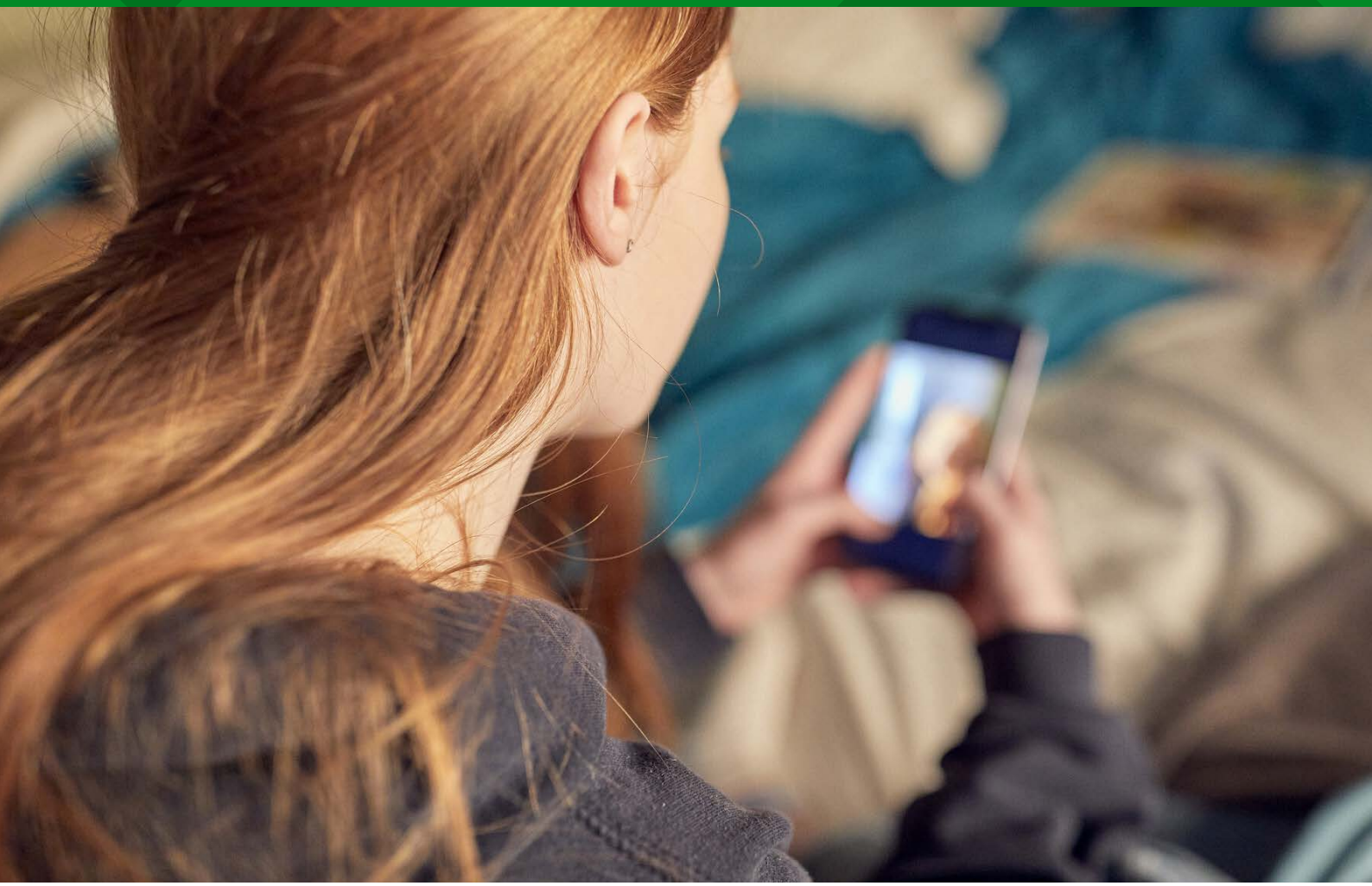
NSPCC Learning

# Evidence Review on Online Risks to Children

## Executive Summary

Dr Jo Bryce, University of Central Lancashire
Professor Sonia Livingstone, London School of Economics
Professor Julia Davidson, University of East London
Beth Hall, University of Central Lancashire
Jodie Smith, University of Central Lancashire

**November 2023**



# Executive Summary

## Rationale

This review provides an outline of the current online risk landscape for children in the UK. The NSPCC commissioned this study to understand what is known about the risks and harms that children face online, and any new evidence that has emerged since 2017 when the UK Council for Child Internet Safety (UKCCIS) published a study in this area.[1]

The review represents the online risk landscape in the period immediately before the Online Safety Act (2023) passed into law. It captures the various risks faced by children in a largely unregulated online environment, before providers of user-to-user online services were legally required to take responsibility for their users' safety. Its timing means that the review maps a baseline from which to assess subsequent change.

## Topics covered

One of NSPCC's priorities for research, policy and influencing work is child sexual abuse. For this reason, the main focus of the evidence review is on children's exposure to online sexual risk and harm.

In addition, a light touch review was carried out of two related topics:

1. Children's exposure to other categories of online risk classified in the Online Safety Act as 'primary priority' content (e.g., online pornography, content that encourages suicide, self-harm and eating disorders) and 'priority' content (e.g., cyberbullying, hate crime). This overview helps place the online sexual risks in perspective and in their wider context, enabling a comparison of their relative scale and the harm they can cause.
2. Technological design features and tools that can increase or decrease children's exposure to risk. Outlining the role that technology can play in modifying levels of risk demonstrates that risk is not an inevitable outcome of being online: it is also influenced by the choices that technology companies make in designing their platforms and implementing safety tools.

The review focuses on victimisation and does not consider the wider literature on offending, prevention and treatment.

1 Livingstone, S., Davidson, J., Bryce, J., Batool, S., Haughton, C. & Nandi, A. (2017) *Children's online activities, risks and safety: A literature review by the UKCCIS evidence group.* UKCCIS. www.lse.ac.uk/business/consulting/assets/documents/childrens-online-activities-risks-and-safety.pdf

## Methodology

Online sexual risks to children were reviewed using a modified Rapid Evidence Assessment (REA) methodology. A systematic and structured search strategy was used to identify academic literature published since 2017 on four categories of sexual victimisation. The review included UK studies and studies using samples from other countries where they added to understanding of victimisation within the UK. A search was also undertaken of grey literature to identify relevant reports by other stakeholders published over the same period.

A light touch review was conducted of other online risks (categorised as 'primary priority' and 'priority' content in the Online Safety Act), and of technical tools that can reduce or increase exposure to risk and harm. This consisted primarily of a 'review of reviews' (REAs, literature reviews). It was supplemented by limited searches for newer academic publications in these areas, together with an examination of research by Ofcom and organisations working in the field of online safety and child protection. For the purpose of the report and data searches, children were classified as anyone under the age of 18.

# Children's exposure to online sexual risks

## Nature of online sexual risks

The review argued that there is value in distinguishing between consensual and non-consensual online sexual interactions, and between different forms of abuse based on the age of those involved. This is because consent and the relative age of those participating in the interactions can have different legal implications, may be experienced differently by participants, and can lead to different outcomes. The review found that these distinctions are not always made in research about online sexual risks and harms.

Working definitions were presented for four categories of online sexual victimisation. For the purposes of this review, two were defined solely as peer-to-peer abuse; one solely as adult-to-child abuse; and the last as the result of actions by either children or adults:

- Online sexual harassment: unwanted online sexual conduct by a child towards another child.
- Intimate image abuse: situations where a child takes or shares sexual images of another child without the consent of the person depicted.
- Technology-assisted child sexual abuse: situations where an adult sexually exploits a child online or uses technology to facilitate the offline sexual abuse of a child.
- Child sexual abuse material: imagery or videos that show a child engaged in (or depicted as being engaged in) sexual activity.

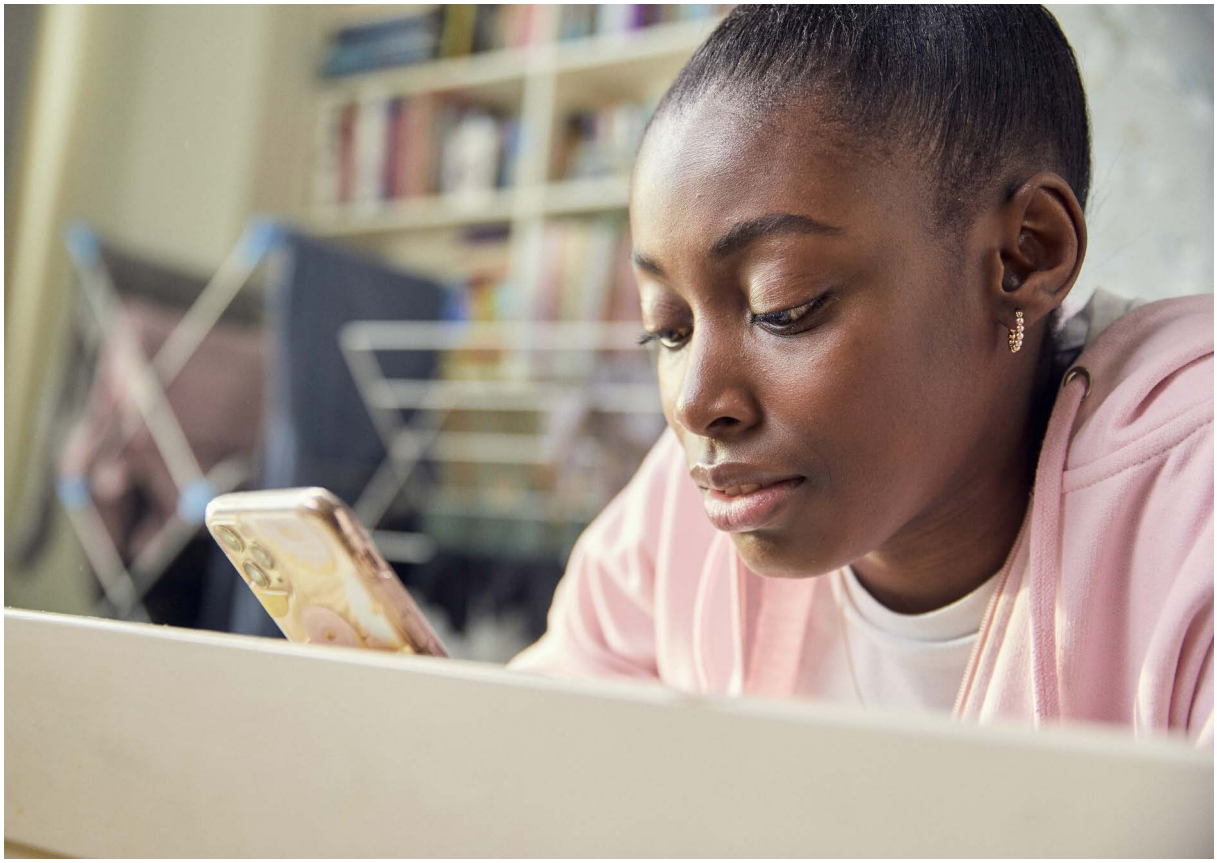
These categories provided a framework for organising and presenting the evidence on online sexual risks.

## Scale of online sexual risks

There has been a general lack of academic research in the UK examining the prevalence of online sexual risk and harm to children since the UKCCIS report was published in 2017. Evidence is, therefore, mainly drawn from non-UK samples of children and studies from the grey literature.

The evidence suggests that a sizeable minority of children – at least one in 20, but potentially up to a quarter – have encountered sexual risks when they have been online. The most common types of sexual risk were online sexual harassment by peers (8–26 per cent prevalence) and technology-assisted child sexual abuse by adults (5–25 per cent prevalence). Intimate image abuse by peers was relatively less widespread (e.g., 5–11 per cent prevalence for receiving unwanted sexual images). The vast amount of child sexual abuse material in circulation evidenced through a variety of data sources suggests there are many undetected victims of sexual exploitation.

It is not possible to confidently establish whether there has been a rise or fall in sexual online victimisation in the years since 2017. Various data sources indicate upward trends over time in the incidence of technology-assisted child sexual abuse and in reports of child sexual abuse material (including first-person produced images). These trends may result from a variety of factors, such as developments in scanning and detection tools, and not solely reflect an increase in levels of perpetration and victimisation.



## Platforms on which online sexual risks occur

It is difficult to determine the platforms on which exposure to sexual risk and harm tends to occur. In part, this is because the relevant data is held by platforms and is hard to obtain. In addition, self-report studies with children rarely collect information on where risks are encountered.

What little data exists suggests that networking platforms that are popular with children (e.g., Snapchat, Instagram) are the most frequent places where children are exposed to online sexual harassment, intimate image abuse and technology-assisted child sexual abuse. There is also evidence that both the open and dark web are important channels for distribution of child sexual abuse material, with livestreaming and first-person produced sexual imagery increasingly involved in the production of such material. However, it is not easy to deduce which platforms host relatively more child sexual abuse material. This is because some platforms make proactive efforts to detect it, whereas others (particularly those that deploy end-to-end encryption) choose not to and, consequently, report fewer instances of it on their platform – potentially misrepresenting how much sexual abuse they actually host.

It is important to develop robust evidence on the prevalence of exposure to risk and harm on individual platforms, as this is currently lacking. Greater empirical understanding is also needed on sexual online risks in gaming environments and on direct messaging services, and how platforms' design features facilitate (or prevent) behaviours associated with the four types of online sexual victimisation.

## Children most likely to be exposed to sexual risk and harm

There has been some development of evidence examining factors that increase children's vulnerability to online sexual risk and harm since the last review in 2017. However, this has been limited in the UK. Recent research confirms earlier findings indicating that children's gender and age are important factors linked to vulnerability.

Girls encounter every category of online sexual risk more frequently than boys: more girls than boys experience online sexual harassment and most types of intimate image abuse; more experience technology-assisted child sexual abuse by an adult; and girls are more commonly depicted in child sexual abuse material. There is also evidence that victimisation through intimate image abuse has gendered dimensions and can have more severe psychological and social impacts for girls.

Older adolescents are more likely than younger children to report experiences of online sexual harassment, intimate image abuse, and technology-assisted child sexual abuse. On the other hand, it is prepubescents and younger children who are most commonly depicted in child sexual abuse material, suggesting that measures like self-report victimisation surveys may be failing to capture the extent of victimisation of younger children.

There is less understanding of the extent to which other demographic, psychological, social, and environmental factors (or their inter-relationships) influence exposure to online sexual risk, or the likelihood that risk will lead to harm. More evidence is also needed about how vulnerability factors operate across multiple domains of online (and offline) risk and harm, leading to the victimisation of the same child in multiple ways and the escalation of harm over time.

## Children's responses to encountering online sexual risks

Children often respond to online sexual risk by deleting messages or images, or by using technical tools to block unwanted communication. Making reports to platforms and seeking help from adults appears to be less common. A minority of children respond by engaging with the perpetrator, but one of the most common responses is to do nothing. In some cases, this is because children think that nothing can be done and see these risks as an inevitable part of being online. In the case of technology-assisted child sexual abuse, it may result from children not recognising a situation as abusive or not realising that images of them are being taken and shared.

There are various barriers to children reporting such experiences, including: concern about the actions that will be taken, or judgements made about them; self-blame and shame; and fear that their reports will not be believed or taken seriously.

> "I was being pressurised into sending sexual photos and videos of myself and was threatened if I didn't. They would go on and on at me when I said no but would carry on with the threats."
>
> Girl, aged 13 (Project deSHAME, 2017)

### Outcomes and impacts for children who encounter sexual online risks

There is a paucity of recent evidence from the UK on this topic. Existing research suggests that there are similar emotional, psychological and social outcomes associated with victimisation across the four categories of sexual risk and harm. These experiences have significant and negative impacts on children's mental health and social relationships.

It is important to note that the psychological impacts of technology-assisted child sexual abuse do not differ significantly from those associated with offline child sexual abuse. The potential production of child sexual abuse material as a result of online abuse, and the possibility of revictimisation through the ongoing distribution of images online, may even exacerbate these negative outcomes.

### Additional gaps in knowledge about sexual online risks

While the current evidence base is sufficient to draw broad conclusions about the landscape of online sexual risks for children today, it could be strengthened with new research that provides a more specific and detailed examination (and distinction between) the four types of online sexual victimisation.

More evidence is also needed to understand how recent technological developments have changed the nature of online sexual risks. Consideration should be given to the following: the way that livestreaming affects interactions between peers, and between children and adults; the ways in which generative artificial intelligence and other technology that may be used to create deepfakes, or synthetic and non-photographic abuse images of children, can impact on the production and distribution of child sexual abuse imagery; and the ways in which technology facilitates interaction between perpetrators on the open and dark web.

## Children's exposure to other types of online risk

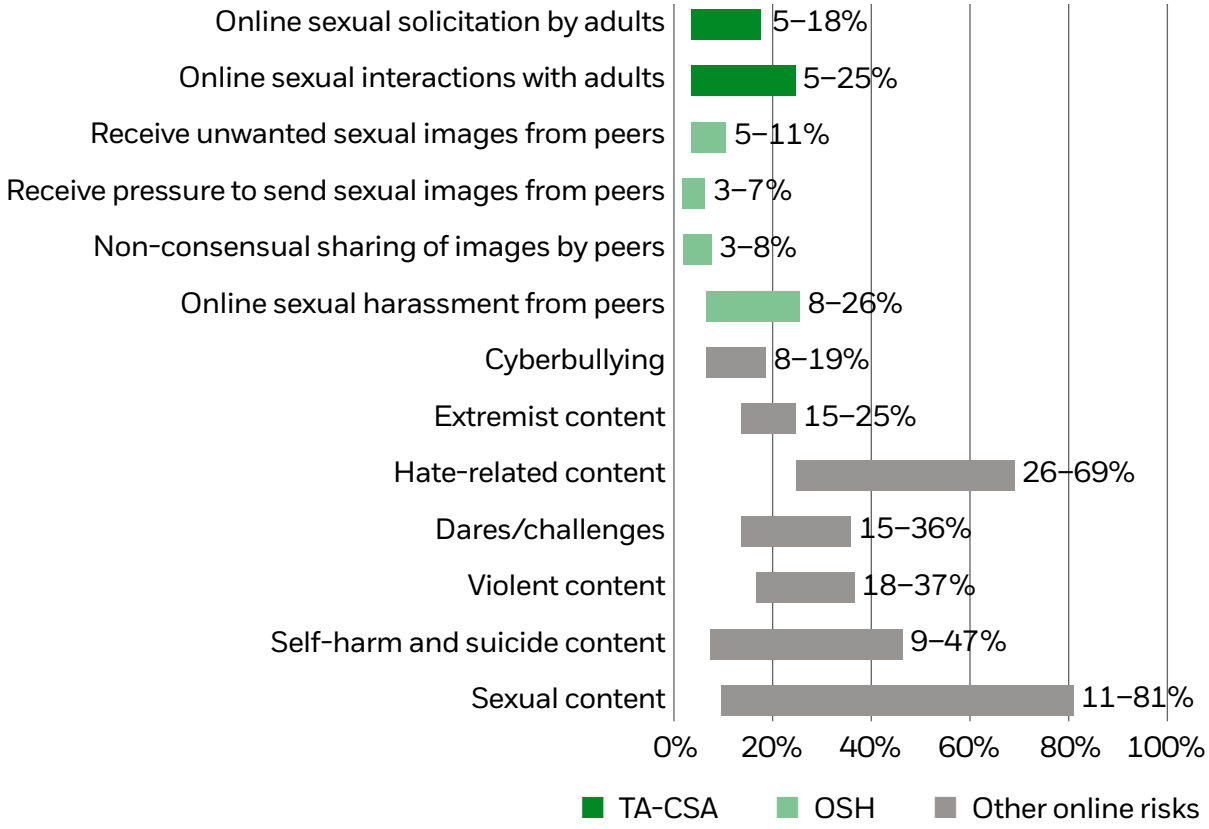
To provide a fuller picture of the online risk landscape for children in the UK before the introduction of regulation, this review also examined the evidence base related to other categories of online risk. A light touch review was undertaken, focusing on the risks classified in the Online Safety Act as 'primary priority' (e.g., online pornography, content that encourages suicide, self-harm and eating disorders) and 'priority' content (e.g., cyberbullying, hate crime). It additionally covered content that promotes extremism and radicalisation, as an example of illegal content not covered by the sexual categories already discussed.

### Relative prevalence of different online risks

The review identified a variety of prevalence figures for children's exposure to 'primary priority' and 'priority' content in the UK. The evidence consistently indicates that a minimum of one-in-12 children has experienced 'primary priority' or 'priority' content.

The evidence also shows that children are more likely to be exposed to content risks than most types of online sexual risk (see Figure 1).

**Figure 1: Prevalence of children's exposure to sexual online risks and a range of other online risks**




## Responses to other categories of online risk

The small amount of self-reported data on children's responses suggests that when children are exposed to these types of content, many tell someone about it, and fewer use technical tools to block content or change their privacy settings. Reporting to platforms is less common, with some children not knowing about this function or feeling it would be pointless to make a report.

The likelihood of telling someone about the experience seems to be higher than for children who experience online sexual risks.

> "I thought it [excessive exercise and disordered eating] was normal behaviour and so it kind of manifested into my own behaviour because I thought all these women are doing it, and I wasn't seeing any healthy behaviours because I was so isolated [because of Covid-19 lockdown]."
>
> Girl, aged 17 (Gill et al, 2022)

### Outcomes and impacts of exposure to other categories of online risk

With the exception of cyberbullying, which has a relatively well-developed evidence base, there is only a modest amount of evidence about the outcomes and impacts of exposure to, and engagement with, 'primary priority' and 'priority' content. The available evidence suggests that the emotional and psychological impacts have some commonalities with those linked to online sexual victimisation. There are also additional attitudinal and behavioural impacts associated with these risks (e.g., health-related outcomes, aggressive behaviour).

## How technological design features can increase or decrease risks

The review examined how a variety of technological features and safety tools can increase or reduce children's exposure to online risk and harm. The review presents the rationale for why the features and tools can make a difference, without attempting to quantify their effect on children's likelihood of encountering risk.

Looking at the associated technology highlights the fact that risk is not an inevitable outcome of being online: risk can be enhanced or mitigated through the design choices of platforms and their decisions regarding the implementation of robust safety tools.

### Features and tools that can increase online risk and harm

Design features that aim to increase user engagement can increase exposure to online risk and harm. The use of recommender algorithms, for example, can provide focused and intense exposure to harmful content and facilitate harmful interactions with other users. This is particularly problematic for children who may be at increased risk of harm due to specific vulnerability factors (e.g., those who have a history of eating disorders or suicidal ideation). The quantification of social activity and use of popularity metrics is also problematic: these features exploit psychological and social needs that are developmentally important for children but simultaneously increase children's exposure to online risks.

Greater empirical understanding is needed of how the design features of platforms interact with a variety of psychological, social and environmental factors to increase risk and harm to children.

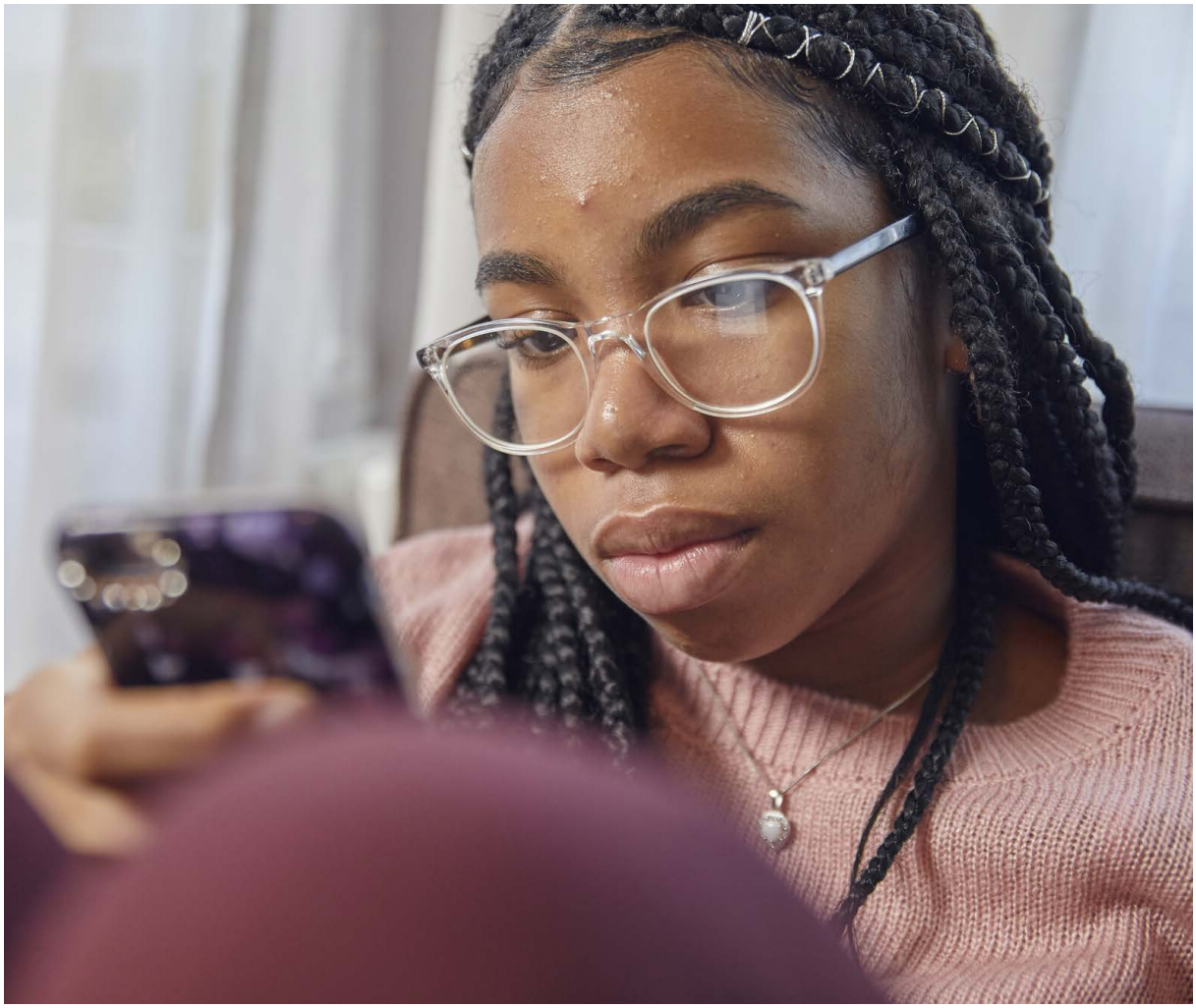
### Features and tools that can reduce online risk and harm

The application of age-assurance methods and a variety of parental control tools can potentially minimise children's exposure to some types of risk – but only when these are robust, consistently applied, easy to implement, and can be adapted to different stages of children's development.

Many platforms equip users with tools to block, mute and report unwanted interactions, or to flag and report harmful content. If used, these can potentially prevent children's continued exposure to risks they have already encountered. There are various barriers to children's usage, however, with some evidence that they are not always easy for children to locate and use effectively, and a degree of mistrust as to whether platforms would respond quickly and appropriately to children's reports.

Content moderation and detection tools are widely used and can be effective in identifying illegal content and behaviour. Their use can reduce children's experiences of risk and harm, assuming that swift and appropriate remedial action is taken by platforms after identification. There are, nevertheless, various challenges to effective moderation, especially given the volumes of material that need moderating and the real-time nature of user-to-user interaction on many platforms. There is also a lack of knowledge about the training datasets, feature detection approaches, and learning models used in creating these tools. Data about detection rates and outcomes is often lacking, making it difficult to determine how well these tools work.

The review also examined emerging challenges for content moderation associated with the wider roll-out of end-to-end encryption (E2EE), which will reduce the ability of platforms and law enforcement to detect technology-assisted child sexual abuse and child sexual abuse material.



## Recommendations about future research

The following recommendations focus on future research that can advance understanding of the overall online risk landscape for children. A fuller set of research recommendations that cover methodological considerations and offers suggestions for building the evidence base on online sexual risks appears in Chapter 9 of the report.

### Recommendations: research about prevalence of online risks

There is a need for more research examining the prevalence, experience and impacts of children's exposure to online risk and harm in the UK. Development of robust measures and datasets as close as possible to the enactment of the Online Safety Act is essential and should be a priority for policy makers and researchers.

These should be designed to enable effective measurement of the frequency of exposure to online risk across the different categories covered by the Online Safety Act. They should address the specific complications associated with measuring online sexual risk (e.g., drawing clear distinctions between adult and peer perpetrators, identifying age gaps between peers who interact sexually online, and determining the degree to which interactions were perceived to be consensual). This should be followed up with questions about the resulting experience and intensity of different potential harms, and identify the platforms involved. Questions should also be asked about related help-seeking behaviour, and data collected examining a broad set of vulnerability indicators. There should also be items examining the opportunities provided by the online environment (e.g., children's rights to information, education, participation) to inform proportionate policy responses.

Data should be collected using a systematic and longitudinal methodology to allow an examination of trends over time and more powerful statistical analyses of the relationships between the measured variables.

There are clear methodological, ethical and resourcing challenges associated with such a research undertaking. This is nevertheless vital, as it will enable the development of robust baseline measures against which change over time and the impact of the Online Safety Act can be assessed. It will allow an examination of the extent to which regulation has the intended effect of reducing risk; and whether this occurs at the cost of the opportunities for children provided by the online environment (e.g., online education, creative production, civic participation).

### Recommendations: Research about vulnerability to online risks

It is important to develop greater understanding of the demographic, psychological and environmental factors that increase children's vulnerability to different types of online risk and harm. More evidence is needed on how these factors can lead to the victimisation of the same child in multiple ways (polyvictimisation); and on how exposure to online risk and harm relates to offline victimisation (e.g., physical violence, sexual exploitation). Greater understanding in these areas will inform the development of more effective strategies to prevent harm, build resilience and develop media literacy for children.

## Recommendations: How technology companies can contribute to building the evidence base

The information provided in transparency reports is currently very limited, making it difficult to draw conclusions about platform safety or changes over time. Platforms should provide more detailed information in the risk assessments required by the regulator. This should include:

- The number of reports and detections across different categories of risk and harm to children, including data on:
  - The amount of child sexual abuse material on their services as a proportion of pages/posts/content viewed.
  - The number of contacts between adult and child users, which are either reported or detected as inappropriate or indicative of technology-assisted child sexual abuse.
- Levels of usage of safety and reporting tools by child users; and the actions taken by the platform in response (including number of takedowns and response times).
- The moderation systems and detection tools used by the platform to protect children from exposure to harmful content.
- The type of age-assurance processes implemented.
- The level of effective liaison with law enforcement, helplines and organisations, such as the Internet Watch Foundation (IWF) and the National Center for Missing & Exploited Children (NCMEC).
- Processes by which safety tools, reporting processes, content moderation, detection tools, and age-assurance processes are evaluated.

The provision of data of this type will also enable a more effective assessment of levels of risk and harm on different platforms; and of the efficacy of platforms in protecting children from online risk and harm.

## Recommendations: How policymakers and regulators can contribute to building the evidence base

Robust datasets should be commissioned to enable more effective understanding of children's experience of online risk and harm in the UK over time. A stronger evidence base will provide essential knowledge for the work of other stakeholders who have responsibility for protecting children online and developing related prevention and response strategies.

Consideration should also be given to building a framework to assess the ways in which technological developments (e.g., generative AI, immersive technologies) potentially influence children's exposure to online risk and harm. This should be created in collaboration with platforms and the safety tech industry, as well as other stakeholders. Having this in place will ensure that policy and regulatory systems are able to effectively respond to the risks associated with emerging technologies.

Regulators should produce and enforce robust and effective regulatory standards that ensure platforms provide:

- Detailed transparency reports that address the recommendations described above.
- Comprehensive risk assessments, including information about recommendation algorithms and content moderation, as well as the efficacy of reporting and detection tools.

## Recommendations from the NSPCC

The findings from the evidence review make clear how the absence of regulation before implementation of the Online Safety Act left children exposed to online risks that were often avoidable.

This review establishes a baseline for the risk landscape as it stands as the Act passes into law. If the research recommendations presented above are followed, the next few years will enable an assessment of any changes linked to the introduction of the new regulatory regime.

Based on the review findings, the NSPCC has prepared a set of additional recommendations for the technology sector, Ofcom and the government to consider.

### Recommendations to technology companies

Safety-by-design as an approach must be built into platforms' development processes. Companies must consider the views and needs of children and survivors when making product decisions, using participatory design approaches and co-production. This is consistent with the child rights framework specified by UN General Comment No.25 on Children's Rights in Relation to the Digital Environment.[2]

Companies need to take the initiative to significantly reduce the levels of harm on their platforms and not wait for Ofcom to use their new regulatory powers. This includes:

- Taking proactive steps to prevent harm through safer design, such as: building robust safety features; applying age-assurance technologies; designing age-appropriate content and spaces; regularly monitoring the efficacy of these features and being transparent about their use.
- Committing to resourcing the research and development of solutions that could mitigate risks to children. This includes investing in technology and engineering resources for the development of new and increasingly effective tools that could examine large volumes of material, including still and moving imagery, text, as well as real-time interactions, and cover a range of languages. These tools should be applied on all their products, including in end-to-end encrypted environments that could potentially harbour child sexual abuse.
- Providing the necessary degree of human moderation to check automated decisions are accurate and that appropriate and prompt action is taken to safeguard users and children. Moderators should be fully trained and supported to perform their work.
- Proactively detecting illicit behaviour and content and taking swift remedial action to permanently remove associated content, accounts, and users. This includes:

2 United Nations (2021) *General comment No. 25 (2021) on children's rights in relation to the digital environment.* UN Doc. CRC/C/GC/25. www.ohchr.org/en/documents/general-comments-and-recommendations/general-comment-no-25-2021-childrens-rights-relation

- acting rapidly on reports from trusted external services like Report Remove on harmful content.
- contacting law enforcement bodies, such as the UK's National Crime Agency and the National Center for Missing & Exploited Children where appropriate to report safeguarding concerns on their platforms.
- collaborating with other technology companies, and sharing information with them, to tackle abuse that starts in one location and moves across platforms (cross-platform risks).

Platforms should use a victim-centred approach in their reporting and complaints mechanisms. This includes:

- Making it possible for children to make a report retrospectively if they cannot relay their concerns in real time or immediately after their encounter. Children should be offered appropriate options to describe their experience and be reassured that their report would be taken seriously, and their confidentiality maintained.
- Ensuring reporting methods are child-friendly and easy for children to locate and use if they wish to make a report or complaint, or flag harmful behaviour. It should be a fundamental requirement on all platforms to promote a positive culture of reporting and offer prompt and effective recourse to users.

## Recommendations to Ofcom

Ofcom must implement the Online Safety Act in full and without delay. We are pleased with the speed at which Ofcom has published the draft codes for illegal harms and look forward to responding. They must actively listen to the evidence provided by non-governmental organisations and independent researchers and take firm action against companies that fail to comprehensively assess the risks posed by their platforms.

Ofcom should use its new regulatory powers to:

- publish up to date codes of practice which demonstrate the steps which platforms can take to protect children. It should use its evidence-gathering powers to ensure that innovative solutions become the baseline of what is expected of companies.
- provide clear age-assurance guidance on how to safely enforce minimum age thresholds.
- issue specific guidance on how companies can work together to tackle cross-platform risks and prevent widespread distribution of child sexual abuse material on undetectable servers, such as private messaging platforms and the dark web.
- provide guidance on how platforms can comply with their duty to mitigate and manage the risk of the service being used for the commission or facilitation of online abuse. This includes behaviours by abusers that do not meet the criminal threshold, such as signposting other abusers to illegal content on third-party messaging apps and the dark web; or sharing child abuse videos that are carefully edited to fall within content moderation guidelines.

Ofcom should put children's voices at the heart of online safety regulation. Regulating the online world to protect children requires an accurate understanding of children's digital lives. Since children's online experiences are often best understood by speaking directly to children, Ofcom should establish mechanisms to meaningfully consult children and young people on how to develop and implement the new online regulatory regime.

Ofcom should keep to its implementation roadmap and develop guidance on tackling online violence against women and girls – which the Government committed to as part of the Online Safety Act – at the latest by Spring 2025. This should be done by:

- listening to girls' experiences, taking into account the views and needs of children and survivors, and consulting with specialist organisations and charities who support girls and women.
- providing technology companies with clear directions on how to prevent violence against girls.
- putting monitoring and evaluation mechanisms in place to ensure companies take up the measures and report on their effectiveness for girls who use their platforms.

## Recommendations to Government

Government should review online safety legislation on a rolling basis to ensure that it is fit for purpose and evolves with the changing risk landscape for children.

Government should continue to show global leadership by funding the development of solutions by the safety technology sector. The Safety Tech Challenge Fund is a positive initiative to support the development of proof-of-concept tools that can detect child sexual abuse material across end-to-end encrypted environments while upholding privacy. Government should commit to supporting the next stage of scaling up and enabling adoption of effective technologies.

Government should continue to collaborate with our international partners to inspire a global response to tackling technology-facilitated child sexual abuse and ensure that there is regulatory harmony with the UK's online safety legislation.

Government must ensure survivors of online abuse are entitled to, and able to access, effective support as victims of abuse. The Victims and Prisoners Bill should be strengthened to:

- Expand the definition of victim to recognise all children, not just those affected by domestic abuse.
- Place a duty on the relevant authorities to consult with providers of children's services to ensure support is in place for child victims.
- Ensure relevant authorities commission sufficient and specific support for children and young people who are victims of crime.
- Promote the establishment of the Child House model and provide central Government funding for the provision.

Government should fund public awareness campaigns and educational material aimed at both children and the adults around them (including parents, carers and professionals) to raise awareness of the types of abuse children experience online and the support available. These resources will help create an improved environment in which children can safely disclose the abuse they have been subjected to and access help.

NSPCC Learning

NSPCC Learning is here to provide you with all the tools, training and resources you need to protect the children you work or volunteer with.

We keep you up to date with the latest child protection policy, practice and research. We deliver expert elearning courses and face to face training for your organisation. And we provide bespoke consultancy, sharing our knowledge of what works to help you deliver services for children and families.

With your support, working together, we can protect more children right across the UK.

**nspcc.org.uk/learning**