights Foundation.

[65] Atabey, A., Sylwander, K. R., & Livingstone, S. (2025). *A child rights audit of GenAI in EdTech: Learning from five UK case studies.* Digital Futures for Children Centre & 5Rights Foundation.

platform exposed child users by default to commercial tracking, including from advertisers associated with adult services.[66]

This sits within a wider body of Digital Futures Commission evidence on children's education data. Their report: *Addressing the problems and realising the benefits of processing children's education data* (Livingstone, Atabey & Pothong, 2021) notes that while education data is often justified in terms of benefits such as supporting teaching, identifying learning needs, facilitating safeguarding, improving school administration and enabling research, these benefits remain largely aspirational and unevenly evidenced.[67]

Claims around personalised learning, AI-driven early intervention, automated assessment and behavioural insight are not consistently validated in practice and may vary significantly across children.[68] The report also emphasises that risks and benefits are unevenly distributed, with some children potentially advantaged while others are disadvantaged or excluded.[69] It further identifies structural issues in how education data is processed. These include extensive and continuous data collection with unclear purpose limitation, weak transparency for children and parents, fragmented and inconsistent governance, and limited practical enforcement of data protection principles such as minimisation and fairness.[70]

In this context, schools are often placed in a difficult position, expected to assess complex legal and technical compliance across global EdTech providers without sufficient capacity, while also managing procurement, pedagogy and safeguarding responsibilities.

These concerns are reflected in children's own reported experiences. In *What do children think of EdTech or know of its data sharing?* (Livingstone & Pothong, Digital Futures Commission, 2022), a nationally representative survey of 1,014 UK children aged 7–16 found that nearly all children use EdTech in school, yet most have limited understanding of how their data is collected, used or shared.[71] Over half reported not being told how their data is stored or processed, and fewer than one third recalled meaningful discussion in school

[66] Atabey, A., Sylwander, K. R., & Livingstone, S. (2025). *A child rights audit of GenAI in EdTech: Learning from five UK case studies.* Digital Futures for Children Centre & 5Rights Foundation.

[67] Livingstone, S., Atabey, A. and Pothong, K. (2021) *Addressing the problems and realising the benefits of processing children's education data: Report on an expert roundtable*. Digital Futures Commission and 5Rights Foundation.

[68] Livingstone, S., Atabey, A. and Pothong, K. (2021) *Addressing the problems and realising the benefits of processing children's education data: Report on an expert roundtable*. Digital Futures Commission and 5Rights Foundation.

[69] Livingstone, S., Atabey, A. and Pothong, K. (2021) *Addressing the problems and realising the benefits of processing children's education data: Report on an expert roundtable*. Digital Futures Commission and 5Rights Foundation.

[70] Livingstone, S., Atabey, A. and Pothong, K. (2021) *Addressing the problems and realising the benefits of processing children's education data: Report on an expert roundtable*. Digital Futures Commission and 5Rights Foundation.

[71] Livingstone, S. and Pothong, K. (2022) *What do children think of EdTech or know of its data sharing? Read our survey findings*. Digital Futures Commission.

about data rights.[72] Children also expressed strong concern about commercial data sharing, with 69% stating that none of their personal data should be shared with companies, particularly sensitive categories such as health or financial information.[73]

This research base also shows that schools tend to frame digital issues primarily through safeguarding narratives, focusing on cyberbullying or stranger danger, while giving less attention to broader data protection questions such as profiling, commercial exploitation and long-term data use.[74] This creates a significant gap in children's understanding of how EdTech systems operate within wider data ecosystems and how their rights apply in practice.

This body of evidence demonstrates why educational contexts cannot be assumed to be low risk. These systems frequently blur boundaries between education, commercial data use and AI model development, with children's data repurposed beyond the immediate educational purpose for which it was collected.[75]

The existence of "child accounts", "teen versions", or parental controls should not provide a basis for exemption. These are mitigations, not evidence that a service is inherently safe or compliant. Evidence from the 5Rights *Instagram Teen Accounts Case Study* (2025), alongside findings from the Molly Rose Foundation (*Instagram Teen Accounts: First-of-its-kind testing of safety tools reveals*), indicates that while "teen accounts" introduce additional safeguards, core platform dynamics remain largely unchanged.[76]

Both organisations highlight that safety tools do not consistently prevent exposure to harmful content or fully interrupt engagement-driven recommendation pathways.[77] Recommender systems, engagement optimisation, profiling and data processing practices therefore continue to

---

[72] Livingstone, S. and Pothong, K. (2022) *What do children think of EdTech or know of its data sharing? Read our survey findings*. Digital Futures Commission.

[73] Livingstone, S. and Pothong, K. (2022) *What do children think of EdTech or know of its data sharing? Read our survey findings*. Digital Futures Commission.

[74] Livingstone, S. and Pothong, K. (2022) *What do children think of EdTech or know of its data sharing? Read our survey findings*. Digital Futures Commission.

[75] Digital Futures for Children centre and 5Rights Foundation (2025) *Better EdTech futures for children;*Atabey, A., Sylwander, K. R., & Livingstone, S. (2025). *A child rights audit of GenAI in EdTech: Learning from five UK case studies.* Digital Futures for Children Centre & 5Rights Foundation; Livingstone, S., Atabey, A. and Pothong, K. (2021) *Addressing the problems and realising the benefits of processing children's education data: Report on an expert roundtable*. Digital Futures Commission and 5Rights Foundation; Livingstone, S. and Pothong, K. (2022) *What do children think of EdTech or know of its data sharing? Read our survey findings*. Digital Futures Commission.

[76] 5Rights Foundation (2025) *Is Instagram now safe for teens? Case study – Instagram Teen Accounts; Molly Rose Foundation, Fairplay, Cybersecurity for Democracy and ParentsSOS (2025) Teen Accounts, Broken Promises: How Instagram is Failing to Protect Minors*.

[77] 5Rights Foundation (2025) *Is Instagram now safe for teens? Case study – Instagram Teen Accounts; Molly Rose Foundation, Fairplay, Cybersecurity for Democracy and ParentsSOS (2025) Teen Accounts, Broken Promises: How Instagram is Failing to Protect Minors*.

structure children's experiences, meaning that risks are not removed but only partially constrained at the margins rather than addressed at source.[78]

Allowing exemptions based on service label, educational framing, or child-targeted features risks underestimating real-world use by children, overlooking embedded AI and social functionalities, and creating inconsistent protections across services with similar risk profiles. It also risks embedding a false assumption that educational purpose equates to low risk, despite clear evidence that EdTech and AI systems can reproduce or amplify the same harms seen in other digital environments.

A more robust approach is therefore to ensure that all services likely to be accessed by children remain within scope, with obligations scaled according to risk, functionality, and design features rather than category. This would provide a more coherent, future-proof framework that reflects how children's digital environments actually operate and ensures that rights protections keep pace with converging technologies.

Artificial intelligence (AI) chatbots

1. The benefits and risks to children

Our position remains consistent with what we set out previously in relation to social media: the question of "benefits" to children is too narrow on its own and risks creating a false binary between benefit and harm. That framing does not reflect how children actually experience digital environments, and it is equally limited when applied to AI chatbots, albeit with important nuances given the pace and pervasiveness of AI development.

As with social media, AI chatbots are not discrete tools that children simply choose to use for defined purposes. They are embedded across services and devices, including increasingly in education contexts, shaping how children access information, learn, create and communicate as part of everyday life. This embeddedness makes it difficult to separate "benefits" from "risks" in any meaningful way, because both arise from the same systems and design choices.

AI is also already pervasive across children's digital environments in education. Ongoing work from the Digital Futures for Children centre's on EdTech and AI-enabled learning environments highlights in early findings that these systems increasingly embed automated decision-making, data-driven personalisation, learning analytics and generative AI tools into core educational processes, shaping how children access content, are assessed, and receive feedback.[79]

[78] 5Rights Foundation (2025) *Is Instagram now safe for teens? Case study – Instagram Teen Accounts; Molly Rose Foundation, Fairplay, Cybersecurity for Democracy and ParentsSOS (2025) Teen Accounts, Broken Promises: How Instagram is Failing to Protect Minors*.

[79] Digital Futures for Children centre and 5Rights Foundation (2025) *Better EdTech futures for children*.

The key point is not whether this is beneficial or harmful in abstract terms, but that it is now a structural feature of children's digital and educational experience, and therefore requires a shift in focus towards governance, design and accountability rather than framing the debate around optional use or isolated impacts.

As highlighted in Internet Matters' *Me, Myself and AI* (2025), children are already using AI tools in overlapping ways, including for learning support, creativity, curiosity-driven questions and companionship, while also encountering issues such as inaccurate outputs, over-reliance and limited understanding of how responses are generated.[80] This reflects the same broader pattern identified in relation to social media: children's experiences of digital systems are rarely singular or static, but involve simultaneous opportunities and risks shaped by design, context and use.

The key difference with AI is that these systems introduce a higher degree of automation, inference and generative output, which can intensify existing concerns around accuracy, persuasion and dependency. This reinforces, rather than replaces, the need for a rights-based rather than binary "benefits versus harms" framing.

This is further supported by Steve Wood's I*mpact of Regulation on Children's Digital Lives: Phase 2* (2026), which highlights that AI-driven systems are rapidly reshaping children's digital environments and that existing regulatory approaches are not yet sufficiently equipped to address the scale and nature of emerging risks.[81] The report underscores the need for stronger, more anticipatory governance where AI systems are concerned, reflecting their capacity to generate new forms of influence and interaction at scale.[82]

Children's experiences should also not be treated as uniform. As 5Rights out in *Digital Childhood: Addressing Childhood Development Milestones in the Digital Environment* (2023), children's capacities, needs and vulnerabilities vary significantly by age and context, shaping how they engage with and are affected by both social media and AI systems.[83]

A more appropriate framing is therefore to move away from debating AI primarily in terms of benefits or harms and instead focus on children's rights and the conditions under which these systems operate. This includes recognising that current deployment is heavily shaped by commercial incentives, and that the policy question is how to ensure children's best interests, privacy and safety are

[80] Internet Matters (2025) *Me, myself and AI: Understanding and safeguarding children's use of AI chatbots*.

[81] Wood, S. (2026) *Impact of regulation on children's digital lives: Phase 2*. Digital Futures for Children, LSE.

[82] Wood, S. (2026) *Impact of regulation on children's digital lives: Phase 2*. Digital Futures for Children, LSE.

[83] 5Rights Foundation (2023) *Digital childhood: addressing childhood development milestones in the digital environment*.

prioritised in design, development and governance rather than secondary to engagement, scale or monetisation.

Children have rights to access information, express themselves and participate in cultural and social life, alongside rights to protection, privacy and the consideration of their best interests under the UN Convention on the Rights of the Child and General Comment No. 25.[84] This means the central question is not whether AI chatbots deliver benefits or harms in isolation, but how they are designed, risk-assessed, governed and deployed in practice, and whether they are aligned with children's rights and developmental needs in a context where these systems are already embedded across both everyday digital life and education.

2. Risky AI chatbot features for children

Consistent with 5Rights' *Children & AI Design Code* (2025), the assessment of AI chatbot features should be grounded in a child rights and risk-based approach that considers how combinations of design features shape children's understanding, behaviour, and exposure to harm.[85]

The Code emphasises that risks do not arise solely from individual features in isolation, but from how AI systems are designed to interact with children, generate content, personalise responses, and encourage ongoing engagement.[86] In this context, structured risk assessment should be a necessary and continuous requirement, not a one-off exercise, given the evolving nature of AI systems and their effects on children.

This is particularly relevant in relation to misinformation, false or misleading content, and children's ability to assess the reliability of AI-generated outputs. The realism of AI interactions, including highly human-like conversational styles and fluent, confident responses, may increase the likelihood that children perceive outputs as authoritative or trustworthy.

This point is reflected in Internet Matters' *Me, Myself and AI* chatbot research (2025), which finds that children and young people often treat chatbot responses as helpful and credible, particularly where answers are conversational, immediate, and personalised, and do not consistently verify accuracy against other sources.[87] This can affect children's ability to distinguish between synthetic and human communication, especially where systems present information confidently regardless of uncertainty or correctness.

---

[84] UN Convention on the Rights of the Child; UN Committee on the Rights of the Child (2021) *General comment No. 25 (2021) on children's rights in relation to the digital environment*.

[85] 5Rights Foundation (2025) *Children & AI Design Code: A protocol for the design, development and deployment of AI systems that impact children.*

[86] 5Rights Foundation (2025) *Children & AI Design Code: A protocol for the design, development and deployment of AI systems that impact children.*

[87] Internet Matters (2025) *Me, myself and AI: Understanding and safeguarding children's use of AI chatbots*.

The Children & AI Design Code also highlights that personalisation, conversational continuity, and the ability to recall interactions across sessions can enable systems to adapt dynamically over time based on previous engagement.[88] In practice, this means children may receive increasingly tailored responses shaped by inferred preferences, behaviours, vulnerabilities, or emotional cues, often without understanding how those inferences are made or how information is being processed.[89]

Internet Matters similarly notes that children can develop relational expectations of AI chatbots, sometimes describing them in ways that suggest companionship or social interaction, particularly where systems are designed to be responsive, affirming, and conversational.[90] Features that mimic friendship, empathy, or emotional responsiveness compound risks by encouraging children to interpret AI systems as socially aware or emotionally trustworthy agents.[91] Where systems flatter users, ask questions back, or encourage sustained interaction, these features can function as engagement and retention mechanisms that increase the persuasive power of inaccurate or misleading content, creating particular risks for children.[92]

The risk of hallucination, fabrication, or false responses is therefore particularly significant in children's contexts. Generative AI systems can produce inaccurate or entirely fabricated information presented in authoritative language, making it difficult for children to distinguish between reliable and unreliable outputs.[93] This raises important questions about how systems communicate uncertainty, signpost reliability, and provide safeguards that reflect children's developmental needs.

The type of content generated also matters. AI systems capable of producing text, images, audio, or video can generate misleading or synthetic content across multiple modalities. The 5Rights framework highlights the importance of safeguards that reduce the likelihood of children being exposed to inappropriate, manipulative, or low-quality outputs, including synthetic content that may not be easily identifiable as AI-generated.[94]

Recommendation systems and engagement optimisation mechanisms should also be assessed for how they prioritise, amplify, or surface misleading or low-

[88] 5Rights Foundation (2025) *Children & AI Design Code: A protocol for the design, development and deployment of AI systems that impact children.*

[89] 5Rights Foundation (2025) *Children & AI Design Code: A protocol for the design, development and deployment of AI systems that impact children.*

[90] Internet Matters (2025) *Me, myself and AI: Understanding and safeguarding children's use of AI chatbots*.

[91] Internet Matters (2025) *Me, myself and AI: Understanding and safeguarding children's use of AI chatbots*.

[92] Internet Matters (2025) *Me, myself and AI: Understanding and safeguarding children's use of AI chatbots*.

[93] 5Rights Foundation (2025) *Children & AI Design Code: A protocol for the design, development and deployment of AI systems that impact children.*

[94] 5Rights Foundation (2025) *Children & AI Design Code: A protocol for the design, development and deployment of AI systems that impact children.*

quality information. As highlighted in the Science, Innovation and Technology Committee report, *Social Media, Misinformation and Harmful Algorithms* (2025), there is a need for stronger transparency and accountability around how algorithmic systems shape information exposure, particularly for children.[95]

Overall, these features should not be treated as inherently harmful in all contexts. However, the evidence shows they can contribute to heightened risks around misinformation, manipulation, emotional dependency and misplaced trust depending on how they are implemented, combined and governed. This reinforces the importance of embedding ongoing, child rights-based risk assessment as a core requirement of AI system design, development and deployment, rather than relying on static or retrospective checks.

More broadly, there is an urgent need for a coherent, holistic strategy for children and AI. Government must move beyond considering individual tools or isolated harms and instead examine how AI systems are being embedded across the full range of children's lives and experiences, especially in contexts such as education where children may have little or no meaningful choice about whether to engage with them.

Government policy in this area is becoming internally contradictory: while some discussions focus on whether certain AI systems pose unacceptable risks to children, government is simultaneously accelerating the rollout of AI technologies into schools and other child-facing environments.[96] Without a joined-up approach to governance, there is a danger that adoption will continue to outpace scrutiny, with children expected to navigate the consequences of systems that have not yet been subject to clear child rights-based standards, safeguards or accountability.

3. Should AI chatbots have minimum age restrictions?

We believe that there should be minimum age requirements for AI chatbots.

We refer to our response to earlier questions regarding age restrictions and reiterate our position that personalised and algorithmically driven services, which would include AI chatbots, should be prohibited for children under 13.

More broadly, we recommend that Government establish a statutory framework to deliver age-appropriate online experiences for children and young people, built around three core elements: a clear minimum baseline for under-13s, a mandatory pre-certification regime for services accessed by under-18s, and a strengthened enforcement framework.

---

[95] Science, Innovation and Technology Committee (2025) *Social media, misinformation and harmful algorithms*. Second Report of Session 2024–25 (HC 441).

[96] Department for Education and Department for Science, Innovation and Technology (2026) *450,000 disadvantaged pupils could benefit from AI tutoring tools*.

As a minimum baseline, personalised services, including behavioural profiling, recommender systems and algorithmically driven content delivery, should not be accessible to children under 13. All in-scope services should be required to implement highly effective age assurance measures to enforce these protections.

For users aged 13–18, access to digital services should be conditional on compliance with a mandatory pre-certification regime based on graduated age-appropriate design and safety standards. Services should be required to demonstrate compliance before being made available to children and young people, with differentiated standards applied across age bands, for example 13+ and 15/16+.

A precautionary principle should underpin the regime, meaning that services, functionalities and significant product changes should be treated as unavailable to under-18s unless and until they have been independently assessed and certified as compliant. This is particularly urgent in the context of AI chatbots and generative AI systems, where the longer-term developmental, psychological and societal impacts on children remain poorly understood.

In these circumstances where evidence of safety is limited, but the potential scale of harm may be significant or irreversible, there is a strong case for temporary “circuit breaker” measures to pause or restrict deployment until systems can be independently demonstrated to be safe, rights-respecting and developmentally appropriate for children.

The framework should be supported by an independent multidisciplinary expert panel responsible for developing and maintaining certification criteria, determining appropriate age thresholds for services and functionalities, and reviewing emerging technologies, AI-driven features and risks on an ongoing basis. Highly effective age assurance should form a core part of the regime and should be risk-based and proportionate, privacy-preserving and secure, effective against circumvention, accessible and inclusive, and transparent and accountable.

The framework must also be supported by robust enforcement. Regulators, including Ofcom and the Information Commissioner’s Office, should work alongside the independent expert panel and be properly empowered and resourced to require evidence of compliance, investigate breaches and impose meaningful sanctions, including business disruption measures, injunctive relief and executive accountability where appropriate.

We also reiterate our concern that Government requires a more holistic and joined-up approach to children and AI. Minimum age restrictions should not be considered in isolation or treated as the sole mechanism for addressing risk. Rather than focusing exclusively on particular chatbot functionalities, regulation should assess how combinations of features, design choices and deployment contexts create risks for children in practice.

That said, stronger safeguards and minimum age restrictions are likely to be particularly important where AI chatbots incorporate functionalities that increase emotional dependency, persuasive influence or behavioural manipulation. This may include systems that simulate human relationships or companionship, encourage prolonged or compulsive engagement, provide personalised emotional affirmation, profile users in order to adapt responses, or are embedded within recommender, commercial or educational ecosystems in ways that children cannot easily avoid or meaningfully opt out of.

Importantly, risks do not arise solely from individual features in isolation, but from how systems are designed, combined and deployed. This is especially significant in contexts such as education, where children may be expected or required to engage with AI systems as part of everyday learning. In such cases, the question cannot simply be whether a chatbot meets a minimum age threshold, but whether the wider system has been designed and governed in a way that is compatible with children's rights, developmental needs and best interests.

For this reason, we encourage Government to move beyond a solely functionality-based approach and towards a broader safety-by-design framework for children and AI, supported by ongoing child rights impact and risk assessments.

4. **The impact of introducing age restrictions on AI chatbots or certain features and functions**

The impact of introducing age restrictions on AI chatbots or certain functionalities will depend significantly on how such measures are designed, implemented, and enforced. Age restrictions may help reduce children's exposure to some of the most concerning risks associated with AI systems, including manipulative interactions, emotional dependency, harmful or misleading information, compulsive engagement patterns, inappropriate content, and profiling-driven personalisation. In particular, restrictions may be justified where systems are designed in ways that increase persuasive influence, simulate human relationships, or encourage prolonged interaction without meaningful safeguards.

However, age restrictions alone are unlikely to be sufficient. If implemented in isolation, there is a risk they become a substitute for addressing the underlying design choices and business models that generate risk in the first place. As we have set out elsewhere in this response, harms do not arise solely from children accessing particular technologies at a certain age, but from the ways systems are designed, deployed and incentivised.

We are also concerned that the Government's current approach is inconsistent and insufficiently joined up. On the one hand, policymakers are exploring whether certain AI systems or functionalities may be sufficiently harmful to warrant age restrictions or other limitations for children. On the other hand, Government is simultaneously encouraging and accelerating the rollout of AI systems into child-facing contexts, including education, where children may have little or no meaningful ability to opt out.

This creates a significant tension in policy thinking. If AI systems are considered potentially high-risk for children in some contexts, there must also be serious scrutiny of what it means to embed those same technologies into environments where use is expected, normalised or effectively mandatory. Without a coherent cross-government strategy for children and AI, there is a danger that deployment will continue to outpace governance, with children expected to absorb the consequences of unresolved policy questions in real time.

For parents and carers, age restrictions may provide some reassurance or clarity, particularly where they are supported by effective age assurance and clear safety standards. However, parental oversight cannot substitute for robust safety-by-design obligations on companies themselves. Equally, any age assurance measures must be proportionate, privacy-preserving and rights-respecting, with careful consideration given to impacts on children's privacy, access to information and participation rights.

Chapter 3 – Enforcement and compliance

Age assurance

1. Key considerations for effective and workable minimum age restrictions

Minimum age restrictions can only be effective if they are embedded within a broader risk-based and safety-by-design framework rather than treated as a standalone safeguard. Age assurance is important, but it is not a single technical solution or gatekeeping mechanism. It is better understood as a spectrum of approaches that should be proportionate to risk, privacy-preserving, and tailored to context.

But How Do They Know It Is a Child? (5Rights Foundation, 2021) explains that age assurance includes a range of methods such as age verification, age estimation, inference, and parental controls, each with different strengths, limitations, and privacy implications.[97] The report stresses that no single method is sufficient on its own and highlights the importance of balancing accuracy, proportionality, transparency, accountability, and data minimisation in the design and deployment of age assurance systems, particularly where they rely on inference, behavioural signals, or extensive data collection.[98]

This is supported by 5Rights' *Age Assurance as a Spectrum: A Risk-Based Approach from a European Perspective* (2026), which similarly emphasises that age assurance should be risk-calibrated and designed to support children's

[97] 5Rights Foundation (2021) *But how do they know it is a child? Age assurance and children's data protection online.*

[98] 5Rights Foundation (2021) *But how do they know it is a child? Age assurance and children's data protection online.*

rights across different developmental stages.[99] While developed from a European policy context, its core argument is relevant to UK discussions: age assurance should enable appropriate design and protection rather than function as a blunt access control mechanism applied uniformly across services.

Both resources underline that age assurance is not a "silver bullet". Its effectiveness depends on the wider system within which it operates. This is where safety-by-design becomes critical. As set out previously, the Safety by Design Code of Practice co-developed by 5Rights Foundation and partners ensures that safety is built in from the outset: risks should be designed out where possible, remaining risks should be managed through technical and policy safeguards, and enforcement or redress should be a last resort rather than the primary safety measure.[100]

In this framework, age assurance works best as a supporting layer rather than a compensatory fix. Where services are designed with children in mind from the outset, age assurance can be applied more precisely and proportionately, targeting genuinely high-risk services or features. Where services are not safely designed, age restrictions alone risk being ineffective or shifting responsibility onto users and families.

Minimum age restrictions become workable when they sit within systems that already reduce risk by design, limit unnecessary data collection, and apply age-appropriate defaults. This allows age assurance to function as intended: not as a blunt barrier, but as one component in a coherent, rights-respecting safety framework.

2. Making age assurance more effective

Age assurance can be made more effective, but the key issue is how effectiveness is defined, measured, and enforced in practice. Under the Online Safety Act, Ofcom has set out criteria for Highly Effective Age Assurance (HEAA), including technical accuracy, robustness, reliability, and fairness, alongside principles of accessibility and interoperability.[101] These provide a necessary foundation, but they do not yet ensure consistent, outcomes-based protection for children.

A primary gap is the absence of clear, measurable performance standards. There are no minimum thresholds for accuracy and no agreed benchmarks for

[99] 5Rights Foundation (2022) *Age assurance as a spectrum: Designing proportionate approaches to protecting children online.*

[100] Online Safety Act Network, 5Rights Foundation and others (2026) Safety By Design Code of Practice.

[101] Ofcom (2025) *Highly effective age assurance: Guidance on children's access and protection duties under the Online Safety Act.*

what constitutes "effective prevention of underage access."[102] This limits comparability between systems and risks compliance being driven by process rather than demonstrable outcomes.

There is also insufficient focus on effectiveness in practice. As set out in *But How Do They Know It Is a Child?* (5Rights Foundation, 2021), age assurance should be understood as a spectrum of methods, including self-declaration, inference, age estimation, and verified credentials, each carrying different levels of assurance, risk, and privacy impact.[103] The report highlights the importance of proportionality, transparency and accountability in how age assurance systems are designed and deployed.[104]

Effectiveness must therefore be assessed on real-world performance, including whether systems reliably prevent underage access without unnecessary exclusion, error or excessive data collection. Privacy is not yet fully embedded within the concept of effectiveness. While GDPR and the Age Appropriate Design Code are relevant considerations, many age assurance systems rely on additional data collection, inference, or profiling.

5Rights' *Age Assurance as a Spectrum: A Risk-Based Approach from a European Perspective* (2026) reinforces that age assurance must be privacy-preserving and proportionate, ensuring that protections for children do not introduce new risks through excessive data processing.[105] From a UK regulatory perspective, this highlights the need to treat privacy as integral to effectiveness, not separate from it.

Accountability and redress mechanisms also remain limited. There are currently no consistent, accessible routes for users to challenge incorrect age determinations or seek meaningful remedy where systems fail. This weakens the ability to test effectiveness from the perspective of children and families in practice.

Improving effectiveness requires a shift towards a risk-based and proportionate approach to age assurance. Age assurance should only be used where necessary and proportionate to risk, and the method selected should reflect the level of risk posed by the service or feature. It should be treated as a spectrum of assurance, rather than a binary requirement.

Effectiveness should be defined in terms of measurable, real-world outcomes. This includes whether systems actually prevent underage access, while also avoiding unnecessary exclusion or disproportionate barriers for users.

---

[102] Ofcom (2025) *Highly effective age assurance: Guidance on children's access and protection duties under the Online Safety Act.*

[103] 5Rights Foundation (2021) *But how do they know it is a child? Age assurance and children's data protection online.*

[104] 5Rights Foundation (2021) *But how do they know it is a child? Age assurance and children's data protection online.*

[105] 5Rights Foundation (2022) *Age assurance as a spectrum: Designing proportionate approaches to protecting children online.*

- Privacy must be embedded as a core component of effectiveness. Systems should be designed to minimise data collection, ensure strict purpose limitation, and avoid secondary use of data. Only the minimum data necessary to establish age or age range should be collected.
- Accessibility and fairness must also be central. Users should not be forced into a single method of verification, and systems must be designed to avoid excluding legitimate users or creating discriminatory outcomes.
- Transparency and accountability are essential. Services should clearly explain what data is used, how age determinations are made, and how long data is retained. Critically, users should have meaningful routes to challenge incorrect decisions and access redress where systems fail.

Without these elements, age assurance risks becoming a compliance exercise focused on deployment rather than a robust system that delivers consistent, rights-respecting protection for children in practice.

3. Assessing the effectiveness of age verification and age assurance technologies

When assessing the effectiveness of age verification and age-assurance technologies, the focus should be on whether systems deliver reliable protection for children in real-world conditions, not simply whether they are deployed or meet technical specifications on paper.

Ofcom's Highly Effective Age Assurance (HEAA) criteria, including technical accuracy, robustness, reliability and fairness, provide a necessary baseline.[106] However, effectiveness depends on whether these principles translate into measurable outcomes that can be observed and enforced in practice across different services and user contexts.

Our recommendations are set out as follows:

Real-world performance

Effectiveness should be judged on whether systems actually prevent underage access in operational environments, including attempts at circumvention and cross-platform use. Laboratory or controlled testing alone is insufficient to demonstrate real-world effectiveness.

Error rates and their implications

A robust assessment must include transparent reporting of both false positives and false negatives. False negatives undermine child safety by allowing underage access, while false positives can lead to inappropriate

[106] Ofcom (2025) *Highly effective age assurance: Guidance on children's access and protection duties under the Online Safety Act.*

exclusion and restrictions on legitimate users. Both must be measured to understand system performance in practice.

Clear and comparable benchmarks

At present, there is no shared operational definition of what "effective" means. Without agreed benchmarks or minimum thresholds, it is difficult to compare services or ensure consistent regulatory enforcement. Effectiveness needs to be defined in a way that enables comparability and accountability across providers.

Robustness against circumvention

Age assurance systems operate in dynamic and adversarial environments. Their effectiveness depends on their ability to withstand evolving attempts to bypass controls, including through workarounds and cross-service manipulation. This resilience should be a core element of assessment, not a secondary consideration.

Privacy and data protection impacts

Effectiveness cannot be assessed without considering how systems process data. As set out in 5Rights Foundation's *But How Do They Know It Is a Child?* (2021) and *Age Assurance as a Spectrum: A Risk-Based Approach from a European Perspective* (2026); systems should be evaluated against data minimisation and purpose limitation principles.[107] Technical effectiveness should not be achieved through disproportionate data collection, profiling, or inference.

Fairness and accessibility of outcomes

Effectiveness also depends on whether systems operate fairly across different groups. This includes ensuring that children and adults are not unfairly excluded and that systems do not create barriers linked to access to technology, documentation, or identity infrastructure.

Transparency and auditability

Finally, effectiveness depends on the ability to scrutinise performance. Providers should publish meaningful, standardised information on system outcomes, including error rates and testing methodologies. Independent auditability is essential to ensure that claims of effectiveness can be verified rather than assumed.

This approach supports an understanding of effectiveness grounded in real-world outcomes, rights-respecting design, and enforceable standards that can be consistently applied across services.

[107] 5Rights Foundation (2021) *But how do they know it is a child? Age assurance and children's data protection online*; 5Rights Foundation (2022) *Age assurance as a spectrum: Designing proportionate approaches to protecting children online.*

Circumvention of age limits

1. Circumvention methods beyond VPNs

Firstly, it is important to note that children should not be positioned as the source of the problem. They are navigating services designed to maximise engagement, where safety protections are uneven, inconsistently applied, and often not embedded in ways that reflect how children actually use platforms.

Evidence from *The Online Safety Act: Are children safer online?* (Internet Matters, 2026) shows that, beyond VPNs or technically advanced tools, children commonly encounter and use relatively simple methods to access restricted content or features where protections are weak or inconsistently applied.[108] These include entering false or adjusted dates of birth during account creation, using older siblings' or adults' accounts, switching between accounts on shared devices, and moving between platforms or features where age checks are not uniformly enforced.[109] The report also highlights that in some services age assurance is only applied at onboarding, meaning once access is granted, children can move across unmoderated or less restricted areas of the platform without further checks.[110]

Importantly, the research points to variation not only between services but within them.[111] This inconsistency means that children may encounter strict age gates in one part of a service, while other features such as content feeds, messaging functions, or embedded tools are subject to lighter or no age assurance.[112] These design gaps create predictable routes through which access restrictions can be bypassed without requiring technical sophistication. This does not indicate intentional circumvention so much as it reflects structural weaknesses in implementation.

Where protections rely on self-declaration, one-off verification, or feature-specific controls rather than consistent, system-wide safety-by-design, workarounds become foreseeable outcomes of system design. The Online Safety Act 2023 and the Age Appropriate Design Code establish important expectations, but the Internet Matters evidence reinforces that implementation remains uneven in practice.[113]

This is also reflected in Steve Wood's *Impact of Regulation on Children's Digital Lives: Phase 2* (2026), which identifies increasing reliance on end-user and parental control measures rather than embedded protections within service

[108] Internet Matters (2026) *The Online Safety Act: Are children safer online?*

[109] Internet Matters (2026) *The Online Safety Act: Are children safer online?*

[110] Internet Matters (2026) *The Online Safety Act: Are children safer online?*

[111] Internet Matters (2026) *The Online Safety Act: Are children safer online?*

[112] Internet Matters (2026) *The Online Safety Act: Are children safer online?*

[113] Internet Matters (2026) *The Online Safety Act: Are children safer online?*

design itself.[114] The result is a safety environment where protections depend too heavily on user behaviour or provider discretion, rather than being reliably embedded across services and features.

Age assurance measures should not be expected to be completely foolproof, just as offline age checks are not infallible. The key question is whether they are proportionate and effective in reducing children's exposure to foreseeable harms while protecting their wider rights. As recognised in by the UN, children's rights to protection must be balanced with their rights to privacy, expression, access to information, and participation.[115]

Where age assurance can help prevent children from inadvertently encountering harmful or developmentally inappropriate content or features, while respecting these wider rights, it can represent a proportionate and worthwhile safeguard.

The policy focus should therefore be on ensuring that age assurance and safety measures are consistently applied, difficult to bypass in predictable ways, and designed into services from the outset. The issue is not how children navigate these systems, but whether products, services and providers are meeting their obligations to make those systems genuinely safe by design.

2. What the government should prioritise to reduce the circumvention of online safety rules in the UK

Rather than focusing primarily on user circumvention measures, the Government should prioritise ensuring that companies fully comply with their existing regulatory duties and that regulators have the powers, resources, and clarity needed to enforce the law effectively and consistently.

The core issue is not simply that some users may attempt to bypass safety measures, but whether online services are meeting their obligations under the UK's existing regulatory framework, including the Online Safety Act 2023 and the Age Appropriate Design Code. Stronger compliance and enforcement would help ensure that safety-by-design duties are meaningfully implemented across services used by children, reducing reliance on reactive or user-level interventions.

A more robust enforcement approach would also create clearer and more consistent baseline protections for children online. Without effective oversight and accountability, there is a risk that safety requirements are treated as

[114] Wood, S. (2026) *Impact of regulation on children's digital lives: Phase 2*. Digital Futures for Children, LSE.

[115] UN Convention on the Rights of the Child; UN Committee on the Rights of the Child (2021) *General comment No. 25 (2021) on children's rights in relation to the digital environment*.

optional or applied unevenly, undermining the effectiveness of the overall regime.

This concern is particularly acute given the limited enforcement action taken to date. Despite the Age Appropriate Design Code being in force since 2021, the ICO has issued only a very small number of formal enforcement actions under the framework, despite widespread and well-documented concerns about children's data practices and potential GDPR infringements across major online services.[116]

5Rights itself has provided evidence to both Ofcom and the ICO relating to serious child safety risks, including concerns around real and AI-generated child sexual abuse material and platform failures to act effectively in response. Yet, even in cases involving the most egregious and unacceptable harms, and clear risks of further harm to children, there remains uncertainty about which regulator will take meaningful action and under what powers.

We must avoid a perception that significant child safety failures can persist without timely regulatory intervention, weakening both deterrence and public confidence in the enforcement regime. Therefore, the Government should prioritise strengthening regulatory enforcement capacity and accountability, ensuring that existing online safety and data protection duties are consistently applied in practice rather than undermined through uneven or delayed enforcement.

3. Age restricting VPNs

We disagree that everyone should go through age checks to access a VPN if it would prevent children from using them.

Age-gating VPNs risks focusing regulatory attention on the wrong point of intervention when it comes to keeping children safe online. The central issue is whether the services children are accessing are meeting their obligations to provide age-appropriate protections and safe experiences by design.

Age assurance is not infallible, but it remains an important component of a broader safety-by-design framework. Effective age assurance introduces meaningful friction, supports the enforcement of minimum age limits, and requires services to "know their customer" sufficiently to apply age-appropriate protections. Responsibility for this should sit with technology companies.

As in the offline world, where businesses are expected to verify age before selling age-restricted goods such as alcohol or knives, digital services must take responsibility for who is accessing their products and services.

[116] Information Commissioner's Office (ICO) (2026) *Reddit issued with £14.47m fine for children's privacy failures*.

Our position is that children under 13 should not have access to personalised digital services, including social media, gaming platforms, AI chatbots, and VPNs, and companies should properly enforce their own terms of service on age limits. This means platforms must either design services that are safe for all users regardless of age or ensure that where a user's age cannot be reliably established, including where VPNs are used, access to higher-risk features is appropriately restricted.

Restricting VPN access in itself is unlikely to address the underlying safety problem. If high-risk services remain poorly designed, weakly enforced, or dependent on inconsistent age assurance measures, children may simply move to other routes or services. Focusing too heavily on VPNs therefore risks diverting regulatory and technical effort away from the more important task of ensuring that platforms themselves are safe by design and compliant with their existing duties under the Online Safety Act.

There are also broader implications for privacy, security, and proportionality. VPNs are widely used by adults and young people for legitimate purposes, including protecting personal data, securing connections on public networks, and reducing exposure to tracking or profiling. Measures to age-gate VPNs could require additional identity verification or data collection, potentially creating new privacy and security risks for users while weakening access to privacy-enhancing technologies.

The effectiveness of any intervention should be judged by whether it improves children's real-world safety and rights online. In our view, the priority should be strengthening platform accountability, consistent enforcement of age limits, and safety-by-design obligations across digital services, rather than placing disproportionate emphasis on restricting access to VPNs themselves.

Chapter 4 – Preparing children for a digital future

Media and digital literacy

1. **Where do children and families most need support with media and digital literacy?**

   Our Youth Ambassadors consistently tell us they want a deeper understanding of how the online world actually works, particularly in relation to how digital services are designed and operate behind the scenes. This includes how their data is collected, used, and shared, as well as how surveillance, profiling, and recommendation systems shape their online experience.

   5Rights' *Disrupted Childhood* report highlights that both children and parents often struggle to understand how digital services function in practice, particularly where design features and data-driven systems operate in ways that

are not visible to users.[117] It points to a wider lack of clarity about how online services shape attention, interaction, and exposure to content through algorithmic and engagement-driven design, leaving families with limited understanding of how social media and similar services actually work.[118]

Recent evidence from Internet Matters also highlights that children and parents are navigating complex digital environments with uneven understanding of how systems, safety tools, and data-driven features operate in practice.[119] This includes how content is surfaced, how automated systems make decisions, and how service rules are enforced.[120]

This is also reflected by Ofcom themselves in their *Children's Media Literacy Report* (2025), which found that while children are highly engaged digital users, many have limited understanding of how online platforms operate, including how algorithmic systems curate content, how commercial influences shape what they see online, and how data is collected and used across services.[121] The report also highlights gaps in children's ability to critically assess online information and understand the broader functioning of digital platforms.[122]

Evidence from the Digital Futures Commission reinforces this gap in educational contexts. In *What do children think of EdTech or know of its data sharing?*(Livingstone & Pothong, 2022), children reported that schools rarely explain how digital systems operate, including how data is collected and used within educational technologies.[123] Over half said they had not been told how their information is stored, used, or shared, and fewer than one third reported meaningful discussions about data rights, contributing to a broader lack of understanding of how data-driven systems underpin everyday digital environments.[124]

Our Youth Ambassadors also highlight that digital literacy education too often frames children as the "risk" or "problematic users," rather than helping them understand how digital services actually function. This is reflected in the Digital Futures Commission findings, which show that schools tend to prioritise safeguarding framings such as cyberbullying and "stranger danger", while underemphasising how profiling, recommendation systems, and commercial

[117] 5Rights Foundation (2023) *Disrupted childhood: the cost of persuasive design*.

[118] 5Rights Foundation (2023) *Disrupted childhood: the cost of persuasive design*.

[119] Internet Matters (2026) *The Online Safety Act: Are children safer online?*

[120] Internet Matters (2026) *The Online Safety Act: Are children safer online?*

[121] Ofcom (2025) *Children's media literacy report 2025: Children and parents – media use and attitudes.*

[122] Ofcom (2025) *Children's media literacy report 2025: Children and parents – media use and attitudes.*

[123] Livingstone, S. and Pothong, K. (2022) *What do children think of EdTech or know of its data sharing? Read our survey findings*. Digital Futures Commission.

[124] Livingstone, S. and Pothong, K. (2022) *What do children think of EdTech or know of its data sharing? Read our survey findings*. Digital Futures Commission.

data practices shape digital environments and user experiences (Livingstone & Pothong, 2022).[125]

The Children's Society's *Good Childhood Report 2025* similarly highlights that young people often feel adults misunderstand their online experiences and focus too heavily on restricting access rather than supporting safe, informed participation.[126] It identifies "online accountability and digital literacy"[127] as a key priority, particularly in relation to appearance pressures, harmful content, and the impact of online environments on wellbeing, and points to the need for better understanding of how digital environments shape what children see and experience online.[128]

There is a clear need for media and digital literacy support that better explains:

- How digital services collect and process personal data;
- How algorithmic systems influence what users see and do, including content recommendation and moderation;
- How surveillance and tracking operate across services, including in educational technology environments; and
- How design choices and commercial incentives shape online environments.

Strengthening literacy in these areas would support more informed, empowered, and rights-aware participation in digital environments.

2. Where and how additional support for children's digital literacy should be provided

There is a need for more support to be available in a range of settings, including but not limited to: schools or childcare settings; community or youth spaces (for example libraries, youth clubs or local charities); parent or carer groups or networks; public services (such as family hubs, GP surgeries or community centres); faith or cultural groups; non-governmental online sources (such as websites, platforms or online communities); and Government websites.

[125] Livingstone, S. and Pothong, K. (2022) *What do children think of EdTech or know of its data sharing? Read our survey findings*. Digital Futures Commission.

[126] The Children's Society, *The Good Childhood Report 2025; Internet Matters, Children's Wellbeing in a Digital World Index Report*, 2025.

[127] The Children's Society, *The Good Childhood Report 2025; Internet Matters, Children's Wellbeing in a Digital World Index Report*, 2025.

[128] The Children's Society, *The Good Childhood Report 2025; Internet Matters, Children's Wellbeing in a Digital World Index Report*, 2025.

To better support children and young people to stay safe and feel supported online, the Government could focus on: providing clear guidance that children can use on their own; supporting parents and carers to support children online; working with online platforms and services that children already use; supporting youth organisations and community groups to help children online; making help or advice easy to access when something goes wrong online; and involving children and young people in designing support.

Support for children with additional needs should be grounded in Universal Design for Learning (UDL) principles, ensuring that digital safety information, tools and skills development are designed from the outset to be accessible, flexible and inclusive for all children. This includes providing multiple means of engagement, representation and expression, so that children with different learning, communication, cognitive or access needs can understand risks, build digital skills and participate safely online without being excluded or requiring separate or retrofitted provision.

Promoting high quality content

1. Determining what is meant by 'high quality' online content for children 13-16

Determining what constitutes "high quality" online content for children aged 13–16 should not be left to online services or their trust and safety teams. At present, too much of what children see online, and increasingly what is considered "appropriate" or "high quality" for them, is shaped unilaterally by technology companies through recommender systems, design choices, and content governance policies embedded within digital products and services. In practice, elements of childhood itself are being shaped by private services.

What children are exposed to, what is prioritised in feeds, what is treated as valuable or suitable content, even at what age childhood ends (with a digital age established at 13 from which children receive the same experience as adults), is driven by commercial incentives such as engagement and retention, rather than by children's rights, developmental needs, or broader societal values. Decisions about what "high quality" content means for children are therefore not purely technical questions, but social, cultural, developmental, and rights-based judgments.

5Rights' research demonstrates how commercial design within digital services can shape children's experiences online in ways that do not align with their best interests. *Disrupted Childhood: The Cost of Persuasive Design* (2023) highlights how engagement-driven design features are widely used to maximise attention, interaction, and data generation, including recommendation systems, autoplay, streaks, notifications, popularity metrics, and endless feeds.[129] The report

[129] 5Rights Foundation (2023) *Disrupted childhood: the cost of persuasive design*.

argues that these systems are primarily structured around commercial objectives rather than children's developmental needs or wellbeing.[130]

Similarly, *Pathways: How digital design puts children at risk* (2021) shows how recommendation systems and platform architecture can rapidly direct children towards harmful, extreme, or age-inappropriate content through automated pathways that are not transparent to children and are not necessarily aligned with their intentions or understanding.[131] It demonstrates that risk often arises not from isolated pieces of content, but from the design and commercial logic of the services themselves.[132]

Recent reporting has also raised broader concerns about the role of major platforms in shaping children's media consumption without the same editorial standards, accountability structures, or child-development oversight traditionally associated with children's broadcasting and educational media.[133]

For these reasons, determining what constitutes "high quality" content should be treated as a matter of public interest and democratic governance, rather than something defined primarily through service policies, recommender systems, or engagement metrics. Government has an important role in setting the overarching expectations and ensuring accountability, informed by children's rights standards and independent evidence.

At the same time, this requires meaningful participation from children and young people, particularly those aged 13–16, alongside parents and carers, educators, developmental experts, youth workers, and child rights organisations. Children must be central to this conversation, given their active participation in digital environments and their direct experience of how content is curated, recommended, and prioritised.

Ultimately, questions about what children should encounter online cannot be determined solely by commercial services whose incentives are shaped by engagement and retention. Standards for "high quality" content must instead be shaped through a rights-based, evidence-led, and society-wide approach, with clear regulatory oversight and meaningful participation from those responsible for supporting children's development and wellbeing.

Supporting positive online spaces and content for young people

[130] 5Rights Foundation (2023) *Disrupted childhood: the cost of persuasive design*.

[131] 5Rights Foundation (2021) *Pathways: how digital design puts children at risk*.

[132] 5Rights Foundation (2021) *Pathways: how digital design puts children at risk*.

[133] Davies, M. (2026) 'Think screens are making your child ratty? You may be right', *The Times*.

1. Taking further action

In considering what should be taken forward to support positive online spaces in practice, it is essential to start with enforcement of existing regulatory duties. Where clear obligations already exist, the central challenge is not the absence of principle but the lack of consistent and robust implementation. As highlighted in Steve Wood's recent report, there remains a persistent gap between regulatory expectations and enforcement in practice, which weakens deterrence and allows harm to persist even within an established legal framework.[134] Strengthening enforcement should therefore be the first priority in supporting safer and more positive online environments.

Where further clarity is needed, this should be addressed urgently to reduce ambiguity for both industry and regulators. Uncertainty around risk thresholds, compliance expectations and the meaning of safety-by-design in practice risks inconsistent application of standards and uneven protection for children across services. Clearer guidance is essential to ensure that positive online spaces are defined not only by what content is available, but by how systems operate and how risks are designed in from the outset.

Alongside stronger enforcement and clearer expectations, a code of practice on safety-by-design is critical to translating high-level duties into practice. This should set out clear, operational requirements for how services are designed, developed and deployed to reduce harm and support children's wellbeing from the outset. It would provide a consistent benchmark for industry, support regulators in taking action, and ensure that the creation of positive online spaces is embedded into system design rather than treated as an after-the-fact moderation issue.

This need is particularly acute given the current gap between the principle of safety by design in legislation and its operationalisation in practice. While the Online Safety Act embeds safety-by-design at a high level, there is limited clarity from regulators on how this should be implemented across systems, data practices and business models.

In response to this gap, 5Rights, alongside the Online Safety Act Network (OSAN) and a coalition of civil society organisations including the Molly Rose Foundation, NSPCC, End Violence Against Women Coalition (EVAW), Refuge, FlippGen, Glitch and the Internet Watch Foundation (IWF), has developed a Safety by Design Code of Practice.[135]

The Code provides a practical framework for translating statutory duties into operational requirements that reflect existing best practice and can be consistently applied across services.[136] It is intended to be submitted as part of

[134] Wood, S. (2026) *Impact of regulation on children's digital lives: Phase 2*. Digital Futures for Children, LSE.

[135] Online Safety Act Network, 5Rights Foundation and others (2026) Safety By Design Code of Practice.

[136] Online Safety Act Network, 5Rights Foundation and others (2026) Safety By Design Code of Practice.

the Network's consultation response and offers a model for how Government and Ofcom can give full effect to the Act's requirement that services are safe by design.

In practice, this means shifting the focus from isolated features towards systemic design choices, including recommender systems, data practices, engagement mechanics and default settings that shape user experience.

Effective implementation should be informed by a broad range of stakeholders, including children and young people, parents and carers, civil society organisations with expertise in children's rights and online harms, regulators, and industry. Crucially, children's lived experience should be central to shaping what "positive online spaces" mean in practice, alongside technical and regulatory expertise on system design and risk.

There is therefore a clear opportunity for Government to strengthen both enforcement and implementation by embedding safety-by-design as the organising framework for future guidance. Ofcom already has powers under section 41 of the Online Safety Act to produce codes of practice covering safety duties, which could be used to mandate a dedicated safety-by-design code.

With a targeted amendment, Government could further clarify this requirement in legislation and introduce a statutory definition of safety by design, ensuring consistency and accountability in how positive online spaces are delivered in practice.

There is also growing consensus behind this approach, reflected in parliamentary debate and increasing recognition of the limits of feature-based interventions alone. This consultation provides an opportunity for Government to adopt and implement civil society's co-drafted Safety by Design Code of Practice ensuring that platforms, including AI chatbot providers, are held accountable not only for content moderation, but for the underlying design of their systems. This would give effect to the broader ambition of delivering online environments that are safe, positive and rights-respecting by design and by default.

Chapter 5 – Supporting families

Parental controls

1. **Should parents have control over the online experiences of their children?**

   Parents and carers can play a crucial role in supporting children's online experiences, particularly younger children, and parental controls should be intended as a useful tool within that wider role. However, it is our view that parental controls cannot be relied upon as the sole or comprehensive safety measure in a child's digital life.

The question should be considered in light of children's rights, including their right to safety, protection from harm, privacy, participation, and access to developmentally appropriate experiences in line with their best interests. The balance between protection and autonomy should be proportionate and evolve as children grow and develop.

For younger children, a higher level of adult oversight will be appropriate, reflecting their developmental stage and need for stronger safeguards. For older teenagers, however, the emphasis should progressively shift towards greater agency, autonomy, transparency, and digital literacy, enabling them to make informed and meaningful choices within a framework of appropriate protections and safeguards.

Parental controls may therefore have a supporting role, but they are limited in practice. Parents are rarely able to exercise meaningful control over the full range of children's online experiences. Digital services are designed and operated by technology companies whose systems, interfaces, recommender algorithms and data practices are far more complex and adaptive than the tools available to most parents.

Research and user experience suggest that parental control tools can be difficult to configure, inconsistently effective, unevenly used, and sometimes capable of being bypassed by children. Additionally, parents may overestimate the protection such tools provide or underestimate their limitations.[137]

Parental mediation can help reduce some risks, particularly where parents have the time, skills, confidence, and ongoing engagement to support their children. However, its effectiveness is mixed and highly context dependent. Restrictive approaches may reduce exposure to some risks, but they can also limit children's opportunities, autonomy, trust, and development of digital skills.

Active dialogue and co-use may better support children's resilience, but this requires significant time and capacity from parents and cannot address systemic risks created by platform design, including commercial profiling, recommender systems, addictive engagement loops and other structural features of services.

For these reasons, parental controls should not be used to shift responsibility for children's online safety away from technology companies and onto individual families. They are one tool among many, but they cannot be the primary or sole mechanism for ensuring children's safety online. Their

[137] Ofcom (2025) *Children's media literacy report 2025: Children and parents – media use and attitudes;* *Kuldas, S., Sargioti, A., Staksrud, E., Heaney, D. and O'Higgins Norman, J. (2024) 'Are confident parents really aware of children's online risks? A conceptual model and validation of parental self-efficacy, mediation, and awareness scales', International Journal of Bullying Prevention, 6(3), pp. 252–266.*

effectiveness is likely to be greatest when embedded within a broader framework of safety-by-design, age-appropriate defaults, clear accountability mechanisms and ongoing parent–child dialogue.

Technology companies design the environments children access. They determine default settings, shape recommender systems, collect and process children's data, and decide how safety protections are implemented and enforced. As in offline contexts, where businesses have clear responsibilities to ensure age-appropriate access to restricted goods and services, digital services should likewise carry responsibility for ensuring that children encounter environments that are safe, age-appropriate and respectful of their rights.

A balanced approach is therefore needed. Parents may appropriately exercise oversight, particularly for younger children, but this must sit within a wider system in which services likely to be accessed by children are safe by design and by default, and in which platforms are clearly accountable for delivering and enforcing age-appropriate protections.

2. **Levels of parental control for children of different ages**

For younger children, such as an 11-year-old or less, a higher level of adult oversight is generally appropriate and in line with their best interests and their more limited capacity to understand commercial design, data practices, recommender systems and online risk. However, this should not mean relying solely on parents to protect children. Services likely to be accessed by children should be safe by design and by default, so that children are protected even where parents do not use, understand or maintain parental control tools.

For older teenagers, such as a 16-year-old, the balance should shift towards increasing autonomy, transparency and meaningful choices while still being protected by appropriate safeguards. At this stage, overly restrictive parental control may be less appropriate and less effective, and can risk undermining trust, privacy and engagement. Safeguards for older teenagers should therefore focus less on direct parental oversight of day-to-day activity and more on clear information, user agency, strong privacy protections, and firm limits on harmful design features.

Across all ages, there are important limits to parental control as the primary mechanism for children's online safety. They may play a supporting role, but they cannot address systemic risks created by recommender systems, commercial profiling, persuasive design, unsafe contact features, or business models that prioritise engagement over children's wellbeing.

Responsibility for children's online safety should therefore not be shifted onto parents. Technology companies design the digital environments children use, set the defaults, control the data practices, and determine how safety protections are implemented. They must carry enforceable duties to provide age-appropriate, rights-respecting and safe digital environments by design and by default.

In this sense, age should shape the balance between protection and autonomy, but not the underlying expectation that services themselves must be appropriate for children. The core principle should be that children's rights including their rights to protection, development, privacy and participation are embedded into digital environments from the outset, rather than managed mainly through parental control tools after risks have already been created.

3. Helping parents and carers more effectively use parental controls

Parents and carers would be better supported in the effective use of parental controls through a combination of clearer information, improved design consistency across services, practical support at the point of use, and stronger regulatory backing. However, parental controls should not be treated as the primary safeguard for children's safety online, nor as a substitute for safety-by-design obligations on services themselves.

At present, one of the main barriers is complexity and fragmentation. Parental controls differ significantly across devices, operating systems, apps and platforms. This makes them difficult for parents and carers to understand, set up, compare and maintain over time. Greater standardisation across core functions such as privacy settings, contact settings and purchasing controls would go a long way in making these tools easier to use, thus reducing the burden on families to learn and navigate multiple inconsistent proprietary systems.

Parents and carers would also benefit from better support at key "points of access". This could include clear, plain-language guidance at the point of device purchase or activation, including during phone or tablet set-up; in-product prompts from platforms that guide parents through relevant safety settings; and ongoing, accessible explanations of what parental controls do, what they do not do, and where their limitations lie. In this regard, guidance as well as the systems themselves should be practical, age-sensitive and easy to revisit as a child grows.

However, improving usability alone will not be sufficient. Parental controls are limited in practice. Their effectiveness can be inconsistent, they can sometimes be bypassed, and their use depends heavily on parents' technical literacy, time, confidence and ongoing engagement. They should therefore be understood as a supporting tool within a broader safety framework, not as a standalone solution.

Government, regulators and industry all have distinct responsibilities in this area. Government and regulators should ensure that companies comply with existing duties to protect children, including duties relating to safety by design and age-appropriate design, rather than placing disproportionate responsibility on parents and carers. Where parental control tools are offered, regulators should also expect minimum standards for their effectiveness, usability, transparency and accessibility.

There may also be a role for technical standards, developed in collaboration with industry, civil society, children's rights experts and regulators, to improve consistency and interoperability across tools. However, any such standards must be rights-respecting by design. They should avoid over-blocking, excessive surveillance, or unintended restrictions on children's rights to information, participation, privacy and development.

Ultimately, the most effective approach is a shared one: better-designed and more standardised parental tools, clearer guidance at the point of purchase and point of use, and, most importantly, stronger enforcement of legal duties on platforms to ensure that safety is built into services themselves. Parents and carers should be supported, but they should not be expected to compensate for unsafe design or to act as the main enforcement mechanism for children's online safety.