NSPCC
Learning

Evidence Review on Online Risks to Children

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

Contents

Acknowledgements

The NSPCC is an independent charity funded overwhelmingly by voluntary donations. The NSPCC is very grateful for the generous support from the Oak Foundation who made this research possible.

Glossary

Child/children	**Someone aged under 18.**
Child Sexual Abuse Material (CSAM)	Imagery or videos that show a child engaged in (or depicted as being engaged in) sexual activity. [See Section 2.2.4]
End-to-End Encryption (E2EE)	Technology that enables users to exchange messages whose content can only be read by the sender and recipient.
Intimate Image Abuse (IIA)	The taking or sharing of sexual images without the consent of the person depicted. In this report, this term is only used in relation to peer-to-peer interactions. [See Section 2.2.2]
Non-photographic Abuse Imagery (NPAI)	Animated depictions of children engaging in sexual activity or being sexually abused.
Online Sexual Harassment (OSH)	Unwanted sexual conduct involving a variety of online behaviours (e.g., sexual and/or gender degrading comments, blackmail, control, coercion or humiliation). In this report, this term is only used in relation to peer-to-peer interactions. [See Section 2.2.1]
Platform	A service that enables users to access user-generated content and facilitates user-to-user interaction (e.g., social media or messaging apps).
Technology-Assisted Child Sexual Abuse (TA-CSA)	Situations where children are sexually exploited online or where technology is used to facilitate their offline sexual abuse. In this report, this term is only used in relation to perpetration by adults. [See Section 2.2.3]

List of acronyms

APPG	All-Party Parliamentary Group
BBFC	British Board of Film Classification
CAID	Child Abuse Identification Database
CCCP	Canadian Centre for Child Protection
CCDH	Center for Countering Digital Hate
DCMS	Department for Culture, Media and Sport
DSIT	Department for Science, Innovation and Technology
IWF	Internet Watch Foundation
NatCen	National Centre for Social Research
NCA	National Crime Agency
NCMEC	National Center for Missing and Exploited Children
NSPCC	National Society for the Prevention of Cruelty to Children
ONS	Office for National Statistics
OSB	Online Safety Bill
UKCCIS	UK Council for Child Internet Safety
UN	United Nations
VAWG	Violence Against Women and Girls

Chapter 1
Introduction

In 2017, the Department for Culture, Media and Sport (DCMS) commissioned the UK Council for Child Internet Safety (UKCCIS) to review the available research examining the online risks and harm to which children in the UK were exposed (Livingstone et al, 2017). This aimed to build upon earlier reviews conducted in 2010 and 2012 (Livingstone et al, 2012; Spielhofer, 2010), and provide an updated evidence base to inform work on the Online Safety Bill (OSB).[1]

The present review builds on the entirety of this earlier work. It presents the evidence on online risks and harms to children that has emerged in the six years since the last review was published. Crucially, this updated picture of the evidence base represents the period immediately before the Online Safety Act passed into law. The review, therefore, serves as an overview of the online risk landscape for children in the UK ahead of the changes associated with implementation of the Act.

The Online Safety Act received Royal Assent on 26 October 2023, marking a step change in the protection of children online. The Act introduces a new regulatory framework that will make technology companies responsible for their users' safety on their platforms and hold them to account for protecting children from illegal and the most harmful types of content. Technology companies will be required to work with Ofcom, as independent regulator, to conduct risk assessments that evaluate platform safety and must take steps to mitigate risks and harm and provide clear and accessible information for users about the actions they are taking.

By capturing evidence of the online risks facing children at a point immediately prior to the enforcement of the Act, this review maps a baseline from which to assess subsequent change.

[1] The Bill was originally part of the remit of the Department of Culture, Media and Sport (DCMS), but later came under the Department for Science, Innovation and Technology (DSIT).

1.1 What has changed since 2017?

The years following publication of the 2017 review have seen a variety of technological, social and political changes that have influenced children's online experiences, including their exposure to risk and harm.

During this period, smartphone ownership and social media use among children continued to rise, with 59 per cent of 3–13-year-olds possessing a smartphone and 63 per cent using social media in 2022 (Ofcom, 2023). A range of new social media platforms and livestreaming became popular with children: TikTok, for example, which was barely known in 2017, is now used by 43 per cent of 3–17-year-olds (Ofcom, 2023). During the same period, immersive online social and gaming spaces (e.g., virtual and augmented reality environments) also developed and started to become more mainstream.

Changes such as these, and ongoing concern about children's online safety, created the impetus to require technology companies to have a duty of care towards their users. Proposals for regulation were first introduced in the Internet Safety Strategy Green Paper (2017) and have been under development throughout the years covered by this review (Woodhouse, 2022). Since 2017, the prospect of regulation and accountability has been present and imminent: children in the UK explored the online world against a backdrop of proposals, counterproposals, inquiries, consultations, debates and scrutiny of the Online Safety Bill.

In the meantime, the UK's socio-cultural landscape has also experienced change. Public and policy debates have emerged about the level of violence and sexual harassment experienced by women and girls in online and offline spaces, leading to the development of the government's Violence Against Women and Girls (VAWG) strategy (HM Government, 2021). Greater awareness of the experience and impacts of prejudice and discrimination on the everyday lives of many UK citizens has led to greater recognition of the need to increase equality, diversity and inclusion in all aspects of society.

The COVID-19 pandemic had a significant impact on everyday life and society during this time, leading to increased use of the online environment for a variety of daily activities, which has widened access and engagement. This period saw children in particular spending more time online (Trott et al, 2022), with evidence suggesting that this was associated with an increase in their exposure to a variety of online risks and harm and a reported rise in levels of child sexual abuse material (IWF, 2021; Romanou & Belton 2020). The same period also saw wider online dissemination of misinformation and disinformation, which exposed a larger audience to beliefs and worldviews that have potentially harmful impacts on the attitudes and behaviour of individuals, as well as wider societal consequences.

It is within this context that the NSPCC commissioned this review to examine the available evidence base related to online risk and harm published since the 2017 UKCCIS report. Since online sexual abuse is a key focus of the NSPCC's research, policy and influencing work, this review pays particular attention to children's experiences of online sexual risk and harm. The review also provides a light touch review of other types of risk and harm (e.g., cyberbullying), similar to those recently reviewed by NatCen (Hudson et al, 2022), and examines the ways in which technological design features can increase or decrease children's online safety.

1.2 Theoretical approach

The CO:RE 4C classification was developed as a comprehensive, child-centred and evidence-based framework that maps the different types of online risk that children can face (Livingstone & Stoilova, 2021). The model is a helpful starting point for understanding the range of risks covered in this report, and for recognising that risks are relational: they result from the dynamic interaction between the child's agency and that of others operating in the digital environment.

Notably, online sexual risks – which are the main focus for this review – cut across all four risk categories in the model (content, contact, conduct and contract), as shown in Figure 1.

Figure 1: The CO:RE classification of online risk to children (Livingstone & Stoilova, 2021)

	Content Child engages with or is exposed to potentially harmful content	**Contact** Child experiences or is targeted by potentially harmful *adult* contact	**Conduct** Child witnesses, participates in or is a victim of potentially harmful *peer* conduct	**Contract** Child is party to or exploited by potentially harmful contract
Aggressive	Violent, gory, graphic, racist, hateful or extremist information and communication	Harassment, stalking, hateful behaviour, unwanted or excessive surveillance	Bullying, hateful or hostile communication or peer activity e.g. trolling, exclusion, shaming	Identity theft, fraud, phishing, scams, hacking, blackmail, security risks
Sexual	Pornography (harmful or illegal), sexualization of culture, oppressive body image norms	Sexual harassment, sexual grooming, sextortion, the generation and sharing of child sexual abuse material	Sexual harassment, non-consensual sexual messaging, adverse sexual pressures	Trafficking for purposes of sexual exploitation, streaming (paid-for), child sexual abuse
Values	Mis/disinformation, age-inappropriate marketing or user-generated content	Ideological persuasion or manipulation, radicalisation and extremist recruitment	Potentially harmful user communities e.g. self-harm, anti-vaccine, adverse peer pressure	Gambling, filter bubbles, micro-targeting, dark patterns shaping persuasion or purchase
Cross-cutting	**Privacy violations (interpersonal, institutional, commercial)** Physical and mental health risks (e.g. sedentary lifestyle, excessive screen use, isolation, anxiety) Inequalities and discrimination (in/exclusion, exploiting vulnerability. algorithmic bias/predictive analytics)			

It is important to recognise that exposure to the online risks in the CO:RE classification does not necessarily lead to harm. The likelihood, severity and nature of harm depends on the combined influence of a variety of risk and resilience factors specific to each child (Livingstone, 2013). Risk exposure also relates to the design of platforms (e.g., content/friend recommender algorithms) and the quality and efficacy of safety tools implemented (e.g., content moderation, privacy settings), as well as wider regulatory and social contexts (5Rights Foundation, 2021b, 2023; Livingstone & Stoilova, 2021).

How, then, does online harm arise? Drawing on qualitative research with children and adults, recent research has developed a 'risky pathways' model of three distinct pathways by which exposure to online risks (termed 'hazards' in the model) can lead to harm (Gill et al, 2022; Ofcom, 2022a):

- **Isolated exposure** to hazards that have immediate but largely transient emotional impacts, leading to minimal harm. This includes limited or one-off exposure to sexual, violent and other types of harmful content. It could also relate to initial contacts or sexual requests made by known or unknown people (adults or peers).
- **Cumulative passive exposure** to hazards over time can lead to more significant harm. This includes more prolonged exposure to harmful content and/or contact from adults and peers (e.g., pro-eating disorders, misinformation). It could also reflect more prolonged experience of sexual contact and content sent by adults/and or peers.
- **Cumulative active engagement** reflects longer-term engagement with content and contact that self-reinforces attitudes or behaviours, leading to significant and severe harm. This includes active membership of pro-anorexia and extremist communities online and could also relate to longer-term exposure to sexual contact and/or contact from adults or peers.

A variety of potential outcomes related to these pathways are also identified by the model. These include transient emotions (e.g., anger, disgust, confusion), longer-term emotional impacts or behavioural changes (e.g., physical aggression, short-term food restriction), as well as severe psychological and physical harm (e.g., depression, anxiety, self-harm, eating disorders).

The 'risky pathways' model recognises that not all exposure to harmful contact or content has these outcomes and highlights the importance of understanding the role of risk or vulnerability factors specific to the individual (e.g., depression, body image, family environment) in determining whether and how this occurs. Some children may be exposed to isolated or cumulative risks that have little negative impact due to other factors in their life (e.g., resilience, family support).

This model was developed in relation to children's exposure to harmful content, contact, and conduct, rather than the specific sexual risks that form the focus of the present review. One limitation of the model, particularly in this context, is the use of the term 'hazards' to refer to online risks. The adoption of a health and safety framework for understanding children's exposure to online risk and harm is useful, but the associated environmental language has the potential to minimise the nature and impact of children's experience of online harm, particularly sexual abuse. Similarly, categorising children's experiences of sexual and other harms as isolated or cumulative does not recognise the significant and negative impacts that isolated exposure can have.

Despite this limitation, the 'risky pathways' model provides a useful framework for understanding the processes by which children can be exposed to risk of harm online. Although it was not developed with specific reference to online sexual risks, it does illuminate the different pathways leading to harm.

1.3 Research questions

The research questions addressed by the review were:

Primary research questions

- What are the new developments in the risk landscape that have arisen over the past five years, and are there any emerging or longer-term trends?
- What does the current evidence suggest about children's exposure to online sexual risks and the associated harms?
- What does the current evidence suggest about children's exposure to other types of online risk and the associated harms?
- What does the current evidence suggest about the ways in which technological design features moderate the likelihood of children encountering or being harmed by online risks?
- What are the gaps in current understanding of children's safety online and the harms affecting them?

Secondary research questions (addressed in relation to each category of online sexual risk)

- What does the current evidence base suggest about:
  - The nature of the risk?
  - The prevalence of the risk?
  - The online platforms or services on which the risk is usually encountered?
  - The ways in which exposure to the risk or susceptibility to harm varies by children's circumstances, age or other demographic characteristics?
  - Children's responses to the risk?
  - The outcomes for children who encounter the risk?
  - Patterns or trends in harmful outcomes to children over the past five years?

1.4 Methodology

Review of the evidence on online sexual risks

A modified Rapid Evidence Assessment (REA) methodology (Davies, 2003) was used to examine the online sexual risks to which children are exposed. This approach has been applied in other published reviews of the literature related to online risk and harm, for examples those by Davidson et al (2019) or DeMarco et al (2018). A systematic and structured search strategy (e.g., inclusion/exclusion criteria, number of databases searched, number of search terms and strings used) was used to identify relevant academic literature published since 2017 on online sexual risk victimisation (within the categories defined in Chapter 2). A search was also undertaken of grey literature published since 2017 in the UK and other countries, to identify relevant reports and publications by other stakeholders (e.g., Ofcom, charities, government departments and data sources). The search for grey literature included the use of online search engines, searches on the websites of the relevant stakeholders, reference to publications already known to the research team, and a call for evidence (circulated to UKCCIS members, other experts, and via relevant mailing lists) in February 2023.

Review of the evidence on other categories of online risk and technical tools

A light touch review was conducted for other categories of online risk classified as 'primary priority' and 'priority' content in the Online Safety Bill (DSIT, 2023). This consisted of a 'review of reviews' (REAs, literature reviews), supplemented by limited academic and other searches for newer publications in these areas (published between 2021 and 2023), together with an examination of research by Ofcom and other relevant organisations. A similar light touch approach was used to review evidence regarding the technical tools that reduce or increase exposure to online risk and harm.

Scope of the review

This evidence review focused on:

- Evidence published since 2017. While the review occasionally makes reference to evidence published prior to that date, it does not attempt to summarise the cumulative evidence on each topic.
- Primarily UK studies, including evidence from England, Wales, Scotland and Northern Ireland; however, it also includes studies that use samples from other countries if those studies add to understanding of victimisation within the UK.
- Children of all ages; however, the studies that were identified usually focused on children aged 11–18, with only some studies including younger children.
- Evidence of online victimisation. The review only addresses perpetrators where they are relevant to understanding how victimisation occurs and does not provide consideration of the wider literature on offending, prevention and treatment.

Assessment and presentation of the evidence

The quality and contribution of identified publications to the UK evidence base was assessed, but formal quality assessment tools were not used.

The results of the review are presented using a narrative summary (Dixon-Woods et al, 2005). This was most appropriate to the aims and timescale of the project, as well as the variety of methodologies that characterise the literature in this area.

More detail about the methodology can be found in Appendix 1. Short summaries of the academic studies included in the review are included in Appendix 4.

1.5 Structure of the report

Chapter 2 provides an overview and definitions of the different categories of online sexual risk and harm, which are the central focus of the review and a key focus of the NSPCC's policy work. It highlights some of the challenges associated with these different categories, and the intersections between them.

Chapters 3, 4 and 5 take each category of online sexual risk in turn, and present evidence on the scale of children's exposure to that risk, the platforms where the risk typically occurs, as well as children's responses and outcomes when they encounter the risk. Each chapter ends by listing gaps in knowledge about that specific form of online sexual risk.

Chapter 6 helps place the online sexual risks described in the previous chapters in perspective and in their wider context, by presenting a brief overview of various other risks to which children are exposed online. The chapter does not attempt to cover every type of online risk but instead focuses on those that are termed 'primary priority' and 'priority' content in the Online Safety Act (e.g., cyberbullying, self-harm and suicide content).

Chapter 7 shifts away from the risks themselves to look at the technology that can facilitate and enhance children's exposure to risk or, alternatively, supress risks or help children steer past them. By looking at the technology involved, the chapter highlights that risk is not an inevitable outcome of being online: it is also influenced by the design choices of platforms and the safety tools implemented.

Chapter 8 brings together the findings from the previous chapters to directly address the five research questions examined by the review (outlined in Section 1.3 above).

Finally, Chapter 9 builds on the challenges and knowledge gaps identified in the previous chapters to offer recommendations about future research needed to provide a fuller and more accurate understanding of the online risk landscape for children in the UK.

1.6 A note on terminology

This review uses the term 'children' throughout to refer to anyone under the age of 18. This is consistent with the language of the Online Safety Act and the legal definition of a child in the UK.[2] The age ranges of children covered by studies is included when available to provide additional context.

The term 'platform(s)' is used to refer to services that enable children in the UK to access user-generated content and facilitate user-to-user interaction (e.g., social media, messaging apps, search services, video-sharing services).

The term 'sexual image' is used throughout the report to cover a range of formats, including photographic stills, videos or livestreamed sexual imagery.

A glossary and list of abbreviations can be found at the start of this report.

[2] It should be noted that the legal definition/age of consent in the UK is 16 and, consequently, the Sexual Offences Act and other related acts do not provide protection for 16–17-year-olds, unless under specific circumstances (e.g., abuse of position of trust).

Chapter 2
Online Sexual Risks

This chapter provides an overview of the different categories of sexual risk and harm to which children are exposed online. This includes 'online sexual harassment' (OSH) and 'intimate image abuse' (IIA) involving peers; 'technology-assisted child sexual abuse' involving adults (TA-CSA); as well as the production and dissemination of 'child sexual abuse material' (CSAM), which can result from peer or adult sexual interactions. Working definitions for each of these terms are provided in subsequent sections of this chapter. These may differ with other definitions found in the literature or used in policy contexts but are the basis for the category-specific evidence review provided in subsequent chapters.

2.1 Distinguishing between abusive and non-abusive online sexual experiences

This report primarily examines online sexual interactions that pose risk or cause harm to children. Before starting, however, it is important to recognise that children's online sexual behaviour and interactions with peers can sometimes be consensual and developmentally appropriate (Döring, 2014). Older adolescents may use social media and other platforms to meet developmental needs associated with the exploration of sexual identity, sexual and romantic relationships, as well as to gain social approval and popularity (Bianchi et al, 2017; Currin & Hubach, 2019). In these situations, the behaviours involved would not be viewed as inherently problematic or constituting abuse, as long as no adult is involved, the children are of an appropriate age, they are close in age to each other, and there is no pressure or coercion involved (Burén & Lunde, 2018; Ringrose et al, 2022).

Distinguishing between online sexual experiences that are developmentally appropriate and those that are abusive, as well as identifying the specific category of risk to which they belong, is dependent on the age of those involved and whether interactions are consensual.

Table 1 lists the different sexual behaviours or interactions that children can experience online, based on the evidence review and work of the NSPCC. This illustrates how the age of the people involved, as well as the consensual or non-consensual nature of the behaviour or interaction, indicates whether it would be classified as developmentally appropriate or under one of the categories of online sexual risk and harm examined by the report. For example, if a child sends a sexual image of themselves to a peer as part of an age-appropriate and consensual relationship, this would not constitute abuse. When this behaviour is unwanted, or the age gap between the sender and receiver is inappropriate, it would be classified as peer-to-peer Intimate Image Abuse (IIA) and in this report would be classified as a form of Online Sexual Harassment (OSH). In situations where a child sends a sexual image to an adult as the result of grooming or exploitation, the experience would constitute Technology-Assisted Child Sexual Abuse (TA-CSA). Regardless of the circumstances and intentions behind its production, the sexual image would be classed as Child Sexual Abuse Material (CSAM) and its onward distribution would revictimise the child involved. This is also the case where images are consensually produced but then accessed by adults online and redistributed to other adults.

Table 1: Classification of children's online sexual behaviours and experiences across specific categories of risk and harm

Experience	Developmentally appropriate*, age appropriate, and consensual sexual interaction between peers	OSH by peers (if age gap is inappropriate or experience is unwanted)	IIA by peers (if age gap is inappropriate or experience is unwanted)	TA-CSA by adults	CSAM production or distribution
Child produces a sexual image of themselves	×				×
Child shares a sexual image of themselves	×	×	×	×	×
Child has a sexual image of themselves taken	×	×	×	×	×
Child receives sexual talk, comments or content	×	×		×	
Child receives a request for sexual act	×	×		×	
Child receives a request for sexual information	×	×		×	
Child receives a request for a sexual image of themselves	×	×	×	×	
Child receives a sexual image of someone else	×	×	×	×	
Child receives an offer of favours in exchange for a sexual act or image		×	×	×	
Child has a sexual image of themselves produced covertly		×	×	×	×
Child has a sexual image of themselves created artificially (through technology)		×	×	×	×
Child has a sexual image of themselves shared with others		×	×	×	×

Experience	Developmentally appropriate*, age appropriate, and consensual sexual interaction between peers	OSH by peers (if age gap is inappropriate or experience is unwanted)	IIA by peers (if age gap is inappropriate or experience is unwanted)	TA-CSA by adults	CSAM production or distribution
Child receives threats that a sexual image of themselves will be shared with others		✕	✕	✕	
Child receives pressure/coercion/ threats to force them to send a sexual image of themselves		✕	✕	✕	
Child receives pressure/coercion/ threats to force them to share sexual information or perform a sexual act		✕		✕	
Child is bullied/ excluded/ humiliated/ discriminated against regarding their sexual behaviour or sexuality		✕		✕	

Key: 'x' indicates that the experience (in the row heading) could be classed as falling within the category indicated in the column heading, depending on the age of those involved and whether or not it is consensual.
** The authors recognise that it would not be developmentally appropriate for a very young child (for example, a six-year-old) to produce or share a sexual image of themselves.*

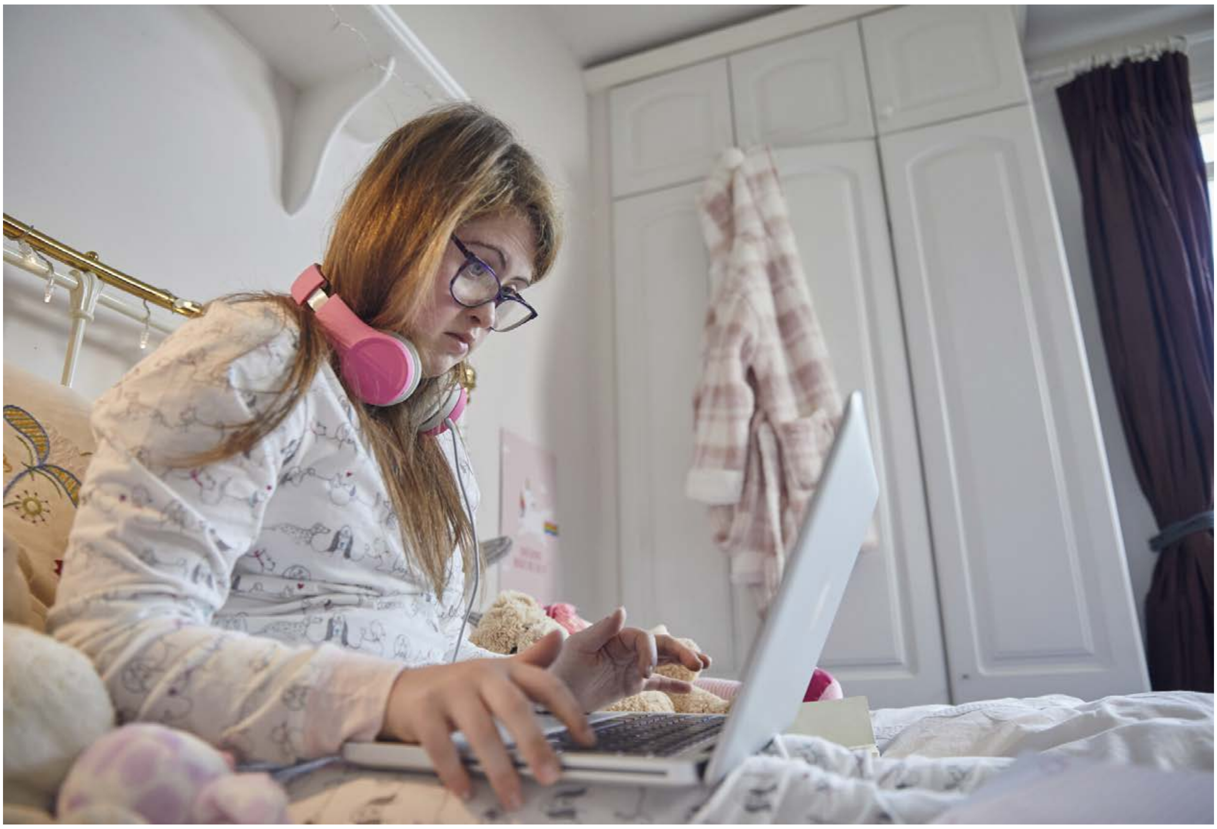
The table is not intended to be definitive, but to illustrate how different sexual behaviours map onto the different categories of sexual risk covered by the review.

These distinctions are worth making. Differentiating between experiences involving an adult and those involving another child is important because of their legal implications (see Section 2.2.5 below) and the power imbalances entailed by the interactions. Separating out wanted from unwanted interactions is also important, as those situations are experienced differently and may lead to different outcomes for children.

A review of the evidence base, however, shows that these distinctions are not always made in research about online sexual risks and harms.

Surveys that explore sexual image exchange, for example, provide prevalence figures that are difficult to interpret. Surveys show that 1–17 per cent of adolescents send sexual images of themselves to others (ONS, 2021; Revealing Reality, 2021a; Katz et al 2020),[3] but it is not clear from these figures whether or not the recipients were willing to receive the images. Surveys also show that 34–46 per cent of children aged 12 to 18 receive requests for sexual images (Burén & Lunde, 2018; Revealing Reality, 2021a), but these figures conflate situations where the requests were made by adults with situations where they were made by (known or unknown) peers. Even in situations where measures are clearly focused on children's interactions with other children, studies do not generally examine the relative ages of those involved. As a result of these types of conflation and ambiguity, it is not possible to unpick what proportion of these experiences occurred within the context of age-appropriate and consensual romantic relationships (McGeeney & Hanson 2017; Burén & Lunde, 2018; Revealing Reality, 2021a; Ringrose et al, 2022) and what proportion represents adult-to-child or peer-to-peer abuse.

More recently, there has been a shift in the way researchers and policy makers conceptualise the production and exchange of sexual images. As detailed in Box 1, there is a move towards acknowledging the varied circumstances, motives and intentions involved in making and sharing images.



[3] The ONS (2021) study collected data from a representative sample of children aged 10-15 in England and Wales; the Revealing Reality (2021a) survey collected data from a non-representative sample of children aged 14-18 in all nations in the UK; and the Katz (2020) survey collected data from a non-representative sample of children aged 13-17 .

Box 1: Recent changes in terminology about sexual image sharing

There are a variety of different definitions and ways of conceptualising the production and sharing of sexual images by children in the research and policy evidence base. Although the term 'sexting' has been used as a general description (Barrense-Dias et al, 2017; Englander & McCoy, 2018), there has been a gradual recognition that this term is problematic as it can refer to sexual images, sexual messages, or both (Madigan et al, 2018).

There is also growing awareness that some children who receive sexual images from peers do not necessarily request or want them (Revealing Reality, 2021a, 2023; Ringrose et al, 2022). Evidence that children themselves distinguish between consensual and unwanted image sharing (Lloyd, 2020) indicates the need for research to more effectively differentiate between the two. This is reflected in the introduction of the terms 'image-based sexual abuse' and 'intimate image abuse' rather than 'sexting' in studies such as by Ringrose et al (2021a), and an accompanying acknowledgement of the importance of consent.

A further development is the increased acknowledgment that there may be varying degrees of agency in the production of sexual images by children. It is now more widely recognised that, while this can occur in developmentally appropriate and fully consensual contexts, they may alternatively do so as the result of coercion by peers or through online grooming by adults (IWF, 2023; NCA, 2021). Victims may appear to be voluntarily engaging in the behaviour, but this could be the result of psychological pressure or threats by adults or peers to ensure compliance. The increase in images produced by children of themselves (see Section 5.1.4) is an important concern for law enforcement, policy makers and researchers (APPG on Social Media, 2021; IWF, 2021, 2022, 2023; Quayle et al, 2018). While a variety of terms are used to describe this type of image production (e.g., 'self-generated', 'self-taken'), these terms problematise developmentally-appropriate production and suggest some level of culpability on the part of the victim when these images are shared in perpetrator networks or identified by detection tools (Tech Coalition, 2021). Discussions about more appropriate terminology have led to the suggestion that 'first-person produced imagery' is a more suitable term, and has been defined as:

> "Sexual visualised depictions of a child that are generated without the full knowledge, consent, and participation (for example, grooming, blackmail and coercion) of the child and without the physical presence of an instigator AND/OR that may have been originally voluntarily produced by the minor child, but then is distributed to or shared with others without the child's full knowledge or consent."
>
> (Tech Coalition, 2021).

This more effectively represents the different circumstances in which sexual images can be produced and does not imply the same level of victim blame. Such changes in terminology are an important way of ensuring that the language used in educational resources, as well as by police, NGOs and policy makers, does not create barriers to children seeking help if they are being exploited in this way.

2.2 Four categories of online sexual victimisation

This report aims to avoid the conflation and ambiguity described above. It distinguishes between peer-to-peer and adult-to-child online sexual interactions and – in the case of peer-to-peer interactions – focuses only on those that are non-consensual. The evidence on online sexual risks was organised into the four categories of sexual online victimisation that appear in Table 1: the peer-to-peer categories of OSH or IIA, the adult-to-child category of TA-CSA, or the category of CSAM (which can result from the actions of either adults or peers).

The sections below offer a working definition for each category. It is worth noting that the definitions focus on victimisation; perpetration is only mentioned where it is relevant to understanding how victimisation occurs and there is no attempt to consider the wider literature on offending, prevention and treatment.

2.2.1 Online sexual harassment (OSH)

In this report, online sexual harassment is conceptualised as unwanted or non-consensual interaction between peers. This can involve children making sexual and/or gender degrading comments towards others online, as well as blackmail, control, coercion and humiliation (Guerra et al, 2021; Henry et al, 2019).

A more specific definition related to the involvement of peers is provided by Project deSHAME (2017), which characterised OSH as unwanted sexual conduct by peers involving a variety of online behaviours and content types across private or public platforms.

The deSHAME Project (2017) developed a taxonomy of peer-related online sexual harassment across four categories:

- **Non-consensual sharing of intimate images or videos:** This covers behaviour that is sometimes referred to as 'sexting' in the literature. It involves sexual images or videos of a child being produced, taken and/or shared without their consent. This has more recently been termed 'intimate image abuse' (see Section 2.2.2 below).
- **Exploitation, coercion and threats:** This relates to children being targeted by sexual threats, being coerced to engage in sexual behaviour online (including image production) or being blackmailed with sexual content.
- **Sexualised bullying:** This involves children being targeted by individual peers or groups using sexual content that humiliates, upsets or discriminates against them (e.g., sexual comments, sexual name-calling).
- **Unwanted sexualisation:** This refers to the receipt of unwanted sexual requests, comments or content online.

While deSHAME's taxonomy has some limitations (including overlapping categories), it demonstrates the variety of experiences that can fall under OSH. Crucial to the definition of OSH is that these peer-to-peer interactions are intended to make the victim feel upset, threatened and coerced, humiliated, sexualised or discriminated against. It largely occurs within existing peer networks and involves an active, engaged online audience.

2.2.2 Intimate image abuse (IIA)

In this report, the non-consensual taking or sharing of sexual images between peers is classed as intimate image abuse. IIA is, therefore, a sub-category of OSH that focuses on sexual imagery. For the purposes of this report, only peer-to-peer interactions will be treated as IIA; situations where similar behaviours take place between adults are not taken into consideration as they are beyond the scope of the review.

The different aspects of IIA examined in the literature involve:

- **Receiving unwanted requests for sexual images:** These requests constitute OSH when they are made by peers and are unwanted, but otherwise they would be viewed as developmentally appropriate.
- **Receiving unwanted sexual images:** Although this can be consensual and developmentally appropriate when the images are sent by peers and are welcomed by the recipient, receiving these images unwillingly is classified as IIA. This is sometimes referred to as being 'cyberflashed'.
- **Receiving pressure (or being coerced) to produce or send sexual images:** This involves children being threatened or coerced into producing sexual images or engaging in sexual activity (O'Malley & Holt, 2022; Wolak et al, 2018).[4] It can also include children sending unwanted sexual images to peers to pressure them to reciprocate, also known as 'transactional' image-sharing (Ringrose et al, 2021a, 2021b).
- **Non-consensual sharing of sexual images:** This refers to the sharing of sexual images without the consent of the person depicted, either by the person they were originally sent to or by others who subsequently receive them.[5] These can be shared on social media, group chats or posted on 'bait out' pages (Lloyd, 2020; Ringrose et al, 2022).[6]

It is important to note that when any of the experiences above involve an adult perpetrator and child victim, they would be classified as sexual solicitation/grooming and constitute TA-CSA rather than OSH or IIA.

In this report, only experiences that are explicitly or implicitly referred to as unwanted, non-consensual, or created under pressure or threat are considered to be IIA. Research studies that are ambiguous in relation to consent were not included in the review.[7] A degree of uncertainty remains in some of the included studies as to the age of the perpetrator, but, in the main, the interactions were not believed to involve an adult.

The available evidence suggests that children perceive IIA, or behaviours that fall under the umbrella of OSH, to be motivated by a desire to gossip, joke or humiliate the person in the image (Project deSHAME, 2017; van Ouystel et al, 2021). More recent qualitative research with older adolescents and young adults in the UK found that perpetrators justified sharing

[4] The term 'sextortion' is also used in research in this area but can also relate to children and/or adults being financially extorted to prevent sharing of sexual images online or with friends and family.

[5] Although this is also sometimes referred to as 'revenge porn', the term is not appropriate where children are involved as the behaviour constitutes abuse.

[6] 'Bait out' pages are social media pages or groups where sexual images are posted non-consensually to humiliate, embarrass and shame the people depicted in them.

[7] Sexual image sharing involving peers with an inappropriately large age gap would also fall under the category of IIA, regardless of whether the image exchange was considered consensual. However, it is rare for research to establish the age of all parties involved or to make a judgement on the appropriateness of the age gap.

the sexual images of others without their consent by claiming that the victim would not mind, that their actions would not cause harm, or that the person in the image would never find out (Revealing Reality, 2023). Retaliation and revenge are also highlighted by researchers as motivations and are a potential explanation for the identified overlap between victimisation and perpetration in some studies (Boer et al, 2021).

Only a small minority of adolescents admit to sharing other children's sexual images without their consent; estimates range from 5–8 per cent of children (age range 12–17 years) in studies in the grey literature in the UK and academic studies from other countries using non-representative samples (Gámez-Guadix et al, 2022; Project deSHAME, 2017). A systematic review of 39 studies found a higher prevalence of 12 per cent across the studies examined (Madigan et al, 2018). This suggests that this behaviour is relatively uncommon, although these studies generally rely on self-report measures and prevalence may be under-reported due to social desirability.

2.2.3 Technology-assisted child sexual abuse (TA-CSA)

The term technology-assisted child sexual abuse (TA-CSA) is used throughout the report to refer to situations where children are sexually exploited by adults online; or where technology is used to facilitate offline abuse by adults (Hamilton-Giachritsis et al, 2020).

There are a variety of different academic and legal or policy definitions of the behaviours and experiences involved in TA-CSA, with terms like online grooming, sexual solicitation and online sexual exploitation being commonly used. In the literature from the last six years, these terms are often used to refer to online sexual interactions with adult perpetrators (Joleby et al, 2021), but can sometimes also refer to peer-related exploitation (Finkelhor et al, 2022; Home Office, 2020). These differences in definitions and terminology create challenges when comparing studies.

For the purposes of this report, TA-CSA will be used to refer only to abuse perpetrated by adults. Unwanted sexual conduct by peers is categorised here as OSH instead (see Section 2.2.1 above).

The different aspects of TA-CSA covered by research include:

- **Sexual solicitation:** This refers to adult requests to children for sexual information, talk and images, as well as to engage in online or offline sexual activities. It can also refer to situations where they receive sexual images, content, comments or threats from adults (de Santisteban & Gámez-Guadix, 2018; Finkelhor et al, 2022).
- **Sexual interaction:** This describes situations where children engage in online sexual talk, sexual activities online (or offline with someone they have met online), or exchange sexual images with adults (e.g., de Santisteban & Gámez-Guadix, 2018; Ortega-Barón et al, 2022). These experiences are also part of the grooming process (see Box 2 for more detail).

Box 2: Recent changes in understanding of online grooming strategies

There is a relatively small body of research produced since 2017 that has examined online grooming processes. This has mostly been based on analysis of chatlogs of grooming cases obtained by the police during investigations or interviews with victims and perpetrators, as illustrated in studies by Kloess et al (2017, 2019).

Although grooming was initially conceptualised as a linear and stage-based process (O'Connell, 2003), more recent evidence suggests that this may not adequately reflect the way in which perpetrators manage contact with victims (Barber & Bettez, 2021; Kloess et al, 2019). As a result, it has been argued that grooming can be conceptualised as involving the use of a variety of different manipulative strategies over varying periods of time (Kloess et al, 2019; Ringenberg et al, 2022; Webster et al, 2012),[8] with perpetrators[9] expediently adopting different strategies to realise their goals of contact and/or non-contact sexual offending (DeMarco et al, 2018).

The grooming strategies identified in the literature include:

Friendship and relationship strategies: Perpetrators may use specific friendship strategies to develop relationships with children if they respond to their initial approaches. These aim to start conversation and develop a friendship, which encourages victims to discuss their lives and develop trust. Perpetrators may present themselves as a mentor and source of support, as well as use compliments, flattery and/or gifts (Joleby et al, 2021; Kloess, et al, 2017, 2019). This can also include the subtle introduction of sexual talk to test the reaction of the victim (Kloess et al, 2019; Webster et al, 2012).

Sexual strategies: The use of sexual strategies by perpetrators can happen very early in online grooming or can be delayed depending on their specific offence goals (Webster et al, 2012). Sexual strategies include sexual conversations, sexual activity via webcams, and image exchange (Chiang & Grant, 2019; de Santisteban & Gámez-Guadix, 2018; Joleby et al, 2021).

Use of threats and coercion: Sexual requests may be non-threatening, but perpetrators can become abusive or use coercion to ensure compliance with their requests if the victim is reluctant or resistant (Chiang & Grant, 2019; Hamilton-Giachritsis et al, 2020; Seymour-Smith & Kloess, 2021). This is consistent with other evidence that some perpetrators use direct, aggressive and sexual approaches from first contact (Joleby et al, 2021).

[8] Although the review period was from 2017 onwards, older studies such as that by Webster et al (2012) have been included where there was no relevant recent evidence available. It is important to bear in mind differences in the technological contexts and online behaviours of children when earlier studies were conducted when considering their implications for children's contemporary online experiences.

[9] Although the review period was from 2017 onwards, older studies such as that by Webster et al (2012) have been included where there was no relevant recent evidence available. It is important to bear in mind differences in the technological contexts and online behaviours of children when earlier studies were conducted when considering their implications for children's contemporary online experiences.

The evidence suggests that perpetrators use a variety of different grooming strategies to manipulate children into engaging in sexually abusive activities online and offline. While some take a longer-term approach to grooming, others are more directly sexual and/or threatening from initial contact (Kloess et al, 2019). The use of non-threatening grooming strategies is similar to the ways in which children develop relationships with peers, making it harder for victims to identify the use of manipulation and recognise contact as abusive (Chiu & Quayle, 2022; Whittle et al, 2013).

This is consistent with evidence that some children perceive themselves to have agency in the relationship they have with perpetrators, which they often view as consensual and reciprocal despite the manipulation involved (Chiu & Quayle, 2022). This may reflect the ability of the relationship to fulfil unmet needs in the victim's life, which result from individual vulnerabilities e.g., psychological, family and peer problems (Chiu & Quayle, 2022).

TA-CSA can occur in a variety of ways. Perpetrators may be known to the child, either as part of their family context (intra-familial), or from other contexts (extra-familial) (Centre of Expertise on Child Sexual Abuse, 2020). Perpetrators may alternatively be unknown to the child. Technological developments create opportunities for offending by adults who are initially unknown to the victim: for example, webcams and video functionality enable livestreaming of child sexual abuse for commercial purposes, with children being manipulated into filming themselves and receiving direction from the perpetrator or a third party facilitating the abuse (NCA, 2021; Napier et al, 2021a, 2021b). Children can also be encouraged to livestream themselves engaging in sexual activity, perceiving this to be for peers, or may do so as the result of being groomed and manipulated by adults (NCA, 2021). Livestreaming has emerged as a significant means of the production of first-generation CSAM in recent years, with the victim often being unaware that content has been captured and shared (IWF, 2022; NCA, 2021).

Virtual reality environments and immersive technologies also have the potential to facilitate TA-CSA by providing opportunities for adult perpetrators to engage in abusive sexual activities with real and virtual children (Pettifer et al, 2022; McIntosh & Allen, 2022). This includes the ability to groom and engage in abusive interactions with real children individually or in group-based situations where others can view and direct abusive interactions (Pettifer et al, 2022). The potential to sexually interact with virtual children carries the additional risk of reinforcing offence-related cognitions, increasing disinhibition and encouraging sexual offences against real children in virtual or offline contexts (Allen & McIntosh, 2023; Pettifer et al, 2022).

2.2.4 Child sexual abuse material (CSAM)

The term Child Sexual Abuse Material is used in this report to refer to any sexual depictions of children in still or moving images.

There are a number of different offending behaviours related to CSAM covered in the literature:

- **CSAM production:** This material can come into existence through photography, videoing, or livestreaming of real children, or synthetically though various technical means (e.g., by digitally editing and mapping the facial or bodily features of a victim onto the face and body of another).
- **CSAM possession:** This refers to perpetrators viewing and/or downloading CSAM, with many creating personal collections of preferred content.
- **CSAM distribution:** Child sexual abuse material is considered to be in circulation if it is stored electronically (e.g., on the open web, dark web,[10] peer-to-peer networks, cloud-sharing apps, on devices) and is accessible to others through searches, weblinks, posts, direct messages or other routes.

There are various ways in which CSAM is produced. The depiction of real children can result from IIA, from TA-CSA, or through non-abusive, developmentally appropriate and age-appropriate sexual interactions between peers. Alternatively, CSAM can be created contextually through the placing of everyday photos of real children in sexualised contexts (NCA, 2021; Pedersen et al, 2020). Recent years have also seen an increase in production of Non-Photographic Abuse Imagery (NPAI) that feature animated depictions of children engaging in sexual activity or being sexually abused (Merdian et al, 2022). Even more recent technological developments in generative AI and computer-generated imagery have now provided the ability to artificially create 'deepfake' or synthetic CSAM that appears to show real children (Eelmaa, 2022; WeProtect Global Alliance, 2021; Thiel et al, 2023).

While CSAM distribution can take a number of forms, it is often difficult to detect and quantify. CSAM perpetrators usually take steps to hide their identities and avoid detection using methods like 'digital pathways'[11] and 'top-level domain hopping'[12] to share material with other internet users (Europol, 2020; IWF, 2021; WeProtect Global Alliance, 2018, 2021).

[10] The dark web is a part of the internet that hosts websites (or 'hidden services') that are not indexed by search engines. Hidden services need specific software, configurations, or authorisation to access. The identities and locations of users who visit the dark web stay anonymous and cannot be tracked due to encryption.

[11] 'Digital pathways' or 'digital breadcrumbs' are methods perpetrators use to direct others to CSAM. This can include edited images or videos that evade content moderation processes because they do not meet the legal threshold for CSAM (NSPCC, 2022b).

[12] 'Top-level domain hopping' occurs when websites hosting CSAM (e.g., 'CSAMhost.tk') change their top-level domain ('.tk') while keeping their second-level domain name ('CSAMhost') to remain online after they have been identified for removal (IWF, 2022). This creates a new website (e.g., 'CSAMhost.ru') with the same content that can still be found by users, and sites may do this multiple times ('hopping') in response to subsequent removals from different top-level domains (IWF, 2022).

2.2.5 Legal frameworks

It is important to recognise that it is an offence for children to produce, send, or share sexual images under the Protection of Children Act 1978: Section 1(1) and the Criminal Justice Act 1988: Section 160. This applies regardless of the circumstances of their production. When the perpetrators of OSH and IIA are children, their behaviour would normally be treated as a safeguarding issue unless there is clear evidence of the use of threats or coercion, the involvement of sexual or physical violence, or the involvement of adults (UKCIS, 2020). If children exchange sexual images and the police are confident no harm has occurred, they can record these offences using the Outcome 21 code: this indicates that formal criminal action would not be in the public interest and avoids criminalising children (College of Policing, 2016; Quayle, 2022). Despite this, research suggests the code is applied inconsistently across police forces and can have a negative influence on the criminal record and career prospects of children (Bond & Phippen, 2019).

The involvement of adult perpetrators (which in this report is categorised as TA-CSA) and the production of CSAM are both offences under the Sexual Offences Act (2003) and other relevant legislation, regardless of whether the child perceived the behaviours involved to be consensual. This highlights the complexity of understanding children's online sexual behaviour and experiences, and the related potential for risk and harm, within developmental, safeguarding and legal contexts.

Appendix 2 contains more information about the relevant legal frameworks in the UK for each of the categories of online sexual risk covered in this chapter.

2.3 Methodological issues in researching online sexual risks

The next three chapters of this report take a closer look at the four categories of online sexual victimisation, synthesizing the evidence available since 2017. The evidence base on each of the four topics shares several limitations that are worth highlighting at this stage:

- There has been a general lack of research in the UK in the last six years examining the four different categories of sexual risk. To some extent, this reflects the ethical and safeguarding challenges associated with asking children – especially those below secondary school age – about sexual victimisation experiences, and in accessing children for this purpose through gatekeepers like schools. However, the lack of funding for such research represents the main barrier to the development of a clearer evidence base about these forms of online victimisation in the UK.
- Studies examining the prevalence of different forms of online sexual risk and harm do not use consistent definitions for the phenomena they are measuring. Ways of measuring sexual behaviours and interactions also differ, with measures typically not distinguishing between peer and adult perpetrators, or consensual and non-consensual experiences. Studies additionally use a range of different methodologies and sampling strategies. Altogether, these inconsistencies may be one potential explanation for differences in prevalence rates between studies.

- Cross-sectional designs with samples of children[13] are commonly used to explore how prevalent particular types of victimisation are at particular points in time. However, these samples are not always representative of the broader population of children. Moreover, these designs are limited in their ability to establish causality: whether other factors lead to the victimisation, or whether victimisation itself leads to those other factors (Ortega-Barón et al, 2022).
- Retrospective studies with samples of young adults, which examine whether and how often respondents experienced a particular phenomenon throughout their childhood, are less common. This is regrettable, as participants may be better able to recognise exploitative interactions retrospectively, improving the validity of the related prevalence figures (Davidson et al, 2017; Quayle, 2022). Where they exist, these studies suffer from an inherent limitation: the pace of technological change and developments in the ways children interact online mean that the results of studies using this methodology quickly become outdated and may not directly inform understanding of contemporary children's experiences online.
- Longitudinal designs explore experiences of online sexual risk and harm of a cohort of children at two or more time points. These designs enable the examination of how victimisation relates to other risk or vulnerability factors and outcome variables. Although such designs are time and resource intensive, they can make an important contribution to the evidence base by allowing stronger claims about causality to be made (Ortega-Barón et al, 2022). Disappointingly, there have been no longitudinal studies addressing the four categories of online sexual risk and harm in the UK, highlighting an important area for future research.
- Research using any of the three different designs described above comes with further limitations. They generally require children to self-report their online behaviour and experiences. The prevalence figures and relationships between measured variables derived from the self-reported data will inevitably be influenced by children's understanding of the questions asked. Social desirability and concerns about privacy of the data may also lead to under-reporting of problematic behaviours and experiences, such as perpetration of OSH and IIA.

[13] Cross-sectional designs provide an assessment of the prevalence of measured behaviours and experiences (e.g., sexual solicitation) within a specific time period (e.g., in the six months before data collection).

Chapter 3
Online Sexual Harassment and Intimate Image Abuse

This chapter provides an overview of the available evidence on peer-related sexual harassment in the online environment (OSH), as well as intimate image abuse (IIA). It examines the scale of the different behaviours involved, responses, outcomes and impacts, as well as its gendered dimensions. This is based on academic and grey literature published since 2017 and draws on research using UK and non-UK samples.

As described in the previous chapter, OSH is an umbrella term for a variety of different harmful behaviours perpetrated by children against other children, with IIA between peers a subtype of this category. Situations that involve adults as perpetrators constitute TA-CSA and CSAM offending, and the related evidence is reviewed in later chapters.

Chapter summary

- **Scale**: Although the UK has produced no new evidence since 2017, prevalence figures from elsewhere suggest that 8–26 per cent of children experience OSH that is not specifically related to sexual images. Fewer children experience IIA, whether that is receiving unwanted sexual images from peers (5–11 per cent); having their own sexual images shared with others without their consent (3–8 per cent); or experiencing pressure or coercion to send sexual images (3–7 per cent).
- **Scale**: It is not possible to determine whether the scale of OSH and IIA victimisation has changed in recent years, as these types of experiences have not been measured in consistent ways over time.
- **Platforms**: There is limited data available about the platforms involved in OSH and IIA, but studies suggest that Snapchat and Instagram Messenger are popular apps for sharing sexual images. Other platforms frequently used by children are also likely to be used to harass peers through commenting, liking and forwarding.
- **Children most likely to be harmed**: Girls are more likely than boys to report victimisation from most types of OSH and IIA, highlighting the gendered contexts in which these experiences occur. Victimisation is also more common among older adolescents compared with younger children.
- **Responses**: The most common response to OSH and IIA is to delete messages and images or use technical tools to block unwanted contact. Reporting to platforms and offline help-seeking is less common, owing partly to concerns about the reactions and judgements that would result from telling adults (parents/carers, teacher). Another common response is to do nothing; this appears to be linked to feelings of self-blame, concerns over judgements, or the perception that victimisation is a normal and inevitable consequence of being online and that nothing can be done once sexual images are leaked.
- **Outcomes and impacts:** The impacts of OSH and IIA tend to be more severe for girls than boys. They include emotional responses (e.g., initial upset followed by desensitisation); psychological impacts (e.g., humiliation, helplessness, fear, shame or self-blame), as well as depression, anxiety and suicidal thoughts; and social impacts, including reputational damage, peer harassment and exclusion, and victim blaming. Children may also experience revictimisation through the onward distribution of their images.

3.1 OSH and IIA: Scale

Overall, the evidence suggests that a minority of children experience OSH and IIA. The variation in prevalence figures outlined below partially reflects differences in measurement and sampling between studies, but it is also likely that children under-report such experiences depending on how the data is collected and concerns over the privacy of their responses.

It should be noted that there is some ambiguity regarding the age of the perpetrator in most of the studies from which the data below is derived. What is not ambiguous is that the behaviour was unwanted and unwelcome. Details of the prevalence data and design of the specific studies from which the figures are taken are provided in Appendix 3.

Prevalence of OSH (general)

There are few studies examining the broader set of behaviours that constitute OSH in the literature. The review did not identify any UK academic studies published since the 2017 review. However, there was evidence in the UK grey literature and academic studies from other countries that have relevance. These found that 8–26 per cent of children in representative and non-representative samples reported some form of online harassment that was sexual in nature (age range 12–17 years) (Barbovschi et al, 2021; Guerra et al, 2021; Ofcom, 2022b; Project deSHAME, 2017).

Prevalence of receiving unwanted requests for sexual images

As indicated in Section 2.1, evidence from non-representative samples indicates that over a third of children receive requests for sexual images, but there is no reliable source of data that distinguishes whether the requests were wanted or unwanted by recipients. Consequently, no prevalence data is available for this type of IIA.

Prevalence of receiving unwanted sexual images

A low proportion of children report this experience online based on evidence from the grey literature in the UK (5–11 per cent in representative and non-representative samples, age range 12–17 years) (Ofcom, 2022b; ONS, 2021; Revealing Reality, 2021a).

Prevalence of receiving pressure (or being coerced) to send sexual images

A small proportion of children have had this experience, based on the grey literature in the UK and academic studies from other countries (3–7 per cent in representative and non-representative samples, age range 12–17 years) (Finkelhor et al, 2022; Gámez-Guadix et al, 2022; Patchin & Hinduja, 2019; Project deSHAME, 2017).

Prevalence of non-consensual sharing of sexual images

There has been a general lack of academic research on this topic in the UK during the review period, but there is evidence from the grey literature and academic studies from other countries that suggest that 3–8 per cent of children in representative and non-representative samples report this experience (age range 12–18 years) (Finkelhor et al, 2022; Gámez-Guadix et al, 2022; Madigan et al, 2018; Ofcom, 2022a).

3.2 OSH and IIA: Platforms

There is limited evidence related to the platforms involved in OSH and IIA, but research suggests that Snapchat and Instagram Messenger are popular apps for sharing sexual images (Revealing Reality, 2021a; Ringrose et al, 2021a). Qualitative research conducted in the UK (Revealing Reality 2021a, 2023) found that Snapchat was commonly used to host 'bait out' pages[14] and the platform most frequently used to send sexual images because images are automatically deleted after viewing in one-to-one conversations (or after 24 hours or in chat groups if viewed by all members). These features are perceived to provide safety for the person sending images, but workarounds (e.g., photographing the image using another device) can also lead to images being 'leaked' (Mandau, 2021). The 'ephemeral' nature of messages can also provide perpetrators with an effective way of sexually harassing peers and sending unwanted sexual images, because of the privacy and secrecy they provide. Although messaging apps are most frequently used for non-consensual sharing of sexual images, wider dissemination in peer networks is likely to occur across a variety of platforms.

3.3 OSH and IIA: Children most likely to be exposed to risk or experience harm

3.3.1 Demographic factors

- **Age:** There is evidence from non-UK academic studies that older adolescents are more likely than younger children to experience OSH and IIA victimisation, as well as to be perpetrators of these behaviours (Gámez-Guadix et al, 2022; Guerra et al, 2022). However, other studies found no variations by age (Madigan et al, 2018; Patchin & Hinduja, 2020). There is a lack of evidence related to interactions between age and gender in this area.
- **Gender:** There is evidence from the UK and studies using samples of Spanish and American children that – with the exception of sextortion (see below) – most forms of IIA are more frequently experienced by girls. Revealing Reality (2021a) found more girls than boys reported receiving unwanted sexual images from someone known to them (13 per cent of girls vs 3 per cent of boys); while the ONS (2021) survey of a representative sample of 10–15-year-olds in the UK found that more girls than boys received an unwanted sexual image in the last 12 months (16 per cent of girls vs 6 per cent of boys). A similar pattern of results for victimisation was found in a non-representative sample of Spanish 12–17-year-olds, although there were no gender differences in perpetration of non-consensual sharing or sextortion (Gámez-Guadix et al, 2022).

> "It was disgusting, I was like 'Eurgh'. I've never asked for them ['dick pics']. If you're having a conversation and it's getting a bit flirty they'll send them. I would just change the subject and not reply as it's not the type of thing you want to be seeing."
>
> Girl, aged 17 (Revealing Reality, 2021a)

[14] 'Bait out' pages are social media pages or groups where sexual images are posted non-consensually in order to humiliate, embarrass and shame the people depicted in them.

Not only are girls more likely to experience OSH and IIA, but evidence suggests these experiences have more negative outcomes and consequences for them compared with boys (e.g., depression, victim blame) (Burén & Lunde, 2018; Gámez Guadix et al, 2022; Ringrose et al, 2021a, 2021b, 2022). This reflects the gendered contexts, expectations and judgements that shape the behaviours involved in these forms of victimisation. Research suggests that boys pressure girls to send sexual images, and those who do so are negatively judged by both genders (Burén & Lunde, 2018). Boys do not experience similar pressures and judgements in relation to sharing sexual images and IIA (Ringrose et al, 2022). Although there is evidence of peer pressure to request and obtain sexual images from girls, boys can gain status for being sent and sharing them with others (Hunehäll Berndtsson & Odenbring, 2021; Livingstone et al, 2017). Female victims of IIA report experiencing blame and shame, whereas there is evidence that the impacts of victimisation are less severe for boys who may also experience an increase in social status and popularity (Gewirtz-Meydan et al, 2018a; Ringrose et al, 2022).

This is further demonstrated by Revealing Reality (2021a), which found that although most children thought people would be sympathetic to a girl who had her sexual image shared, girls perceived the outcomes for the female victim to be more negative than for boys. They were more likely to perceive that the victim would lose status, be teased and shamed, and for people to be angry with them. When the same questions were asked about a boy in the same situation, a lower proportion of participants (both boys and girls) thought that they would have these experiences.

The gendered nature of IIA and OSH was highlighted in the 2017 UKCCIS report, with the current work highlighting wider social and cultural contexts that shape children's online behaviour and experiences. It is important to recognise that boys also have negative experiences related to IIA, and in fact may be more likely than girls to experience sextortion (Patchin & Hinduja, 2020) and homophobic bullying (Project deSHAME, 2017). It is also possible that boys may be under-reporting IIA experiences. Nevertheless, the stark disparity between girls' and boys' experiences suggests there is a need to address the gendered dimensions of OSH in education and support for children who are victimised, consistent with other actions in society to address VAWG as outlined in government policy (e.g., the VAWG strategy [HM Government, 2021]). There is also a need to examine how non-binary children or those with other gender identities experience these forms of online victimisation and the related impacts.

3.3.2 Other factors

There is a general lack of research examining the extent to which factors associated with vulnerability to offline victimisation (e.g., parental conflict, loneliness) increase the risk of experiencing OSH and IIA. One notable exception is a study of Portuguese 12–20-year-olds, which found that being a victim of non-consensual sharing was related to behavioural and emotional problems, as well as experiences of neglect and abuse in childhood (Barroso et al, 2021). This highlights the importance of further research examining the role of these factors and how they may lead some children to experience multiple forms of online harm or polyvictimisation (Finkelhor et al, 2007).

3.4 OSH and IIA: Responses

The literature has identified several ways in which children may respond to OSH and IIA. The most common response is to delete unwanted messages or images they receive and block users who request or send unwanted images or make unwanted sexual comments (Budde et al, 2022; Project deSHAME, 2017).

Another commonly reported response is to do nothing (Project deSHAME, 2017; van Ouytsel et al, 2021). This can reflect the perception that there is nothing that can be done once images are non-consensually shared, as well as concerns over the judgements that will be made of the victim and potential police involvement (Project deSHAME, 2017). Non-consensual sharing is viewed by many children as an inevitable consequence of sending sexual images and is associated with victim-blaming by peers and adults (Budde et al, 2022; Ringrose et al, 2021a, 2022). For example, qualitative research has found that children often take a negative view of this behaviour and blame victims, claiming that they should not have created or shared sexual images (Budde et al, 2022; Lloyd, 2020). Similarly, an Israeli study found that school counsellors perceived victims to be responsible for sextortion if they had been actively involved in the exchange of sexual images (Dolev-Cohen et al, 2022). This is consistent with evidence suggesting that children are less likely to report their experiences to adults (e.g., school staff, family members) or social media platforms due to concerns about the reactions and judgements of others (Budde et al, 2022; Project deSHAME, 2017; Ringrose et al, 2022).

A lack of action in response to OSH and IIA may also be related to victims' feeling of self-blame, or a potential lack of recognition that this behaviour is abusive (Lloyd, 2020; Revealing Reality, 2023; Setty, 2019). Alternatively, or additionally, it may reflect the perception of many children that receiving unwanted requests and sexual images is a normal and inevitable consequence of being online, which has been identified as a barrier to reporting online abuse or seeking help (Budde et al, 2022; Project deSHAME, 2017; Ringrose et al, 2021a). One UK qualitative study found that inaction by schools can normalise the experience of IIA and discourage future reporting by victims (Lloyd, 2020).

> **"Rumours started about me having sex with a boy in the school. The rumours soon got out of hand and spread online and more throughout school. Some of them were unfortunately true but most of them were false and horrible that made me feel awful. I told a teacher and eventually they got sorted but some are still going around."**
>
> Girl, aged 15 (Project deSHAME, 2017)

3.5 OSH and IIA: Outcomes and impacts

A number of studies identified during the review period have examined the emotional, psychological and social impacts of victimisation through OSH and IIA. As indicated in Section 3.3.1, evidence suggests that negative outcomes and impacts are more severe for girls compared with boys (Burén & Lunde, 2018; Gámez Guadix et al, 2022; Ringrose et al, 2022).

Emotional responses

Barbovschi et al (2021) found that girls were more upset when receiving sexual messages than boys in a sample of European 12–18-year-olds. This is consistent with the results of the survey conducted by Revealing Reality (2021a), which found that girls were more likely to report negative emotional reactions than boys when receiving unwanted sexual images (70 per cent of girls compared with 17 per cent of boys). Other research suggests that girls initially react with bemusement or disgust when receiving unwanted sexual images, but desensitisation over time subsequently leads them to do nothing in response (Ricciardelli & Adorjan, 2019; Ringrose et al, 2021b). Girls also report negative feelings after complying with requests to send sexual images (e.g., self-blame, anxiety at school), and report experiencing negative consequences if they do not send images in response to pressure (e.g., spreading rumours that they have done this anyway) (Hunehäll Berndtsson & Odenbring, 2021).

Psychological impacts

Victimisation through OSH and IIA can have severe psychological impacts. This includes the experience of humiliation, helplessness, fear, distress and shame or self-blame (Burén & Lunde, 2018; Mandau, 2021; Nilsson et al, 2019). It has also been found to be associated with higher levels of depression and anxiety (Guerra et al, 2021; Ståhl & Dennhag, 2021), though these symptoms may be shortlived (Mitchell & Štulhofer, 2021). These are similar to the outcomes identified for victims of cyberbullying (see Section 6.2.1), although the sexual and gendered dimensions of OSH and IIA create specific impacts that disproportionately affect girls (Copp et al, 2021; Ståhl & Dennhag, 2021).

There is also evidence of a reciprocal relationship between IIA and mental health outcomes, e.g., depression and anxiety (Gámez-Guadix et al, 2022). However, it is unclear whether this is an outcome of victimisation, or whether psychological difficulties increase vulnerability to this experience as has been found for other forms of online sexual victimisation (de Santisteban & Gámez-Guadix, 2018).

Social and behavioural impacts

OSH and IIA can have a variety of social and behavioural impacts on victims. These include the experience of reputational damage, peer harassment and exclusion, negative peer judgements and victim blaming, as well as being the target of rumours and gossip, which intensify the negative impacts of victimisation (Quayle & Cariola, 2019; Strassberg et al, 2017). IIA has also been found to be associated with subsequent unwanted peer sexual advances and pressure to send additional sexual images (Gámez-Guadix & Mateos- Pérez, 2019; Van Ouytsel et al, 2017). The potential for revictimisation through the forwarding of images, or through receiving 'likes' and comments on the images by peers, can further exacerbate the emotional and psychological impacts of victimisation (Quayle & Cariola, 2019; Strassberg et al, 2017). These effects are further intensified by the loss of control of images and their ability to circulate online indefinitely, as well as the difficulties of deleting content across multiple platforms (Wolak et al, 2012). Impacts are similar to those experienced by children who experience TA-CSA and victimisation through CSAM production, indicating that revictimisation is ongoing with each viewing of the shared image (see Chapters 4 and 5).

3.6 OSH and IIA: Current knowledge gaps

- There is a need for more research examining the prevalence, experience and impacts of OSH and IIA on children in the UK as there has been limited development of the evidence base in this area since 2017.
- It is also important to develop a clearer understanding of the prevalence of exposure to risk and harm on individual platforms, as well as how platform design features facilitate (or prevent) behaviours associated with OSH and IIA. This evidence will be needed when the new regulatory system is implemented with the enactment of the Online Safety Act. Ofcom will need to establish that risk is high on specific platforms and identify why this is the case if they are to require technology companies to identify and change features that facilitate victimisation.
- More empirical studies are needed that examine the psychological, social and environmental factors that may increase the vulnerability of children to online victimisation through OSH and IIA.
- There is also a gap in knowledge regarding the overlap between peer-to-peer OSH on the one hand, and adult-to-child TA-CSA on the other.

Chapter 4
Technology-Assisted Child Sexual Abuse

This chapter provides an overview of the available evidence related to technology-assisted child sexual abuse (TA-CSA). It examines the scale of the problem, the platforms involved, factors that may increase susceptibility to risk exposure and harm, as well children's responses and outcomes. This is based on academic and grey literature published since 2017 and draws on research using UK and non-UK samples.

As described in Chapter 2, in this report TA-CSA refers to sexual exploitation of children by adult perpetrators. The evidence presented in this review focuses on victimisation by adult perpetrators only, while recognising the difficulties of identifying the age of offenders in some situations. Unwanted sexual conduct by peers is categorised in this report as OSH or IIA instead, and the related evidence is reviewed in Chapter 3.

Chapter summary

- **Scale**: Evidence published since 2017 from outside the UK suggests that 5–18 per cent of children experience online sexual solicitation and 5–25 per cent have online sexual interactions with adults.
- **Scale**: While various data sources on the scale of TA-CSA show upward trends over time, it is difficult to say with confidence whether victimisation in the UK has increased. This is because factors like greater public awareness and reporting, as well as the increased success of investigations, may partly account for some of those trends.
- **Platforms**: The limited data that is available on this topic indicates that sexual approaches by adults occur more commonly on platforms that are widely used by children (most notably, Snapchat and Meta-owned platforms). We currently lack detailed empirical investigation of the extent to which gaming platforms and direct messaging services are used by adults for TA-CSA.
- **Children most likely to be harmed**: Research suggests that older adolescents and girls more commonly experience TA-CSA. A variety of psychological and environmental factors can increase the vulnerability of children to TA-CSA, although there is a general lack of empirical research examining this.
- **Responses**: The evidence base, although underdeveloped in this respect, suggests that children are more likely to block contacts when they are concerned about sexual solicitation by adults, than to use reporting functions on platforms or seek help offline.
- **Outcomes and impacts**: The available literature suggests that TA-CSA can have significant and long-term psychological impacts on victims that do not differ significantly from those related to offline sexual abuse and exploitation. The negative reactions of others after disclosure, and the possibility of revictimisation through the ongoing distribution of any CSAM linked to their abuse, can exacerbate these negative outcomes.

4.1 TA-CSA: Scale

While it is difficult to estimate how widespread TA-CSA is due the secretive nature of offending, there are a range of data sources to draw on that can give an indication of its scale.

Details of the prevalence data discussed below, and of the design of the specific studies from which the figures are derived, are provided in Appendix 3.

4.1.1 Police recorded crime

In 2021/22, 13,110 child sexual offences were flagged as ‘online crime’ in police recorded crime figures relating to England and Wales (ONS, 2023a). This figure reflects a range of sexual offences against children. More specific figures for the offence of ‘sexual communication with a child' are available from NSPCC's analysis of data provided by police forces in England and Wales, Scotland and Northern Ireland in response to Freedom of Information Act (FOIA) requests: 6,350 such offences were recorded in 2022/23 (NSPCC, 2023a). Details of the number of related convictions and sentences for these offences have not been published.

Analysis of FOIA police data by the NSPCC shows that 'sexual communication with a child' offences increased by 84 per cent between 2017/18 and 2021/22 (NSPCC, 2022a). Published police recorded crime figures for 'sexual grooming' offences[15] have also increased year-on-year, rising from 4,478 in 2017/18 to 6,910 in 2021/22 (ONS, 2023b). The upward trends are likely to reflect increases in reporting linked to greater public awareness that online grooming is a crime, as well as the increased success of investigations; they cannot automatically be attributed to a rise in perpetration of TA-CSA or increase in the number of victims.

4.1.2 Reports to National Center for Missing and Exploited Children (NCMEC)

The most recent figures from NCMEC (2022a) indicate that 80,524 reports were received in 2022 related to 'online enticement of children for sexual acts' (which is the US legal category covering TA-CSA). There has been an uneven trend of increases in reports over recent years, with a large increase between 2019 and 2020 (98 per cent), a lower increase between 2020 and 2021 (17 per cent), and another large increase of 82 per cent between 2021 and 2022 (NCMEC, 2022a, 2022b). As the majority of reports come from platforms themselves, these are likely to reflect improved scanning and detection processes rather than solely indicating increases in levels of offending.

The COVID-19 pandemic in 2020 saw a surge in the number of reports of TA-CSA to organisations like NCMEC and the Internet Watch Foundation (IWF) (WeProtect Global Alliance, 2021). The volume of CSAM reported to NCMEC saw a corresponding rise (see Section 5.1.2). During this period, livestreaming and first-person produced imagery increased dramatically (IWF, 2022), as did every indicator of online child sexual abuse monitored by Europol (Europol, 2020). This may reflect a change in the amount of time that children and perpetrators were spending online, and in how victims and perpetrators met and interacted online, as well as the type of grooming strategies used. However, it is difficult to determine the extent to which increases in reports represent changes in levels of offending due to the nature of the different data sources and the potential influence of other factors.

4.1.3 Prevalence of sexual solicitation by adults

The main source of evidence for estimating the prevalence of TA-CSA is research examining the proportion of children who have experienced sexual approaches and requests from adults. There is a general lack of academic research on this topic in the UK. However, since 2017 there has been a developing evidence-base in Europe and the USA. A variety of measures of solicitation are used, and these include different combinations of specific behaviours. Some studies specify the involvement of adults only, while others include peers

[15] The 'sexual grooming' category reflects offences of sexual communication between an adult and a child under 16, and offences where the offender met or arranged to meet a child with the intention of sexually abusing them.

or do not ask about the identity of perpetrator(s), creating challenges when interpreting the research evidence (Stoilova et al, 2021).

The sexual solicitation figures in this evidence review are derived from studies where the perpetrator was known to be an adult (or, in the case of one study, someone more than five years older than the victim). Retrospective studies in the UK and USA that use representative and non-representative samples suggest 5–17 per cent of young adults experienced sexual solicitation when they were children (Davidson et al, 2017; Finkelhor et al, 2022; Greene-Colozzi et al, 2020). Non-representative samples of Spanish children aged 12–17 suggest that 7–18 per cent of children had this experience over the past year (Calvete et al, 2021; de Santisteban & Gámez-Guadix, 2018; Gámez-Guadix & Mateos-Pérez 2019; Ortega-Barón et al, 2022).

4.1.4 Prevalence of sexual interactions with adults

There are fewer studies that examine the prevalence of online sexual interactions between children and adults, and no available evidence from the UK. Research from non-representative samples of Spanish children aged 12–17 years suggests that 5–8 per cent had sexual interactions with adults within the previous year (de Santisteban & Gámez-Guadix, 2018; Ortega-Barón et al, 2022).

Some of these studies used longitudinal research designs, addressing an identified gap in the literature (Livingstone et al, 2017; May-Chahal & Palmer, 2018). The studies had mixed results: while some show that children's experiences of sexual solicitation and interactions with adults can remain stable over time (Gámez-Guadix & Mateos-Perez, 2019), others suggest that experiences of victimisation increase in number as children get older (Calvete et al, 2019; Ortega-Barón et al, 2022). Where an increase has been found from one time point to the next, this has been attributed to the sexual and social development of research participants over the course of the study period (Ortega-Barón et al, 2022) or interpreted as evidence of revictimisation (Gámez-Guadix & Mateos-Perez, 2019).

Two other studies published since 2017 suggest that as many as 23–25 per cent of children had an online sexual relationship with an adult during the course of their childhood that was either sustained (Greene-Colozzi et al, 2020) or secretive (WeProtect Global Alliance, 2020).[16]

> **"So I was at home, in my room, and I get a Facetime call from a guy, and he was rubbing his belly, but it was a girl account. He just showed me his face. He opened up his top and started rubbing his belly. And he was like do you want me to open my trousers? And I just blocked him."**
>
> Girl, Year 8 (Ringrose et al, 2021a)

[16] These figures may include experiences at ages 16 and 17. Notably, sexual interaction between an adult and a 16- or 17-year-old would not constitute a 'sexual communication with a child' offence in the UK.

4.2 TA-CSA: Platforms

Academic research does not generally provide evidence about the platforms on which TA-CSA occurs. This is due to difficulties associated with accessing relevant data. Where such information is available, it is those platforms that are most popular with children that appear to be the location of the most frequent sexual approaches by adults. For example, Hamilton-Giachritsis et al (2020) reported that the victims in their UK sample experienced initial contact through a variety of channels (e.g., online gaming, social media). Joleby et al (2021) found that the platforms involved in the criminal cases they examined included Facebook, Instagram, Kik, Momio and Omegle. Another study found that online sexual interactions between children and someone they believed to be an adult were most frequently reported on Snapchat (15 per cent) and Instagram (13 per cent), with slightly lower percentages for WhatsApp (11 per cent), Facebook (10 per cent) and Facebook Messenger (10 per cent) (Thorn, 2021).

Snapchat and Meta-owned platforms were the platforms most commonly associated with offences of Sexual Communication with a Child, according to an analysis of police recorded crime figures obtained by the NSPCC through FOIA requests in 2021/22.[17] Snapchat was involved in 33 per cent of cases where the platform was recorded; while Facebook, Instagram and WhatsApp were involved in 38 per cent of cases (NSPCC, 2022a).

Earlier research highlighted the use of gaming environments by perpetrators to contact potential victims (Webster et al, 2012). However, the present review found no recent evidence related to sexual solicitation or TA-CSA on gaming platforms, highlighting an important area for future research.

Cross-platform risk or off-platforming

While a variety of platforms and messaging apps are involved in TA-CSA, there is also evidence of 'cross-platform risk' or 'off-platforming' where perpetrators encourage relationships to migrate from public platforms (e.g., Facebook) to private encrypted messaging apps where communication cannot be monitored (e.g., WhatsApp, Telegram) (see Section 7.2.4 for a discussion of the detection challenges associated with encrypted services). These private environments reduce opportunities for detection and facilitate the development of greater intimacy and secrecy between perpetrators and victims, as well as greater security for sexual interactions and image exchange (Kloess et al, 2019; Thorn, 2022; WeProtect Global Alliance, 2021). Thorn (2022) found that 65 per cent of their sample of 9–17-year-olds had been asked to move their interactions in this way by an online-only contact, and 52 per cent of the sample subsequently did this.

[17] Note that law enforcement agencies are not required to record platform-level data for offences, and do not collect this data consistently.

4.3 TA-CSA: Children most likely to be exposed to risk or experience harm

4.3.1 Demographic factors

- **Age/developmental stage:** Several studies have found that older adolescents are more likely to experience adult sexual solicitation and sexual interactions compared with younger children (Calvete et al, 2021; de Santisteban & Gámez-Guadix, 2018). This is likely to reflect engagement in a wider range of online activities and use of platforms by older adolescents, as well as the developmental tasks that characterise this period e.g., exploration of sexual identity, romantic relationships (Hamilton-Giachritsis et al, 2020; Whittle et al, 2013). However, younger children are also increasingly experiencing sexual solicitation and grooming, as indicated by the 13 per cent increase in the number of children aged 7–10 in the CSAM assessed by the IWF in 2022, which was most likely produced as a result of grooming (IWF, 2023).
- **Gender:** Evidence suggests that girls are more likely to report adult sexual solicitation and interactions than boys (Gámez-Guadix & Mateos-Pérez, 2019; de Santisteban & Gámez-Guadix, 2018; Thorn 2022). Moreover, older girls are more likely to do so compared with those who are younger (Gámez-Guadix & Mateos-Pérez, 2019; Thorn, 2021b).
- **Sexual orientation:** Some studies have examined whether sexuality influences frequency of the experience of sexual solicitation, with evidence suggesting that this is more common in LGBTQ+ children. For example, Thorn (2021b) found that 32 per cent of LGBTQ+ participants reported an online sexual interaction with someone they believed to be over 18, compared with 22 per cent of non-LGBTQ+ participants.
- **Ethnicity and disability/SEN:** The review did not identify any academic studies that examined the influence of ethnicity or disability status on the frequency of sexual solicitation or experiences of TA-CSA, with the majority of studies focusing on white, non-disabled children in Western countries. The extent to which there may be differences according to these factors is generally unknown and exploring this further can inform the development of appropriate prevention and response strategies for specific groups of children who may be at greater risk of harm (Lundy et al, 2019).

4.3.2 Online behaviours that can expose young people to risk and harm

Research has examined how specific routine behaviours online can increase the vulnerability of children to TA-CSA by creating opportunities for sexual solicitation and contact with potential perpetrators (Stoilova et al, 2021). More frequent engagement in activities like accepting friend requests from unknown contacts and sharing personal information was found to be associated with higher levels of adult sexual solicitation (DeMarco et al, 2017; de Santisteban & Gámez-Guadix, 2018; Longobardi et al, 2021; Stoilova et al, 2021). These behaviours are often referred to as ‘risky online behaviours’ in the literature because of their ability to expose children to sexual solicitation and TA-CSA (and other risks); however, this does not imply that they are to blame for any resulting harm, as these are routine online activities and encouraged by the design and architecture of platforms. Psychological and environmental factors will also influence the extent to which engagement in such behaviours leads to risk, exposure and subsequent harm (see below).

4.3.3 Psychological and environmental factors

There are psychological (e.g., social anxiety, depression) and environmental factors (e.g., family structure, parental conflict, exposure to violence and neglect) that potentially increase vulnerability to TA-CSA (Livingstone et al, 2017; Stoilova et al, 2021). These may lead children to engage with adult sexual solicitation or be susceptible to grooming techniques by perpetrators that focus on providing the support, attention and affection that may be missing in other areas of their lives (Chiu & Quayle, 2022).

There is a general lack of research specifically examining how combinations of these different factors influence experience of TA-CSA. Some studies have been published during the review period that examine specific vulnerability factors (e.g., family functioning, offline sexual abuse) (Augusti et al, 2021; Jonsson et al, 2019), but they do not clearly distinguish between adult and peer sexual exploitation. For example, Augusti et al (2021) found that being female, having a lower family income and lower family functioning (e.g., poor parental mental health) were significant predictors of online sexual victimisation.

Overall, these studies suggest that children with specific vulnerabilities, or in specific situations that exacerbate them, may be at greater risk of TA-CSA. Research additionally suggests that there is a relationship between online and offline vulnerability to sexual exploitation, with children who are already at risk offline being similarly vulnerable in digital environments (Livingstone et al, 2017; Stoilova et al, 2021). Clearly, not all children who receive sexual approaches subsequently engage with adults (de Santisteban & Gámez-Guadix 2018): this highlights the need to develop greater understanding of the characteristics of those that do.

4.4 TA-CSA: Responses

There has been little research examining children's responses to online sexual solicitation by adults published since the 2017 review.

Evidence shows that some children engage with adult approaches for a variety of reasons. These include developmental (e.g., curiosity, sexual experimentation) or vulnerability factors (e.g., loneliness), boredom, or thrill-seeking (Greene-Colozzi et al, 2020; Kloess et al, 2017).

The main response covered by the literature relates to children's help-seeking and reporting behaviour. The evidence suggests that when children are concerned about this type of contact, they are more likely to use block or ignore functions, and less likely to opt for reporting functions or to seek help offline. A retrospective sample of young adults found that 59 per cent of those who had received sexual content from adults or someone unknown when they were younger deleted or blocked the sender in response, while fewer reported the problem online (28 per cent) or told a trusted adult or peer (23 per cent) (WeProtect Global Alliance, 2020). A study of children's online sexual interaction conducted by Thorn (2021b) – which did not differentiate between interaction with adults and peers – found that 83 per cent of participants used platform-based tools in response to unwanted sexual contact (e.g., 65 per cent blocking; 47 per cent reporting); over a third sought help offline from parents, carers or friends (37 per cent); around a quarter told no one about it (26 per cent); and 14 per cent ignored the sexual contact. When asked how they would respond if they received requests for sexual interactions from an adult, younger children were more likely than older adolescents to say they intended to tell a parent.

4.5 TA-CSA: Outcomes and impacts

There are numerous ethical and safeguarding challenges associated with working with children who have been victimised. This may partly explain why there has been so little empirical research on the impacts of TA-CSA in the UK and other countries.

Although under-developed, the available evidence demonstrates that TA-CSA can have significant and long-term psychological impacts on victims (Hamilton-Giachritsis et al, 2020). A study conducted in Spain found that health-related quality of life decreased among children aged 12–15 after experiencing sexual solicitation or interaction with adults, and that this risk was double compared with children who did not have those experiences (Ortega-Barón et al, 2022). Guerra et al (2022) found that being asked sexual questions by an adult was a significant predictor of depression in a non-representative sample of Brazilian 15–17-year-olds. Research further suggests that psychological distress (e.g., depression) may not just be an outcome, but also act as a risk factor to further sexual exploitation (Chiu & Quayle, 2022; de Santisteban & Gámez-Guadix, 2018).

Recent research has shown that TA-CSA can also occur in virtual reality environments (Allen & McIntosh, 2023). The immersive and embodied nature of these spaces (Steffen et al, 2019; Dincelli & Yayla, 2022) can cause abusive sexual interactions to be experienced in ways that are similar to offline victimisation and have the same impacts (Pettifer et al, 2022).

This is consistent with evidence from a recent UK multi-method study that the significant and negative psychological impacts of TA-CSA (e.g., anxiety, depression, PTSD symptoms, shame) do not differ significantly from those associated with offline CSA (Hamilton-Giachritsis et al, 2020). Despite this, research with professionals shows that TA-CSA is often perceived as lower risk, less serious, or of less immediate concern than offline CSA due to the lack of physical contact (Hamilton-Giachritsis et al, 2020).

While the experience of TA-CSA itself can cause harm, the negative reactions of others after disclosure of TA-CSA have been found to intensify the psychological impacts on the victim (Hamilton-Giachritsis et al, 2020). Moreover, the potential production of CSAM as a result of TA-CSA, and the possibility of revictimisation through their ongoing distribution online, could also exacerbate these negative outcomes (Hamilton-Giachritsis et al, 2020).

> **“Like as soon as I went on [Habo Hotel gaming site], em, I was approached by like 40, 50 year olds that were on there and wanted me to like show myself on webcam, and I was only nine at the time, but I sort of just did it anyway”**
>
> Girl, aged 17 (Hamilton-Giachritsis et al, 2017)

4.6 TA-CSA: Current knowledge gaps

- There is a need for research that examines TA-CSA in the UK. Cross-sectional and retrospective studies of children's experiences of online sexual solicitation and interaction with adults could provide insight into how widespread this is and the characteristics of children most likely to be victimised. The use of longitudinal designs would be beneficial in developing greater understanding of the influence of specific vulnerability factors and pathways to harm.
- There is a current lack of evidence about sexual interaction between children and adults on specific platforms, particularly on direct messaging services, which provide perpetrators with a private environment in which to offend. This is needed in order to understand how platform functionalities and design features facilitate the grooming process.
- Despite earlier research suggesting that perpetrators use online gaming platforms, gaming interests and performance to make contact with potential victims (Webster et al, 2012), there remains a lack of evidence related to the prevalence and nature of TA-CSA in these environments.
- Given the evidence that livestreaming has emerged as a significant location for TA-CSA, more empirical evidence is needed about how grooming strategies are used in these situations, as well as the potentially group-based nature of exploitation.
- The potential for virtual reality environments to facilitate TA-CSA also requires further research to assess how they can facilitate offending behaviour, as well as the way in which immersion and embodiment may influence the psychological impacts of victimisation.

Chapter 5
Child Sexual Abuse Material

This chapter examines the exploitation of children through the production, possession and distribution of Child Sexual Abuse Material (CSAM). It provides a review of the available evidence related to the amount of material in circulation, victim characteristics, as well as the outcomes and impacts for children who are sexually exploited in this way. This is based on academic and grey literature published since 2017 and draws on research from the UK and beyond.

Chapter summary

- **Scale**: Various data sources show that a large amount of CSAM is being produced, circulated and viewed online. This suggests that a sizeable number of adults have an interest in this material and, consequently, there are many undetected victims. However, the different data sources have limitations and can only provide a partial estimation of the volume of material in circulation.
- **Scale**: Most data sources show upwards trends over time in the amount of CSAM being detected. While this could indicate a change in the number of perpetrators or victims, it may also reflect improvements in scanning and detection tools and investigations, as well as greater public awareness and reporting.
- **Scale**: There has been a large increase in the amount of 'first-person produced' CSAM identified in recent years, particularly since the COVID-19 lockdowns. It is difficult to determine the extent to which this material is voluntarily produced and then shared non-consensually by peers or adults, or results from TA-CSA.
- **Platforms**: Both the clear and dark web are important channels for the production and distribution of CSAM. Some of the biggest global platforms, such as Facebook and Instagram, make proactive efforts to scan and detect CSAM and are, therefore, able to report large volumes of CSAM to the authorities (NCMEC).
- **Children most likely to be harmed**: Girls are more frequently depicted in CSAM and exploited through its production, although boys are also victims. A significant proportion of material depicts prepubescent children, with abuse severity often higher for the younger or youngest victims.
- **Responses**: It is unusual for victims to report their experiences. This may be because they are not aware that their image is in circulation, or recognise that its production constituted abuse, but there are also barriers to disclosure (threats by perpetrators, shame or self-blame, or concerns of being taken seriously).
- **Outcomes and impacts**: Evidence suggests that victimisation through the production and distribution of CSAM has significant psychological impacts similar to those of TA-CSA. These are further compounded by the persistence of images and the potentially indefinite circulation of material online.

5.1 CSAM: Scale

There are different data sources that can be used to indicate the scale of CSAM in circulation. These suggest that a large amount of this material is being identified though detection, reporting tools, investigations or online monitoring and web crawling. Each source has limitations (see Wager et al, 2018), suggesting that individually, or taken together, they can only provide a partial picture of the volume of CSAM being produced and disseminated at any given time. We cannot deduce from them what proportion of children in the population are being victimised in this way.

Most sources show upwards trends in the amount of CSAM in circulation over recent years. While this could indicate a change in the number of perpetrators or victims, it may also be linked to improvements in scanning and detection tools, better investigations, as well as greater awareness by the public that CSAM should be reported (ONS, 2023c). It also reflects the rise in offending during the COVID-19 lockdowns, when perpetrators had more opportunities to interact with victims online (WeProtect Global Alliance, 2021).

5.1.1 Police recorded crime

In 2021/22, 30,925 offences involving indecent images of children were recorded by the 41 police forces in England and Wales, Scotland and Northern Ireland who responded to an NSPCC Freedom of Information Act (FOIA) request (NSPCC, 2023b).

Analysis of FOIA police data by the NSPCC shows that offences relating to indecent images of children have increased by 66 per cent between 2016/17 and 2021/22 (NSPCC, 2023b). Increases can also be seen in published police recorded crime figures for Obscene Publications[18] in the UK, with recorded crimes in this category rising from 23,236 in 2016/17 to 40,472 in 2022/23 (Home Office, 2023a). The subcategory of Obscene Publications flagged as 'online crime' has similarly risen, from 18,222 in 2017/18 (ONS, 2020) to 25,311 in 2022/23 (ONS, 2023d).

5.1.2 Reports to the National Center for Missing and Exploited Children (NCMEC)

The most recent figures provided by NCMEC indicated that there were 31,901,234 reports related to CSAM in 2022 (NCMEC, 2022a). There has been a general trend of yearly increases in reports, with large increases between 2019 and 2020 (28 per cent), and between 2020 and 2021 (35 per cent), and a lower increase of 9 per cent between 2021 and 2022 (NCMEC, 2022a, 2022b).

5.1.3 Unique images

The Child Abuse Identification Database (CAID) holds data about CSAM obtained by the UK police, the National Crime Agency (NCA) and approved industry bodies during criminal investigations; the database is used to help identify victims and perpetrators of sexual abuse (ONS, 2020; Home Office, 2018). The most recent figures available regarding CAID indicate

[18] It should be noted that 'Obscene Publication' police data does not disaggregate offences that were specifically against, or relate to, children, as this category of offence includes other types of indecent images (e.g., extreme pornography).

that 8.3 million unique indecent images were uploaded to the database between 2015 and 2019; 7,600 children in the UK were safeguarded as a result of related investigations in the year ending September 2019 (NCA, 2021). Figures for the current number of images in the database are not available, but it has been estimated that the database increases by an average of half a million images every two months (Home Office, 2019b).

5.1.4 Webpages on the open web

Web crawlers, such as those developed by the Internet Watch Foundation (IWF) and Project Arachnid, are automated tools that scan the open and dark web[19] for webpages hosting known and first-generation CSAM images (IWF, 2022; Canadian Centre for Child Protection, 2023). These organisations also provide relevant data about the volume of material they identify in their yearly reports.

The most recent figures from the IWF indicate that they identified 255,571 webpages containing CSAM in 2022 (IWF, 2023). As each individual webpage can contain multiple images, the number of individual images these figures represent will be much higher. The IWF figures show a similar trend of yearly increases to those provided by NCMEC. There was a 16 per cent increase between 2019 and 2020, a 64 per cent rise between 2020 and 2021 (reflecting the influence of the COVID-19 pandemic), and a much smaller increase of 1 per cent between 2021 and 2022 (IWF, 2021, 2022, 2023).

Project Arachnid identified 5,257,763 webpages that contained CSAM on the open web between 2018 and 2020 (Canadian Centre for Child Protection, 2021), with an upward trend in detections across this period.

It is worth noting that the last few years have seen an increase in the amount of first-person produced imagery within the webpages that host CSAM. This has been attributed to the impact of COVID-19 lockdowns (IWF, 2022). Of the 255,571 webpages assessed as containing CSAM in 2022, 78 per cent contained first-person produced imagery, with a large proportion being captured from livestreaming (IWF, 2023). The proportion of first-person produced imagery rose by 28 per cent between 2020 and 2021, and by a further 6 per cent between 2021 and 2022 (IWF, 2022, 2023).

5.1.5 Webpages and forum traffic on the dark web

The volume of CSAM that is directly hosted or accessed on the dark web is relatively small compared with the open web. The IWF identified 1,067 hidden services displaying CSAM in 2022, though figures have increased yearly: by 27 per cent between 2020 and 2021, and by a further 15 per cent between 2021 and 2022 (IWF, 2022, 2023). Project Arachnid made 48,770 CSAM detections on TOR in 2020 (Canadian Centre for Child Protection, 2021).

While the dark web may not host high volumes of CSAM, it can facilitate its distribution (Canadian Centre for Child Protection, 2021). It does this by directing users to the presence of this material on the open web and anonymising the traffic between users and the websites they are visiting. The level of activity on the dark web related to CSAM is demonstrated in a

[19] The dark web is a part of the internet that hosts websites (or 'hidden services') that are not indexed by search engines. Hidden services need specific software, configurations, or authorisation to access. The identities and locations of users who visit the dark web stay anonymous and cannot be tracked due to encryption.

study that examined traffic on one dark web forum for 13 days and found that each thread received an average of 10,234 views during this period (Owens et al, 2023).[20]

5.2 CSAM: Platforms

There are a variety of online distribution channels for CSAM. Images can be spread through links or posts on the open web (e.g., on social media platforms) and exchanged between individuals and groups using encrypted messaging apps (e.g., WhatsApp, Telegram), or through the use of peer-to-peer filesharing networks and the dark web (e.g., IWF, 2022; NCA, 2021). These media allow perpetrators to access, download and share CSAM (Martellozzo et al, 2020; NCA, 2021).

A recent report by NCMEC (2023) found that 90 per cent of the reports they received for CSAM in 2022 came from the largest platform providers (e.g., Facebook, Instagram, Google, WhatsApp). This reflects the fact that some of the biggest platforms with global reach proactively use scanning and detection tools, whereas many other platforms do not. The volume of reports received from platforms (31.8 million in 2022) demonstrates that the open web is widely used to distribute CSAM and facilitate livestreaming.

Owens et al (2023) highlighted the trend of the migration of hosting sites and communities to the dark web, with TOR hidden services[21] being identified as the preferred platform for perpetrators to communicate, share links and be directed to sources where they can easily access large amounts of CSAM (Mordock, 2019; Owens et al, 2023; WeProtect Global Alliance, 2018). Although it has been estimated that these hidden services represent approximately 1 per cent of websites within the TOR network, they have higher traffic volumes than other categories of site (Intelliagg, 2016; Owens et al, 2023).

> **“There’s evidence of it [imagery linked to TA-CSA victimisation] and I don’t know who else has seen that…It does make me feel a bit sick. I don’t know, it does stress me out. I think it makes me anxious because I don’t know what was recorded, when it was recorded.”**
>
> Girl, aged 18 (Hamilton-Giachritsis et al, 2017)

[20] It is important to note that number of views does not necessarily equal the number of individual users accessing an image as it could be viewed multiple times by the same person.

[21] TOR, short for ‘the onion router’, is the name of the largest network of websites (or ‘hidden services’) on the dark web.

5.3 CSAM: Children most likely to be exposed to risk or experience harm

The review identified a small amount of evidence published since 2017 on the characteristics of the victims of the production and distribution of CSAM. The evidence is mainly derived from annual reports provided by the IWF, as well as studies that examine the CSAM collections of convicted perpetrators, or samples of images from enforcement databases (e.g., IWF, 2023; Tejeiro et al, 2020). The figures below suggest that prepubescents and girls are more likely to be victimised through CSAM, although this is likely to depend on the specific victim gender and age preferences of individual perpetrators and/or distribution channels.

5.3.1 Demographic factors

- **Age:** Evidence shows that a large proportion of victims in these studies are prepubescents or younger. The mean age of victims is around 10 (Tejeiro et al, 2020) or 11 (Quayle et al, 2018). The IWF found that most images (56 per cent) were of 11–13-year-olds, but that two fifths were of children aged under 10 (IWF, 2023). Tejeiro et al (2020) found that 36 per cent of images showed those aged 10–13, but the biggest age group in the images examined (45 per cent) were of younger victims aged 4–9 years. Both studies highlighted that abuse severity increased as victim age decreased (IWF, 2023; Tejeiro et al, 2020).
- **Gender:** The evidence suggests that the majority of victims shown in CSAM are girls (IWF, 2023; Tejeiro et al, 2020). The most recent IWF report found that among the 255,571 webpages identified as CSAM in the previous year, 96 per cent of images showed girls (IWF, 2023). Only a small proportion of images showed boys (3 per cent), but a higher proportion of these were classified as Category A (those showing the most severe abuse).[22]

5.3.2 Other factors

- **Family relationships:** There are few studies that examine the relationship between victims and perpetrators in relation to the production of CSAM, and only one UK study was identified by the review during the specified time period. The available evidence suggests that there are a variety of contexts of production and that these vary by age, with a large proportion of this material being produced by family members or other known adults.

  Quayle et al (2018) found that images of children aged seven and under were more likely to be taken by family members (69 per cent), whereas images of those aged 12–17 were more likely to be produced by victims who had been groomed (48 per cent). Gewirtz-Meydan et al (2018b) found that 52 per cent of CSAM victims reported that the perpetrator was a family member and 41 per cent had been victimised by an acquaintance. This is consistent with evidence from NCMEC (2022a) that the majority of CSAM examined in 2022 was produced within family and local community contexts, although a sizable proportion was first-person produced or resulted from online grooming.

[22] Category A images show penetrative sexual activity with children, as well as those involving sexual activity with an animal or sadism. Category B images depict non-penetrative sexual activity, and Category C images are those that do not fall within categories A or B (Sentencing Guidance Council, 2014). These classifications are used when sentencing perpetrators and in research examining perpetrator collections.

5.4 CSAM: Responses

There has not been specific empirical attention to victim responses during the review period. The evidence suggests that it is unusual for victims to report their experiences: Gewirtz-Meydan et al (2018b) found that victimisation was only reported to police or welfare agencies by 23 per cent of participants in their study, with 14 per cent indicating that no one else knew about it, and 41 per cent saying they felt embarrassed about police, social workers and others seeing the images.

The evidence did identify a number of challenges that CSAM victims face in terms of reporting and seeking help. If the victim is very young, or if CSAM is produced during the online grooming process or captured during livestreaming (a process known as 'capping'), the victim may not be aware that images have been produced and/or are being disseminated online (IWF, 2023). Even where victims are aware, they may be manipulated into believing that image production and sharing is part of a romantic relationship, and not recognise this interaction as exploitative (Bryce, 2010; Webster et al, 2012). Alternatively, victims may be reluctant to report their experiences because of direct or implied threats by perpetrators, the experience of self-blame and shame, or concerns that their reports will not be believed or taken seriously (Bryce, 2010).

A recent development enabling children to take action to report and remove their sexual images from online circulation is the Report Remove tool (IWF, 2023). This is an anonymous service that enables children to report their sexual image without shame, blame or other repercussions. Images are assessed by the IWF and given unique image hashes (equivalent to a digital fingerprint) if classified as illegal (see Section 7.2.4 for more information on image hashing). The hashes are then shared with platforms to enable the removal of images and prevent reposting and sharing. The IWF received 187 reports through the Report Remove tool in 2022 and action was taken on 101 reports (IWF, 2023). This tool is important as it provides children with the ability to take action in response to this form of exploitation, and its use will hopefully increase as public awareness of its availability increases.

5.5 CSAM: Outcomes and impacts

There is a general lack of research examining the impacts of children having sexual images of themselves in circulation. However, as CSAM is generally produced as the result of contact and/or non-contact offending, the same emotional and psychological outcomes have been identified in the literature as those experienced by victims of TA-CSA (e.g., guilt, shame, self-blame, depression) (Gewirtz-Meydan et al, 2018b; Hamilton-Giachritsis et al, 2020).

There are also specific aspects of CSAM victimisation that can intensify these impacts. Research suggests that victims experience anxiety about being recognised from images or being seen as willing participants in the abuse depicted (Gewirtz-Meydan et al, 2018b). This may be particularly relevant when CSAM is first-person produced and/or captured through the livestreaming of abuse that appears to be voluntary. The impacts are also exacerbated by the permanence of images, lack of control over their online circulation, and associated revictimisation (Hamilton-Giachritsis et al, 2020). This highlights the need for these specific aspects of victimisation to be addressed by practitioners developing interventions and supporting victims (Gewirtz-Meydan et al, 2018b).

5.6 CSAM: Current knowledge gaps

- The production of CSAM in the context of TA-CSA, as well as that which appears to be voluntarily produced and/or is the result of IIA, highlights the need for research examining the intersection between these different modes of production.
- There is also a need to develop greater understanding of how perpetrators interact with each other on the clear and dark web, including how they share links and direct each other to CSAM.
- This review found almost no evidence related to the scale of production of CSAM through generative AI, deepfakes, and Non-Photographic Abuse Imagery (NPAI) of children. More research is needed on these topics, and greater understanding on how perpetrators engage with this material.
- Although there is some evidence relating to the impacts of CSAM victimisation and barriers to reporting, further research is needed to inform the development of more effective support for victims and educational resources for children on this topic. Empirical research examining the knowledge gaps identified in this chapter are important in informing these developments, as is the need to develop further understanding of victim perceptions of barriers to reporting, the support they receive and how this can be improved.

Chapter 6
Other Categories of Online Risk

This chapter provides an overview of research relating to other types of risk that children can be exposed to online. This information is intended to complement and provide context to the evidence on sexual online risks presented in the previous chapters and enable a comparison of their relative scale and the harm they can cause.

Since this evidence review aims to provide a baseline of the online risk landscape to children in the UK at a point immediately before the introduction of regulation through the Online Safety Act, only those categories that are within the scope of the Act are considered here (DSIT, 2023). Consequently, this chapter covers 'primary priority' content (e.g., online pornography, content that encourages suicide, self-harm and eating disorders), and 'priority' content (e.g., cyberbullying). It also covers content that promotes extremism and radicalisation, as an example of illegal content not covered in the previous chapters. The fact that other types of harmful content and conduct are not covered (e.g., online fraud, other forms of cybercrime) should not be taken to imply that they are less prevalent or intrinsically less harmful.

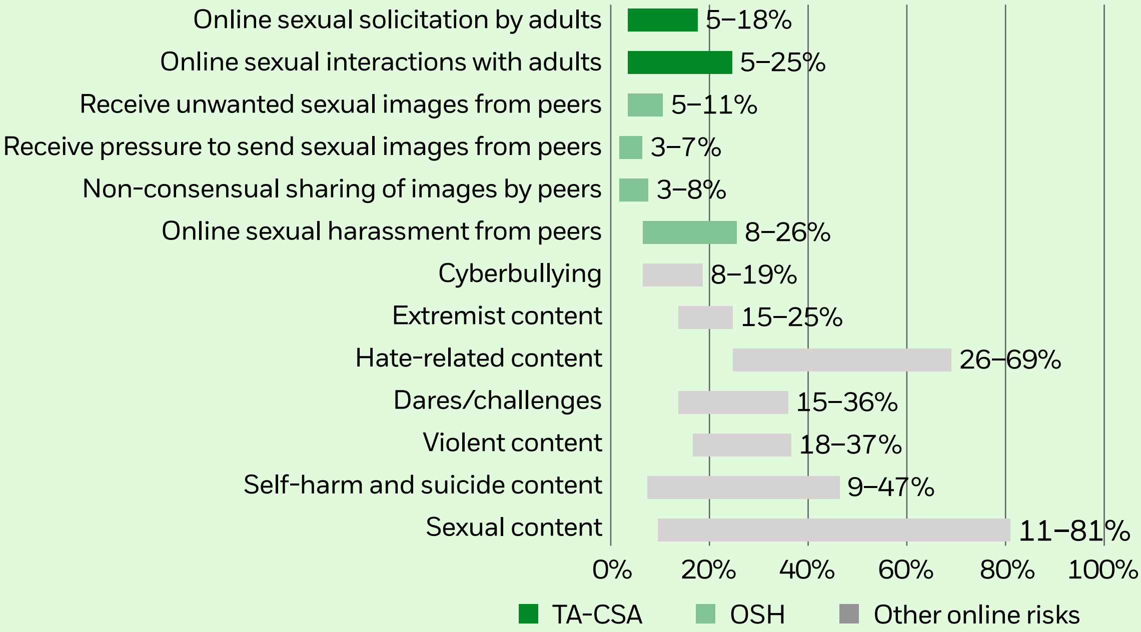
The chapter provides an outline of the prevalence and outcomes associated with seven categories of online risk. It does not attempt to provide a detailed examination of each of these categories, but instead summarises the results from a 'review of reviews', with some additional searches for recently published studies. Reference is made, where appropriate, to the results of a Rapid Evidence Assessment (REA) recently published by NatCen (Hudson et al, 2022) that examined the evidence on this topic in detail. It should be noted that the NatCen REA covered a longer time period (2011 to 2022) and includes prevalence figures from studies that are earlier than those examined in the present review.

Chapter summary

Six types of content risk were examined in this chapter, and one type of conduct risk (cyberbullying).

- **Responses**: The small amount of evidence on children's responses to being exposed to 'harmful' content suggests that children may tell someone about their experience or block a sender; they are less likely to report the problem to a platform, either because they do not know about this function or feel it would be pointless to do so.
- **Scale**: There are various prevalence figures for children's encounters with these risks. Prevalence figures from different studies are not directly comparable to one another, not least because of methodological factors (different studies, for example, asked children to think of different reference periods when giving their answers). Nevertheless, evidence from different sources consistently suggests that a minimum of 1 in 12 children have experienced each of these risks when they were online.
- **Scale**: Children are more likely to be exposed to content risks than most of the online sexual risks examined in previous chapters.

Prevalence of children's exposure to sexual online risks and a range of other online risks



- **Outcomes and impacts**: There is a lack of research examining how exposure to, or engagement with, the different categories of online risk examined in this chapter affects children (with the notable exception of cyberbullying). The available evidence suggests there are some commonalities with the emotional and psychological impacts identified for online sexual risk categories; in addition, there may be specific attitudinal and behavioural impacts and potential health-related outcomes.

6.1 Content risks

Children can be exposed to a variety of content risks online. The 'primary priority' and 'priority' content covered by the review includes sexual and violent content, depictions of self-harm, suicide and eating disorders, as well as online challenges and dares. Illegal content that promotes extremism or radicalisation is also covered.

Methodological factors in measuring exposure to these types of content

Most of the studies examining these categories of content are based on self-report data. Self-report data is limited by a reliance on the participant's own perception of what constitutes content of that type – for example, their assessment of whether an image is sexual or violent. There are also potential problems with recall, and whether participants can reliably attribute their experiences to a particular point in time or reference period.

The subsections that follow show a variety of prevalence figures associated with each category of content. In part, this variation results from different ways of measuring exposure among children. For example, studies may use different definitions for a particular type of content; they may focus on different age groups (e.g., the experiences of younger children only, or of adolescents, or of a broad age group); or they may use different timeframes e.g., asking research participants to recount experiences in the past month, the past year, or during their entire lifetime.

It is also important to recognise that the precise nature or severity of the content that children are exposed to is rarely captured in these studies. Studies do not generally distinguish whether a piece of content was promoting, glamourising, inciting others or providing instructions for the harm, or whether it was simply portraying it.

As explained previously in Section 1.2, the fact that these types of content are often referred to as 'harmful' does not automatically mean that a user who encounters them would be harmed. A variety of factors play a part in children's outcomes, including the precise nature of the content, a child's personality traits, and social and environmental factors in children's lives. It is beyond the scope of this report to examine these factors or the complex interactions between them, and whether they lead to harm.

Responses to encountering these types of content

There is a general lack of evidence examining children's responses to 'harmful' content online. The main sources of data in the UK instead focus on measuring frequency of exposure.

The Ofcom media use and attitudes report (Ofcom, 2023) found that among 8–17-year-olds who had seen something worrying or nasty online, 84 per cent told someone about it. This varied by age, with younger children more likely to do so. The study also found that 66 per cent of 12–17-year-olds had blocked someone on social media, and 33 per cent had changed their privacy settings to reduce the number of people who could see their profile. Only 14 per cent had used a reporting or flagging function to report inappropriate content. There were lower awareness levels for this function (35 per cent) compared with blocking people on social media (84 per cent), which might explain why it is less frequently used. This suggests that increasing knowledge about reporting tools should receive greater focus in educational resources.

These results are consistent with another study (Children's Commissioner, 2022), which found that 50 per cent of children aged 8–17 in a representative UK sample who had seen harmful content online did not report this to the platform. A large proportion of these participants (40 per cent) said this was because they did not feel there was any point, and the platform would not take any action.

Elsewhere, children have expressed concerns about the repercussions of telling their parents or carers about these encounters for fear of their internet usage being limited or monitored by adults, either as a safety measure or as punishment (Nominet, 2022).

6.1.1 Sexual content

Definitions

This category of harmful content relates to children being exposed to sexual content online intentionally or by accident. It refers to a variety of types of sexually explicit material that can be accessed on different online platforms (Hudson et al, 2022; Stoilova et al, 2021), but excludes that which is consensually or non-consensually produced (see Chapter 3 for more detail).

Prevalence

There are a broad range of prevalence figures for exposure to sexual content online, with the available evidence suggesting that 11–81 per cent of UK children have ever had this experience. For example:

- The recent review conducted by NatCen (Hudson et al, 2022) reported a range of prevalence figures for children being exposed to this type of content online (11–81 per cent across studies reviewed).
- Katz et al (2023) found that 32 per cent of a sample of 11–16-year-olds in the UK had ever seen unwanted sexual content.
- A recent study found that 64 per cent of a representative sample of 16–21-year-olds in the UK had seen pornography online, with 50 per cent of these participants having seen this by age 13 (Children's Commissioner, 2023).
- Ofcom found that 9 per cent of children aged 13–17 had seen inappropriate sexual content *in the last 4 weeks* (Ofcom, 2022b).

Research suggests that older adolescents are more likely to see sexual content online. In a representative study of UK children, 79 per cent of 16–17-year-olds had seen sexual content online, while only 51 per cent of 11–13-year-olds had done the same (BBFC, 2020). There is also evidence that boys are more likely to report having seen sexual content than girls (Hudson et al, 2022; Stoilova et al, 2021). For example, Martellozzo et al (2020) found that 56 per cent of boys aged 11–16 had seen this type of content online compared with 40 per cent of girls of the same age.

Outcomes

A variety of emotional, attitudinal, and behavioural outcomes have been identified as being associated with exposure to sexual content online. Emotional reactions include shock, fear, anxiety, sexual arousal, and curiosity (Hudson et al, 2022; Martellozzo et al, 2020). Exposure has been found to potentially impact on children's attitudes towards gender stereotypes and norms about appropriate behaviour in romantic relationships and sexual interactions e.g., use of coercion and violence (Barter et al, 2022). These attitudinal impacts may also influence sexual behaviour and encourage use of sexual coercion, sexual violence, and harassment. For example, Barter et al's (2022) study, which included UK participants, found that frequent exposure to online pornography among boys was associated with perpetration of emotional and sexual abuse in their relationships; however, causality could not be inferred due to the cross-sectional nature of the study.

6.1.2 Self-harm, suicide and eating disorder content

Definitions

The literature does not provide a clear definition of this category of online risk and harm. Instead, it provides descriptions of the types of self-harm, suicide and eating disorder content available online and the user interactions that occur around it (Hudson et al, 2022). These forms of content promote the associated behaviours and provide information about methods of harm. They can be shared by individuals on their feeds, shared in online communities showing supportive attitudes to suicide, self-harm and eating disorders, or be presented to users through content recommendation algorithms and/or the social media feeds of others (Stoilova et al, 2021). A number of platforms have been identified as hosting these types of harmful content, including Twitter, Roblox, TikTok and Tumblr (CCDH, 2022; Jacob et al, 2017).

Prevalence

There is a general lack of UK-based research examining the prevalence of children's exposure to this category of online risk and harm (Hudson et al, 2022). The available evidence suggests that 9–47 per cent of children have ever seen this type of content online. For example:

- The NatCen review reported a range of prevalence figures for children being exposed to content of this type (9–25 per cent across studies reviewed) (Hudson et al, 2022).
- Katz et al (2023) found that 45 per cent of children aged 11–17 in the UK had ever seen suicide-related content, 47 per cent had seen eating disorder content and 25 per cent pro-self-harm content.
- Ofcom found that 7 per cent of children aged 13–17 had seen suicide and self-harm content *in the previous four weeks* and 11 per cent had seen eating disorder content (Ofcom, 2022b).

> "I can't remember how I found it, but you type in cutting and then everything comes up, all links, most of the things that come up are linked to Tumblr. So you type in self-harm ... And then people share like pictures of self-harm, which is normal, it's Tumblr, that's what people use it for really."
>
> Boy, aged 16 (Jacob et al, 2017)

Literature reviews of the available evidence suggest that girls, non-binary and trans children are more likely than boys to view or share self-harm content (Kostyrka-Allchorne et al, 2023; Stoilova et al, 2021).

Outcomes

The negative impacts of exposure to this type of content include triggering, distress, reinforcement and normalisation of 'pro' attitudes, suicidal ideation and negative body image, as well as increased engagement in the related behaviours (Hudson et al, 2022; Kostyrka-Allchorne et al, 2023; Mento et al, 2021; NSPCC, 2022c).

These effects may be further exacerbated where children are members of online communities that are supportive of these behaviours, as the social interaction they provide can further normalise and encourage them (Jacob et al, 2017). This is consistent with research in which children described the role of negative online experiences (in this area and others) in the development of their mental health problems (Livingstone et al, 2022).

6.1.3 Violent content

Definitions

There is no general definition of violent content in the literature (Hudson et al, 2022). This broad category encompasses a range of behaviours and content types, including violent content in gaming environments, violent clips on YouTube or user feeds, gore sites, and content glamorising gang culture and knives. It can also include hate speech (covered in section 6.1.1) and content related to extremist beliefs and radicalisation e.g., Far-right, Incel, jihadist content (covered in section 6.1.6).

Prevalence

The evidence base for prevalence of exposure to violent content in the UK is underdeveloped. There are specific literatures that cover the categories mentioned above, but it is beyond the scope of the review to examine the specific evidence for each here. Studies that ask about exposure to violent material in general terms suggest that 18–37 per cent of children have ever viewed it:

- The review conducted by NatCen reported a range of prevalence figures for children being exposed to violent content online (18–30 per cent across studies reviewed) (Hudson et al, 2022).
- Katz et al (2023) found that 37 per cent of children aged 11–17 had ever seen violent content online.
- Ofcom found that 16 per cent of children aged 13–17 had been exposed to this type of content online *in the last 4 weeks* (Ofcom, 2022b).

Outcomes

Research about the impacts of exposure to violent content online can be found in different literatures (e.g., violent gaming, hate crime, radicalisation) and does not examine exposure in general terms (Hudson et al, 2022). This material has the potential to reinforce the acceptability of aggression and encourage violent behaviour in some children, depending on the influence of other personality, social and environment factors. For example, there is emerging evidence that exposure to gang-related violent content on social media, as well as depictions of knives or recordings of fights, may be a catalyst in serious offline violence (Irwin-Rogers & Pinkney, 2017; Patton et al, 2019). There is also a developed but contested body of research examining the effects of playing violent video games, although many of the studies focus on impacts on adult samples.

6.1.4 Challenges/Dares

Definitions

It is difficult to define content that relates to dangerous stunts and challenges as this covers a range of different behaviours that carry different levels of risk (Bada & Clayton, 2020; Hudson et al, 2022). Hilton et al (2021) define online challenges as situations where children film themselves engaging in some form of risky behaviour and then share it online to encourage others to do the same.

Prevalence

The available evidence base for the UK suggests that a relatively large proportion of children have ever seen such content online (15–36 per cent). For example:

- The review conducted by NatCen reported a range of prevalence figures for children being exposed to content promoting challenges/dares (15–21 per cent across studies reviewed) (Hudson et al, 2022).
- Katz et al (2023) found that 36 per cent of children aged 11–17 in the UK had ever seen challenges/dares online.
- Ofcom found that 18 per cent of children aged 13–17 had seen this type of content online *in the last 4 weeks* (Ofcom, 2022b).

Outcomes

There is very little evidence related to the outcomes of exposure to this form of content, but there is a risk of physical injury as a result of repeating the viewed activities. For example, Hilton et al (2021) found that 21 per cent of a sample of 13–19-year-olds had engaged in an online challenge, with 2 per cent reporting that it was risky or dangerous.

6.1.5 Hate-related content

Definitions

There are a variety of definitions used in relation to this form of online risk and harm (Brown, 2017; Wachs et al, 2022). 'Hate speech' has been defined as harmful communications motivated by the desire to justify or disseminate hatred against specific social identities or groups (Kansok-Dusche et al, 2022; Wachs et al, 2022). This is reflected in the UK definition of hate crime as:

> *"Any criminal offence which is perceived by the victim or any other person to be motivated by hostility or prejudice, based on a person's disability or perceived disability; race or perceived race; or religion or perceived religion; or sexual orientation or perceived sexual orientation or transgender identity or perceived transgender identity."* (Home Office, 2022)

Prevalence

While the NatCen review did not provide any prevalence figures for this content category (Hudson et al, 2022), other evidence suggests that exposure to hate-related content and communications online is relatively common: 26–69 per cent of children have ever been exposed to this type of material, without necessarily being targets themselves. For example:

- Almost a third of 11–16-year-olds (31 per cent) reported ever having seen content which encouraged racist views in the most recent UK Cybersurvey (Katz et al, 2023).
- A recent systematic review found a wide range of prevalence rates for exposure (26–69 per cent), victimisation (7–24 per cent) and perpetration (5–32 per cent), with males more likely to see and engage with such content online (Kansok-Dusche et al, 2022).
- Ofcom (2022b) found that 13 per cent of children aged 13–17 had been exposed to hateful or offensive[23] content online *in the last 4 weeks*.

Outcomes

The impacts of exposure to hate-related content on children have not been widely studied. However, developmental factors could mean that children are more receptive than adults to experiencing emotional reactions to this type of content (e.g., fear, distress), and may be more susceptible to attitudinal impacts (e.g., stereotyping of social groups, prejudice, radicalisation) or behavioural impacts like violent or aggressive online and offline behaviour (DRCF, 2022; Kansok-Dusche et al, 2022).

[23] This is defined in the questionnaire as 'hateful, offensive or discriminatory content that targets a group or person based on specific characteristics like race, religion, disability, sexuality or gender identity' (Ofcom, 2022b).

6.1.6 Extremist content and radicalisation

Definitions

There is a general lack of evidence in the UK related to the exposure of children to extremist material and radicalisation online. Although the literature on radicalisation and the role of the online environment has developed in recent years, this generally focuses on adult samples. Extremist material is defined as online content that challenges democratic values, communicates stereotypes of specific social groups and blames them for social and political conditions, as well as inciting violence (Nienierza et al, 2021).

Prevalence

The NatCen review did not provide prevalence figures for this category of harmful content (Hudson et al, 2022), but recent research suggests that 15–25 per cent of children in the UK have ever seen extremist material online. For example:

- Katz et al (2023) found that 25 per cent of a sample of 11–16-year-olds in the UK had ever seen content supporting extremism or terrorism in the previous year.
- Research with German adolescents found that 15 per cent of a sample of 14–19-year-olds had seen extremist content frequently or very frequently (Nienierza et al, 2021). This occurred most commonly on social media platforms (26 per cent), video-sharing platforms and messaging apps (15 per cent).
- Ofcom found that 2 per cent of children aged 13–17 reported seeing this type of content *in the last 4 weeks* (Ofcom, 2022b).

Outcomes and impacts

There are very few studies examining the outcomes of children's engagement with this category of online content. It has been argued that children may be particularly vulnerable to the impacts of this material due to their developing identity and political beliefs (Nienierza et al, 2021). Trends in radicalisation referrals through Prevent[24] and children's social care show an increase in reports related to extreme right-wing ideologies and ideologies that are mixed, unclear, or unstable; as well as an increase in reports linked to the role of online influencers (Langdon-Shreeve et al, 2021). The same study highlights the importance of understanding the range of vulnerabilities (e.g., autism spectrum disorder, social isolation, low self-esteem) that may be involved in children engaging with extremist ideologies and radicalisation. There is currently a lack of empirical evidence assessing the influence of these factors on the outcomes of exposure, and how the prevalence and the impacts of engagement with extremist content relate to radicalisation and hate crime (Stoilova et al, 2021).

[24] 'Prevent' is a national programme in the UK that aims to stop people from becoming terrorists or supporting terrorism.

6.2 Conduct risks

Conduct risks refer to children witnessing, experiencing or engaging in potentially harmful peer-related contact (Livingstone & Stoilova, 2021). Sexual conduct risks are covered in detail in Chapters 2 to 5, and the relevant evidence is, therefore, not repeated here.

Research has identified a relationship between being a victim of online sexual harassment (OSH) or intimate image abuse (IIA) by peers and being a victim of cyberbullying (Gámez-Guadix et al, 2022). There is also evidence that OSH, IIA and cyberbullying have a similar pattern for perpetration, as there is some overlap in the behaviours involved (e.g., threats, name-calling); the impacts of OSH and IIA can be more severe due to the involvement of sexual images and the ways in which victims are judged by others (Gámez-Guadix et al, 2022).

6.2.1 Cyberbullying

Definitions

Cyberbullying is defined in the literature as an: '(a) intentional aggressive behaviour that is, (b) carried out repeatedly, (c) occurs between a perpetrator and victim who are unequal in power, and (d) occurs through electronic technologies' (Kowalski et al, 2014, p.1,073). It involves a range of different behaviours (e.g., direct threats, gossip, exclusion from online groups) that can occur across a variety of platforms (e.g., social media, video-sharing platforms, gaming sites) (Stoilova et al, 2021).

Prevalence

The available evidence suggests that prevalence of cyberbullying is similar to that of OSH – reflecting the overlap in behaviours – but that cyberbullying occurs more frequently in the population than IIA. The literature on cyberbullying in the UK shows prevalence figures of 8–19 per cent. For example:

- The NatCen review reported a prevalence range of 8–19 per cent across studies reviewed (Hudson et al, 2022).
- Mateu et al (2020) found that 17 per cent of their UK sample of 11–19-year-olds had been victimised, and 29 per cent had been perpetrators of the behaviour.
- Katz et al (2023) found that 17 per cent of children aged 11–17 had ever experienced cyberbullying.
- Ofcom found that 13 per cent of children aged 13–17 had experienced bullying, abusive behaviour or threats online *in the last 4 weeks* (Ofcom, 2022b).[25]

There is evidence that this form of victimisation is more prevalent in girls and older adolescents (Brooks et al, 2020; Hudson et al, 2022; Smith et al, 2019).

Outcomes

The experience of cyberbullying can have a variety of negative psychological and social impacts, including distress, anxiety, and depression (Stoilova et al, 2021). Recent UK studies have found evidence that victimisation is associated with symptoms of post-traumatic stress, self-harm and suicidal ideation (John et al, 2023; Mateu et al, 2020). However, researchers have highlighted the predominance of cross-sectional designs in this area, which reduces the ability to establish causality and highlights the need for longitudinal studies (John et al, 2023). Notably, these outcomes are similar to those identified for OSH and IIA victimisation.

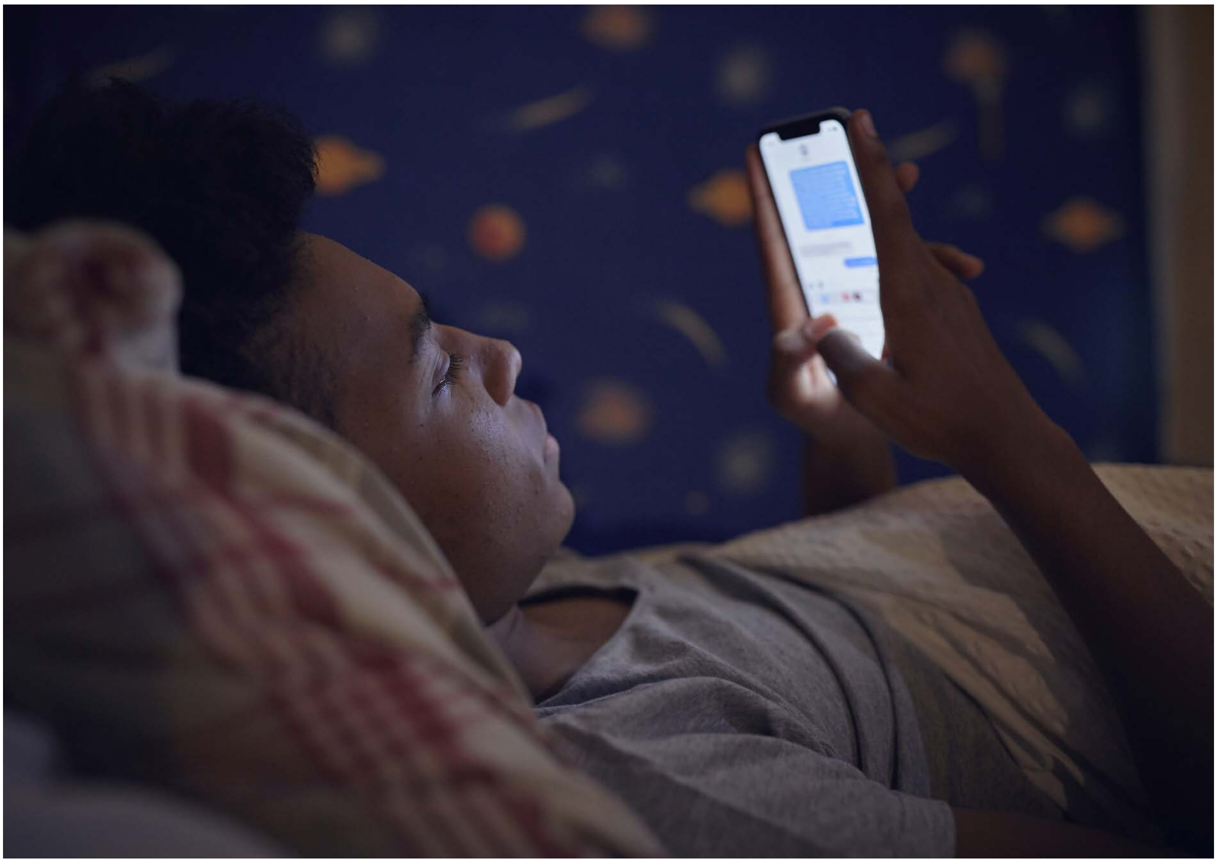
[25] Other research, for example by Ofcom (2023), has found a higher proportion of children reporting experiences of someone being nasty or hurtful to them through communication technology. However, this cannot be classified as 'cyberbullying' as there is no indication that this behaviour was repeated.

Chapter 7
Technological Design Features and Tools that Increase and Decrease Risk of Harm to Children

The previous chapters examined evidence relating to the different categories of risk and harm to which children are exposed online. This chapter focuses on the ways in which technological design features increase or decrease risk and harm. By doing so, it demonstrates that risk is not an inevitable outcome of being online: it is also influenced by the design choices of platforms and the safety tools implemented.

The tools and features that form the focus of this chapter are those most commonly discussed in the literature around the safety and protection of children online. The chapter provides a descriptive account for readers without technological expertise and presents the rationale for why the feature or tool may enhance or reduce risk. There is no attempt to quantify the effect of the various features and tools on children's likelihood of encountering risk.

Section 7.1 examines the ways in which platforms can potentially increase risk of harm. Section 7.2 examines the different technical tools available to decrease risk and protect children, and also presents some of the challenges associated with their use.



Chapter summary

The table below provides an overview of the different design features and tools examined in this chapter that can increase or decrease online risk and harm to children.

Table 3: Features and tools that can increase or decrease online risk and harm

	Exposure to inappropriate content	**Exposure to potential perpetrators**	**Exposure to inappropriate content**	**Exposure to potential perpetrators**
	What can increase this?		**What can decrease this?**	
Algorithmic recommendations	Inappropriate recommendations or search results; recurring recommendations that amplify harmful messages.	Friend, follower, or random contact recommendations.	Functionality and features designed using 'safety by design' principles and assessed for potential risks before rollout and during implementation.	
Notification and quantification of activity	Notifications, feedback requests and updates that increase activity.	Popularity metrics that increase networking.		
Moderation	Ineffective or minimal moderation of content.	Ineffective or minimal moderation of user-to-user interactions.	Effective and consistent moderation, leading to rapid removal and/or disruption of harm.	
Detection	Ineffective or minimal detection of illegal content, such as CSAM (e.g., in end-to-end encrypted environments).	Ineffective or minimal detection of illegal interactions, such as grooming (e.g., in end-to-end encrypted environments).	Effective and consistently implemented detection tools, leading to rapid removal and/or disruption of harm.	
Age assurance	Ineffective or absent age-assurance measures.		Effective and consistently implemented age-assurance measures, leading to age-appropriate online experiences.	
Parental controls	Ineffective or difficult to use parental controls.		Effective and easily implemented parental controls.	
User reporting and safety tools	Ineffective or absent reporting and safety tools.		Comprehensive and easy to use reporting and safety tools, leading to rapid and effective removal and disruption of harm.	

7.1 Features and tools that can increase online risk and harm

Fundamentally, the profitability of platforms relies on high levels of usage: the more time users spend on a platform interacting and generating content, the higher the platform's revenue (from advertising, for example). To achieve this, technology companies deliberately design their platforms to maximise the amount of time users spend engaging with content, the size of their social network, as well as the amount of content they create and share. The design choices involved to achieve the platforms' business goals are referred to as their 'choice architecture' (5Rights Foundation, 2021b; Ofcom, 2022a).

Choice architectures form specific usage habits. The content presented to users, along with features that demonstrate popularity and social approval (e.g., 'likes', 'thumbs up', 'heart' emojis), can incentivise behaviour, reinforce habits, and maximise engagement with the platform. However, they do so by exploiting the interests and developmentally important needs of children (e.g., pursuit of popularity and social approval, concerns about body image). In doing this, choice architectures may also influence children's exposure to risk and harm (5Rights Foundation, 2021b, 2023; Ofcom, 2022a).

It has been claimed that choice architectures are inherently 'risky by design' (5Rights Foundation, 2021b, 2023; Livingstone et al, 2022). That is not to say that experiences of risk and harm are contingent on this factor alone as psychological, social, and environmental factors also play a part. However, greater recognition of the role of choice architectures can give a more comprehensive understanding of how risk and harm come about. The features examined below can unintentionally expose children to both content and contact risks.

7.1.1 Algorithmically recommended content

Some platforms use algorithms to automatically present users with content that is targeted to appeal to them. This is based on analysing user behaviour and interests from their previous views and 'likes' (5Rights Foundation, 2021b). This removes the need to actively search for content and, combined with the use of automatic play functions for video content, encourages users to spend long periods of time on platforms (5Rights Foundation, 2021b). Similarly, users can be presented with a continual stream of personalised content when they scroll down their screen (referred to as their 'feed'), without the need for active searching (5Rights Foundation, 2023; Ofcom, 2022a). These features encourage users to spend more time on the platform without alerting them if this becomes excessive. Some platforms do have tools for monitoring or setting time limits, though these usually need to be activated by users and are not commonly used by children (Ofcom, 2022a).[26]

The amount of time spent on platforms in combination with content recommender algorithms can increase risks to children by leading to prolonged and narrowing exposure to harmful content (Livingstone et al, 2022; Ofcom, 2022a). The presentation of content that is triggering or supportive of self-harm, suicide, and eating disorders, as well as violent or sexual material, can negatively impact mood, attitudes, behaviour and mental health (DRCF, 2022; Livingstone et al, 2022). This is particularly problematic for children whose circumstances, mental health and other vulnerabilities make them disproportionately susceptible to the effects of specific types of harmful content (Livingstone et al, 2022).

[26] Exceptions do exist, for example TikTok has recently introduced a default daily screen time limit of 60 minutes per day for users aged under 18 (BBC, 2023).

Recent research testing the type of content recommended by algorithms using simulated accounts for child users: this found that potentially harmful content was automatically targeted at users who self-declared as children (CCDH, 2022; 5Rights Foundation, 2021b). These studies also found that this was particularly the case for child accounts with indicators of vulnerability (e.g., eating disorder terms in their usernames), and that harmful content could easily be searched for and accessed by users of this age. The studies suggest that although children may actively seek out harmful content, they can also be automatically exposed to these online risks by algorithms that use profile keywords to recommend content (Hudson et al, 2022; Kostyrka-Allchorne et al, 2023).

7.1.2 Friend recommendation algorithms

Friend recommendation algorithms use data related to users' common connections, their location, and their personal interests to suggest new contacts (5Rights Foundation, 2021b; Ofcom, 2022a). By making it easy to find and connect with people, users are encouraged to develop large social networks, which in turn encourages engagement and an increasing amount of activity on user feeds. In addition, some platforms allow other users to add children as friends without them actively accepting contacts, as long as they have public profiles and a common connection; or they allow other users to send direct messages to children (Ofcom, 2022a). These features can be disabled, but this requires action and the necessary know-how by the user.[27]

As a result, children may have many unknown followers or contacts on their friends lists – some of whom were not added through active choice (5Rights Foundation, 2021b) – who could pose a risk to them by having access to their profiles and posts (Ofcom, 2022a). Children may be targeted by these unknown contacts with unwanted requests for sexual interactions and images, coercion, manipulation or grooming. This can also become a route for access to other forms of inappropriate content and contact that can have negative psychological, attitudinal and behavioural impacts (e.g., pro-suicide and self-harm material).

7.1.3 Quantification of social activity and popularity

Some platforms display metrics that indicate how popular users, or their posts, are (e.g., by showing number of 'likes' or 'followers'). The ability to demonstrate one's own popularity, and the ease with which users can give and receive feedback via single click responses, can encourage users to continuously find or create content to post (5Rights Foundation, 2021b). Children are especially susceptible to this as they place particular importance on social acceptance and popularity and may feel driven to increase their contacts or followers by making contact with recommended users, or by accepting friend requests from unknown others (Ofcom, 2022a). The use of notifications to update users about activity (e.g., 'likes', new posts) can also promote more interaction and engagement by encouraging users to keep up with newly posted content. Again, children who may feel anxious about missing out on social activities ('FOMO' – fear of missing out) may be particularly susceptible to this. Altogether, these features serve to boost online interaction and create self-reinforcing usage patterns (5Rights Foundation, 2021b).

[27] Many platforms have default private profiles for users under 18, but research shows that children often change them to 'public' as a means of expanding their social network and gaining further social approval (Ofcom, 2022a).

These features can increase children's potential exposure to harmful content and contact by encouraging them to make new (and sometimes indiscriminate) contacts and to spend long periods of time on platforms.

7.2 Features and tools that can decrease online risk and harm

This section examines features and safety tools that platforms can use to reduce children's exposure to online risk and harm. It should be noted that the application of the tools described below runs counter to platforms' fundamental business model: implementing them creates friction and barriers to the user experience, which then impacts on users' levels of interaction and engagement (i.e., usage habits central to their business model). For this reason, platforms may be reluctant to implement them (5Rights Foundation, 2021b).

7.2.1 Age assurance

Age assurance refers to a range of techniques used to verify the age of users when they register an account and use platforms, or when they access online content that is age-restricted or age inappropriate (e.g., adult pornography, graphic content, gambling) (5Rights Foundation, 2021a; DCMS, 2020). There are a variety of different approaches and tools available for this, although these have varying levels of reliability in determining the age of users (5Rights Foundation, 2021a; Smirnova et al, 2021). Approaches that combine human, technical and hard identification have been highlighted as most likely to be successful in determining age and preventing circumvention by children (Smirnova et al, 2021).

Age assurance can potentially reduce children's exposure to online risks. However, there are several challenges to safeguarding children by this means:

- **Limited adoption:** the available literature suggests that the current use of age-assurance tools is fragmented, and there is a reluctance among platforms to implement them as they may result in additional responsibilities in relation to the safety of children (5Rights Foundation, 2021a).
- **Potential for circumvention:** the most commonly used approach to verifying age – namely, self-declaration of age at sign-up – is relatively ineffective. There are limited checks to verify the information provided, allowing children to enter a higher age (Ofcom, 2022d) without understanding how this may influence their online experiences and potential exposure to inappropriate content and contact (5Rights Foundation, 2021a; Smirnova et al, 2021).

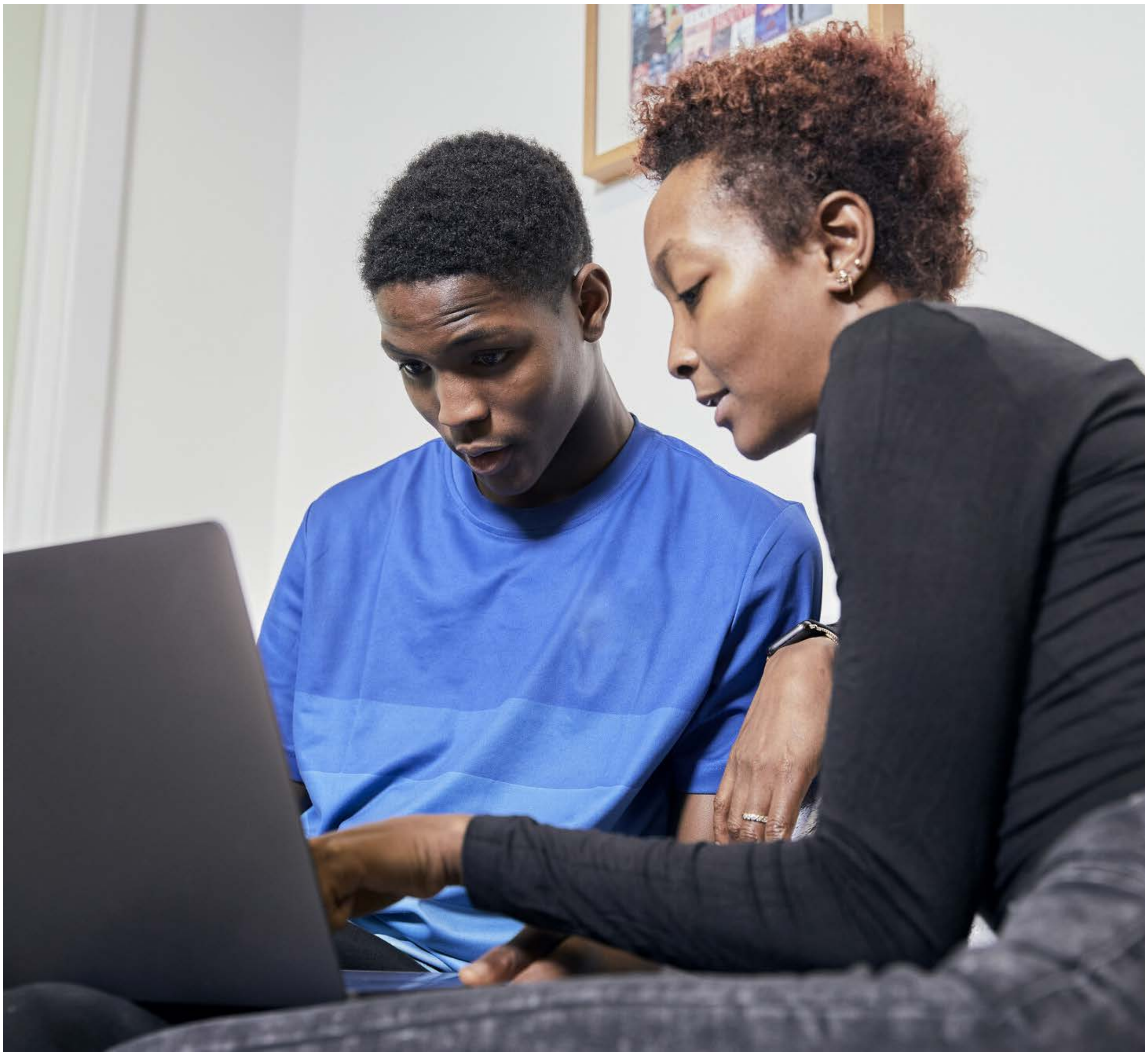
Widespread implementation of robust age assurance would potentially result in greater reductions in children's exposure to online risks. A key recommendation of recent reports examining age-assurance approaches is the development of a set of common standards in the UK with specific performance metrics that can be used to evaluate their effectiveness (5Rights Foundation, 2021a; Smirnova et al, 2021). Alternatively, or alongside this, platforms could modify design features that are particularly risky for children or redesign their services to ensure they are suitable for audiences of mixed ages (5Rights Foundation, 2021a).

7.2.2 Parental controls

There are a variety of different technical tools that parents and carers can use to improve the safety of their children online: these can set screen time limits, block and filter inappropriate content, and monitor online behaviour and interactions (Smirnova et al, 2021). The tools are offered by broadband and mobile providers, and platforms and search engines; they can also be purchased and installed on home networks and on children's devices.

Parental tools can play an important role in reducing children's exposure to online risk and harm, but their efficacy can be limited by a variety of factors (Smirnova et al, 2021):

- **Ineffective design or poor usability:** Tools must be well designed, effective and easy for parents and carers to use, even if they have differing levels of skills or knowledge. Parents need to be aware of the availability of tools, their function and limitations, as well as how to set them up (Smirnova et al, 2021).
- **Insufficient tailoring to children's development:** Tools must be age-appropriate, customisable and be able to respond to developmental changes in children's knowledge and capacities over time (Smirnova et al, 2021).



- **Ineffectiveness:** Przybylski & Nash (2018) highlighted the limited protection provided by filtering tools, which their study found to under-block inappropriate sites or content, over-block those that were age appropriate, and that proved ineffective in preventing children's exposure to online sexual content.
- **Poor levels of uptake:** The most recent data for the use of parental control tools in the UK found that there is a relatively low level of use (12 per cent), with parents and carers more likely to talk to their children about their online behaviour (39 per cent) (Ofcom, 2023). Tools that are built into platforms or offered by broadband providers are more likely to be used than those that are purchased and added to home networks or devices independently (Smirnova et al, 2021). Smahel et al (2020) also found a relatively low uptake of tools (22 per cent across 19 countries), with active parental mediation being more commonly used. This may be due to a lack of awareness of the availability of tools and knowledge to set them up. However, the evidence suggests that this reflects a preference for other forms of mediation (e.g., discussion, rule-setting).
- **Negative perceptions of children:** Children may feel that the use of parental controls invades their privacy, creating conflict and the motivation to find workarounds that can inadvertently increase their exposure to risk and subsequent harm (Smirnova et al, 2021).
- **Not complemented by discussions about online safety:** The use of parental controls can create a false sense of security for parents and carers, replacing the perceived need for discussions with their children about online safety (Smirnova et al, 2021).

The use of parental tools alongside a range of mediation strategies would provide a more effective approach to reducing online risk and encourage children's development of resilience to online risk and harm (Smirnova et al, 2021). The ongoing development of tools is also important, particularly in relation to their interoperability across different platforms.

7.2.3 Content moderation

Content moderation refers to a range of activities concerned with removing or reducing the visibility of potentially harmful content online. It is used by platforms to assess content and take appropriate action when illegal and harmful material is identified (Cambridge Consultants, 2019; Gillespie, 2022). Platforms use different content moderation approaches depending on their architecture, content formats and the types of user interactions they support (Cambridge Consultants, 2019). Some moderation is automated, relying on Artificial Intelligence (AI) and machine learning, perceptual hashing, metadata, and text-based analysis (Cambridge Consultants, 2019; Draper, 2022). The degree to which human moderators are involved (e.g., community volunteers, internal analysts) varies between platforms (Spence et al, 2023).

By removing potentially harmful content from view, content moderation can theoretically reduce the risk to children. The actual efficacy of AI moderation is hard to assess, however, given the lack of transparency in how such models works, including the training datasets, feature detection and learning models they rely on (Cambridge Consultants, 2019).

The identified limitations of content moderation include:

- **Lack of capacity for handling vast volumes of material:** While AI can moderate content more quickly than humans, the sheer volume of user-generated content that needs to be pre-moderated makes it difficult to resource. The variety of formats involved, as well as the need for metadata analysis to identify user accounts posting illegal or harmful content or interactions, adds to these resource requirements (Cambridge Consultants, 2019; Gillespie, 2022).
- **Limited contextual understanding by AI:** Automated tools need to be trained to identify illegal and harmful categories of content or communication. Identification may require image analysis or linguistic examination but is also likely to require an understanding of context (Cambridge Consultants, 2019; Gillespie, 2022). AI is less competent than humans in understanding context, for example how language and associated symbols are used (e.g., slang, sarcasm, emojis), as well as how they develop over time (Cambridge Consultants, 2019).

Implementation of more sophisticated AI tools with greater capacity for understanding context and moderating vast volumes of material could potentially further reduce children's exposure to online risk. In the meantime, human moderation will still be required for many categories of illegal and harmful content and to verify decisions made by AI moderation (Cambridge Consultants, 2019).

7.2.4 Detection tools

A key task of automated content moderation tools is to scan images, videos and text to detect online sexual abuse, such as CSAM and TA-CSA (Draper, 2022; Cambridge Consultants, 2019). Platforms may implement a variety of detection tools, as described below.

Detection of known CSAM

Hashing is a way of detecting content that has previously been identified as illegal. During content moderation, online images are scanned and compared against databases of 'hashes' (unique digital fingerprints associated with known CSAM images or videos): where the image matches to a hash, the image is recognised as illegal and prevented from being uploaded or shared, and action can be taken on the user account (Cambridge Consultants, 2019; Draper, 2022). A variety of different tools may be used in this process. For example, PhotoDNA (Microsoft) and PDQ (Meta) are used in scanning and detection of still images of CSAM, while TMK+PDQF (Meta) is used for videos (Ofcom, 2022c).

By removing these images from circulation, the risk of sexual victimisation or revictimisation of children is reduced. Hashing solutions are highly accurate; however, they can only detect illicit images that are already known (Cambridge Consultants, 2019). The lack of compatibility of hash databases also means that it can be difficult to detect content featuring specific victims across platforms.

Detection of new or previously undetected CSAM

CSAM that has not previously been identified as illegal – known as 'first generation' CSAM – can be detected using a variety of image classifiers and tools (e.g., Thorn's CSAM Image Classifier). AI can flag new content that is very similar to patterns of previously confirmed CSAM; once flagged, the content is reviewed by human moderators and if it is confirmed as CSAM it is prevented from being uploaded and shared (Draper, 2022). The new image can then be reported to organisations like NCMEC or IWF so that it can be hashed and added to relevant hash databases. Tools have also been developed to detect the live production of new imagery through livestreaming (e.g., SafetoWatch).

The use of these tools can theoretically reduce the upload of sexual images related to IIA and TA-CSA and prevent revictimisation through the continued circulation of CSAM online. However, it is worth noting that these tools are less accurate than hashing. Difficulties in estimating the ages of children in images and assessing the precise nature of what is depicted mean that they can either fail to detect some CSAM or falsely identify benign images as CSAM (WeProtect Global Alliance, 2021).

Detection of online grooming

Online grooming detection tools use AI to analyse different features of online interactions in order to identify patterns of communication that indicate suspected grooming (Draper, 2022). Metadata analysis (e.g., user account details) can also be combined with these approaches, which can serve to identify the adult users involved and take appropriate action. These tools can be integrated with platform architecture, as well as included in parental tools that can be installed on children's devices.

While these tools theoretically disrupt TA-CSA and reduce the victimisation of children, it is not easy to assess their efficacy. This is because they are often proprietary in nature, and platforms provide limited information about their use in transparency reports,[28] or explanation of how they work (e.g., coding, training datasets, classifiers and feature detection) and the learning models they rely on (Borj et al, 2023). Academic researchers have developed and evaluated grooming detection techniques and highlighted a variety of related challenges (Borj et al, 2023; Razi et al, 2021). Most notable are the complexities involved in identifying potential grooming interactions (e.g., sexual talk or requests for images) in real time from a mixture of online material that can include text, emojis, and culturally or locally specific slang. Like other content moderation tools, detection tools can lack sufficient contextual understanding, making it difficult to effectively distinguish grooming activity from other forms of sexual and non-sexual interactions online (Razi et al, 2021).

[28] Transparency reports are published on a quarterly or yearly basis by platforms and provide information about the number of reports or detections of illegal and harmful content in different categories (e.g., CSAM).

Box 3: The impact of end-to-end encryption on detection

The detection technologies described above are not readily deployable within end-to-end encrypted (E2EE) environments. E2EE enables users to exchange messages whose contents can only be read by the sender and recipient (Children's Commissioner, 2020; Kardefelt-Winther et al, 2020). As messages are encrypted during transit, they cannot be read by third parties, whether the platforms themselves, law enforcement, or other agencies (WeProtect Global Alliance, 2021). Platforms like WhatsApp, Telegram, and Signal, which all use E2EE, do not apply these technologies (aside from on non-encrypted data) and consequently detect fewer – or no – cases of CSAM or TA-CSA on their platforms (NCMEC, 2023).

E2EE hinders the detection of instances of sexual abuse, and so increases the potential for children to experience ongoing harm. The decision by some platforms to use E2EE highlights the importance of developing alternative solutions to identify CSAM and TA-CSA, including robust solutions at the device or server levels and encryption-related solutions that give due consideration to users' privacy (Draper, 2022; Peersman et al, 2023).

Plans by Meta to use E2EE across all its messaging applications carry significant implications for children's safety, and has raised concerns from government, law enforcement and NGOs (Children's Commissioner, 2020; NSPCC, 2021; WeProtect Global Alliance, 2021; Home Office 2023b). Wider implementation is predicted to significantly decrease the number of child abuse referrals made by platforms (Europol, 2020; Home Office, 2019a; NSPCC, 2021). It will negatively impact on the ability of law enforcement to obtain communications that can serve as evidence to prosecute perpetrators involved in TA-CSA (Draper, 2022; NSPCC, 2021). It is also likely to increase platform migration by perpetrators from social media to encrypted messaging apps where these tools are not deployed and detection opportunities are reduced (Europol, 2020; NSPCC, 2021). This is particularly problematic given the cross-platform risks involved in TA-CSA (see section 4.2).

7.2.5 Reporting and safety tools

There are a variety of safety tools and features built into platforms, which provide users with the ability to take action if they encounter illegal or harmful content and contact. Functions are available that prevent further content and communication from specific users (block, ignore and mute functions): these are relatively easy to use and access. Tools are also available for flagging content as illegal or inappropriate on video sharing platforms. This content is then reviewed by human moderators and appropriate action taken. In addition, users have the option to report offensive content and other users to platforms via a menu of different options.

Reporting and safety tools can prevent ongoing risk to child users who have already encountered online material or conduct that distresses them. However, there are various limitations associated with them:

- **Poor usability:** Some reporting tools use menus that are difficult to navigate; consist of options that use technical language that is hard to understand; or omit options for some of the risks that children encounter online. As a result, children can struggle to specify the experience they wish to report. A recent study found that it was difficult to locate reporting tools on the platforms examined, and to flag content specifically as CSAM (Canadian Centre for Child Protection, 2021).
- **Mixed uptake:** While children commonly use block and ignore functions, few make use of reporting tools (Thorn, 2022). This may partly be due to uncertainty about whether their report will lead to any action (Children's Commissioner, 2022). Research has identified a lack of information from platforms about the action that will be taken in response to reports, and the amount of time it will take for users to receive a response; moreover, the outcome of reports is often not provided (Canadian Centre for Child Protection, 2021). Transparency reports by individual platforms confirm that only small proportions of the illicit content that they refer to the National Center for Missing and Exploited Children (NCMEC) is derived from reporting or flagging by their users (e.g., Facebook, 2023; Google, 2023).

These platform features provide a helpful route to recourse for children who feel able to proactively respond to online risks. Their effectiveness would improve if they were easy to find and use and framed in child friendly language, and if they offered users feedback about any action taken by the platform.

Chapter 8
Conclusions

The aim of the review was to explore children's exposure to online risk and harm, and how this has changed since the publication of the last comprehensive evidence review on this topic, published in 2017 (Livingstone et al, 2017). This updated picture of the evidence base provides an overview of the online risk landscape for children in the UK ahead of the changes associated with implementation of the Online Safety Act.

This section organises the results of the evidence review to directly address each of the five research questions posed in Section 1.3.

Research Question 1: What are the new developments in the risk landscape that have arisen over the past five years, and are there emerging or longer-term trends?

- A number of technological advances that have occurred since 2017 have had an impact on children's exposure to online risk and harm. This includes the launch of new platforms like TikTok and the prominence of social media influencers. Increases in the popularity of livestreaming, ephemeral media, and messaging platforms, as well as the development of virtual reality technology, have changed the dynamics of online sexual interaction and, consequently, the dynamics of Technology-Assisted Child Sexual Abuse (TA-CSA), unwanted peer-to-peer sexual conduct, and the production and distribution of Child Sexual Abuse Material (CSAM). The COVID-19 pandemic saw changes in children's and adults' online usage, as evidenced in reports of increases in TA-CSA and the amount of CSAM being detected, particularly first-person produced sexual imagery. This has been accompanied by developments in Artificial Intelligence (AI) that have provided additional means for producing synthetic CSAM (e.g., deepfakes), but also greater efficiency in moderating and detecting material that may be harmful to children. There has also been increasing use of algorithms by platforms to 'feed' content and recommend contacts to children, and greater recognition of algorithmic risk as an example of how platform design – or so-called 'choice architecture' – can have a direct impact on children's potential experience of harm.
- It is difficult to determine the extent to which these technological developments have impacted on children's exposure to risk and harm online. This is only one of various different factors that can potentially influence the dynamics of risk and harm. Moreover, there is a lack of robust, longitudinal datasets that can be used to assess trends and changes over time.
- Looking ahead, there is no easy way to predict whether risk will reduce or increase after 2023. The Online Safety Act has initiated the regulation of user-to-user services; this, alongside the increasing use of AI-based tools to improve content moderation and grooming detection, may lead to a reduction in online risk and harm, particularly if other preventative measures are also applied. Conversely, the wider implementation of end-to-end encryption (E2EE) that is set to take place in the coming months, and growing usage of generative AI and virtual reality, may increase harmful online experiences. The combined effect of such developments is unlikely to be straightforward.

Research Question 2: What does the current evidence suggest about children’s exposure to online sexual risk and associated harm?

This 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. Working definitions were presented for four categories of online sexual victimisation. Online sexual harassment (OSH) is an umbrella term for a variety of different unwanted sexual behaviours perpetrated by children against other children. This includes intimate image abuse (IIA), which relates to the non-consensual production and sharing of sexual images. Technology-assisted child sexual abuse (TA-CSA) refers to situations where children are sexually exploited online by adults. Child Sexual Abuse Material (CSAM) refers to imagery of children being sexually exploited or engaging in sexual behaviour and can be produced through peer or adult sexual interactions online or be first-person produced.

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

Scale of online sexual risks

- A sizeable minority of children – at least 1 in 20, but potentially up to a quarter – have encountered sexual risk online.
- Some sexual risks are more likely to be encountered online than others. Drawing on evidence published since 2017, the review found that the most common types of sexual risk were OSH by peers (8–26 per cent prevalence) and TA-CSA (5–18 per cent experienced online sexual solicitation and 5–25 per cent experienced sexual interaction with adults). IIA by peers was relatively less widespread (e.g., 5–11 per cent prevalence for receiving unwanted sexual images).
- It is not possible to confidently establish whether there has been an increase or decrease in online sexual victimisation in the years since 2017. Various data sources indicate upward trends over time in the incidence of TA-CSA and the amount of CSAM being reported or detected by different agencies (including first-person produced images). These trends are likely to be influenced by a variety of factors, such as developments in scanning and detection tools, and not solely reflect an increase in levels of perpetration and victimisation. It is nevertheless worth noting that no data source shows dips or downward trends.

Platforms on which online sexual risks occur

- There are indications that perpetrators commonly move victims from public platforms to private online environments, but little research to identify which specific platforms OSH, IIA or TA-CSA occur on, and little to no investigation of these risks on gaming platforms or direct messaging services. What little data exists suggests that networking platforms that are popular with children (e.g., Snapchat, Instagram) are the most frequent places where they are exposed to OSH, IIA and TA-CSA.

- CSAM is widespread on the dark web, but also on the open web including large global platforms like Facebook. However, it is not possible to determine with any certainty which platforms host relatively more CSAM. This is because some platforms make proactive efforts to detect this kind of material, whereas others (particularly those that deploy end-to-end encryption) choose not to and, consequently, report fewer instances of sexual exploitation, potentially misrepresenting how much sexual abuse they actually host.

Children most likely to be harmed

- Girls have more encounters than boys with every category of online sexual risk. More girls than boys experience OSH, most types of IIA, and TA-CSA, and girls are more commonly depicted in CSAM. There is also evidence that victimisation through IIA has gendered dimensions and can have more severe psychological and social impacts for girls.
- Age is an important factor in children's vulnerability to online sexual risk and harm. Older adolescents report experiences of OSH, IIA and TA-CSA more frequently than younger children. On the other hand, prepubescents and younger children are most commonly depicted in CSAM, suggesting that measures like self-report victimisation surveys may be failing to capture the extent of victimisation of younger children.

Responses to encountering sexual online risks

- Children respond to abusive situations in a variety of ways: some engage with the perpetrator; others delete unwanted messages or images or use technical tools provided by platforms to mitigate further risk, for example by blocking the sender. Making reports to platforms and seeking help from adults appears to be less common.
- One of the most common responses is to do nothing. In cases of IIA, this may be linked to the perception that victimisation is a normal and inevitable consequence of being online. In cases of TA-CSA and CSAM, it may result from children not recognising the situation as abusive or not realising that images of themselves are being taken and shared. There is evidence that children have concerns about the reactions and judgements of adults due to the sexual nature of all these types of abuse, and that this is a significant barrier to reporting victimisation. Further barriers include threats by perpetrators, the experience of self-blame and shame, or fear that their reports will not be believed or taken seriously.

Outcomes and impacts for children who encounter sexual online risks

- There are similar emotional, psychological and social outcomes associated with victimisation across the different categories of sexual risk and harm examined. This demonstrates that online sexual victimisation, regardless of whether it is caused by adults or peers, can have significant and negative impacts on children's mental health and social relationships. The psychological impacts are further intensified by the persistence of sexual images online and the difficulties of preventing their ongoing dissemination.

Research Question 3: What does the current evidence suggest about children's online exposure to other types of online risk and harm?

This review also examined the evidence base related to other categories of online risk and harm, focusing on those classified in the Online Safety Act as 'primary priority' and 'priority' content (DSIT, 2023). The categories examined were as follows: sexual content and online pornography; content that encourages suicide, self-harm and eating disorders; violent content; content promoting dares or challenges; hate-related content; and cyberbullying. Extremist and radicalisation content was also examined, as an example of illegal (but not sexual) content.

Studies on these topics mainly focus on the prevalence of children's exposure and engagement. Fewer studies have examined children's responses or outcomes. The literature on cyberbullying is a notable exception: the evidence base on this topic is more developed and provides a fairly detailed understanding of risk factors for victimisation and perpetration, as well as the psychological impacts for children who experience it.

Scale of exposure to other types of online risk

- While there is a notable degree of variation in estimates of prevalence, the evidence consistently suggests that a minimum of 1 in 12 children has encountered 'primary priority' or 'priority' content risks.
- Exposure to 'harmful' content is more common than most types of online sexual victimisation.

Responses to encountering other types of online risks

- Research suggests that children who are exposed to the types of content reviewed in Chapter 6 may tell someone about their experience or block a sender; they are less likely to report the problem to a platform, either because they do not know about this function or feel it would be pointless to do so. The likelihood of telling someone about the experience seems to be higher than for children who experience sexual online risks.

Outcomes and impacts

- The light touch review that was undertaken found only modest amounts of evidence about the outcomes and impacts of exposure to, or engagement with, these categories of online risk and harm. The available evidence suggests there are some commonalities with the emotional and psychological impacts identified for online sexual risk categories, as well as specific additional attitudinal and behavioural impacts and potential health-related outcomes.

Research Question 4: What does the current evidence suggest about the ways in which technological design features moderate the likelihood of children encountering or being harmed by online risks?

How the design of platforms can increase risk

- Recommender algorithms can provide focused and intense exposure to harmful content, which can be particularly problematic for children experiencing mental health problems. Friend recommendations can result in children having many unknown followers or contacts who can pose risks to them. Popularity metrics indirectly encourage children to make new (and sometimes indiscriminate) contacts and to spend long periods of time on platforms, increasing their potential exposure to harmful content and contact.

How the design of platforms and technological tools can decrease risk

- There are a variety of tools that can reduce children's exposure to online risk and harm. These include age-assurance methods and parental controls, which are most effective if they are robust, consistently applied, easy to implement, and can be adapted to different stages of children's development. Tools to block, mute and ignore unwanted interactions, or to flag and report harmful content can potentially prevent children's continued exposure to risks they have already encountered. However, there are various barriers to children using these, 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 to reduce the risks to children. These can be effective in identifying illegal content and behaviour, but data about detection rates and outcomes is often lacking, making it difficult to measure how well they work. The fact that automated tools have a poor degree of contextual understanding can limit their efficacy.
- The wider use of end-to-end encryption by platforms in the future will reduce the ability of platforms and law enforcement to detect TA-CSA and CSAM, with implications for child protection.

Each of the technical tools that can reduce children's exposure to risk and harm online has a role to play in their safety, and in the creation of digital environments that enable children to maximise positive opportunities and digital participation. However, this is counterbalanced by platform design choices that increase exposure to harmful experiences. This tension needs to be considered within regulatory and child rights frameworks to ensure that protection from harm, privacy, and security are effectively balanced with the current and future design choices and business models of platforms.

Research Question 5: What are the gaps in current understandings of children's safety online and the harms affecting them?

The review has highlighted several key gaps in the evidence published since 2017 that require further empirical and policy attention.

Gaps regarding the scale of online sexual risk

- There has been a general lack of academic research published in the UK since 2017 examining the prevalence of online risk and harm to children (Livingstone et al, 2017). As a result, the evidence is fragmented, with much of this review drawing on grey literature studies or research that uses non-UK samples of children. This makes it difficult to assess the current landscape of online risk and harm for children in the UK, or to identify changes over time, particularly in relation to sexual risks and harm. It also creates challenges in establishing a baseline understanding against which the impact of the Online Safety Act can be measured.
- This is complicated by the methodological issues highlighted throughout the report. It is also difficult to directly compare the results of research studies due to the variety of definitions, measures and sampling strategies used. Many do not clearly differentiate between adult and peer perpetrators, or consensual and non-consensual online sexual interactions. It is disappointing that greater progress has not been made in this area given that the need for greater research standardisation in these areas was recommended in the 2017 report and has more recently been highlighted again as an important issue (Finkelhor et al, 2022; Revealing Reality, 2021b; UN, 2019).

Gaps regarding the evolving nature of online sexual risk

- The review indicated that recent developments in livestreaming, direct messaging and gaming have influenced children's exposure to sexual risk and harm. However, greater empirical understanding of how this has influenced the use of grooming strategies and sexual interactions between children and adults is needed.
- Another knowledge gap relates to the impact of generative AI and deepfake technology on the production of CSAM and non-photographic abuse imagery (NPAI), as well as how offenders engage with this material and whether it encourages or can deter further sexual abuse of children.
- There is also limited understanding of how perpetrators interact with each other on the open and dark web: for example, how they share links and direct each other to CSAM.

Gaps regarding the platforms on which online sexual risks occur

- There is inadequate understanding of the prevalence of children's exposure to online risk and harm on specific platforms, as well as of how platforms' design features facilitate (or prevent) these experiences. The information provided in transparency reports is limited, making it difficult to use them to draw conclusions about platform safety or changes over time. It is also difficult to assess the prevalence of sexual risk and harm on platforms that use E2EE, where there is limited use of tools to detect TA-CSA and CSAM.

- Greater transparency is needed in relation to the operation of algorithms, content moderation systems and the detection tools used by platforms. The current lack of information creates challenges for policy makers, researchers, the online regulator, and the public when assessing whether (and to what extent) these technological factors increase or decrease children's exposure to online risk and harm.

Gaps regarding risk factors and vulnerability to online sexual risk

- There is currently a paucity of UK research examining factors that increase children's vulnerability to online sexual risk and harm. While the literature suggests that age and gender are important factors, understanding of the influence of other demographic characteristics is underdeveloped, as is the role of intersectionality. There is also less understanding of the interactions between the demographic, psychological and environmental factors that may increase the likelihood that online risk exposure will lead to harm.
- Relatively little is known about how vulnerability factors operate across multiple domains of online risk and harm, leading to the victimisation of the same child in multiple ways (polyvictimisation). Online sexual exploitation by adults and peers is generally addressed in separate research literatures, with little empirical attention to the extent to which these experiences are inter-related, or whether victims engage with other forms of harmful content and contact as a means of coping with the resulting distress (e.g., accessing self-harm material and communities). There is some evidence of relationships between victimisation through cyberbullying and IIA (Barroso et al, 2021; Zych et al, 2019), but it is important to develop greater understanding of polyvictimisation, particularly in the context of risk factors, vulnerability, and pathways to harm.

Gaps regarding the outcomes and impacts of online sexual victimisation

- The review identified a variety of emotional, psychological and social outcomes associated with exposure to online sexual risk and harm. These were very similar across different categories of online victimisation. However, there is a lack of evidence about the impacts experienced by children in the UK, as well as how this can be exacerbated by specific vulnerability factors and polyvictimisation.

Chapter 9
Research recommendations

This report has highlighted a number of knowledge gaps that limit an effective assessment of the current online risk and harm landscape for children in the UK. There is a clear need to improve the evidence base given that there has been limited development since 2017, particularly in relation to sexual risk and harm.

This chapter provides recommendations for future research and data transparency. While it may be the responsibility of academics and other researchers working in this field to cover the topics below and effectively design their research questions and studies, responsibility also lies with policymakers, regulators, and the technology industry. It is up to policymakers, Ofcom and other regulators to initiate, fund, and oversee the research recommended below; and it is the technology industry's responsibility to work with government and regulators, feeding in data and expert insight where required.

9.1 Overall recommendations for future research

Prevalence of online risk and harm

- 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 the capture of data at the very beginning of the enactment of the Online Safety Act are essential and should be a priority for policy makers and researchers.
- The research 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 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. However, such an approach will enable the development of baseline measures against which change over time and the impact of the Online Safety Act can be assessed. It will also allow an examination of the extent to which it 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).

Methodology

- There should be greater standardisation of the definitions used by researchers, and specificity of research measures to distinguish more effectively between different categories of sexual risk and harm (e.g., IIA and TA-CSA). This is an important step in enabling greater comparability of studies, as well as ensuring consistency between research evidence and policy.
- Use of longitudinal research designs is also important: these can provide greater understanding of the relationships between exposure to risk and harm, vulnerability factors, and outcomes that may be delayed or manifest gradually over time. Consideration should be given to how the results of longitudinal research are interpreted, given that technology constantly evolves, and that the online behaviour of children (and of perpetrators) will always be subject to change over time.

Exposure to sexual risk and harm

- Research that provides a more specific and detailed examination of each category of sexual victimisation (OSH, IIA, TA-CSA and CSAM) is essential, as this is a significant knowledge gap at present, particularly in the UK.
- Empirical evidence is needed to develop greater understanding of the dynamics of sexual interactions between children and adults on individual platforms, particularly those that are most popular with children (e.g., Snapchat, Instagram). This should include examination of how design features and functionalities can potentially help to initiate or maintain sexual communication, as well as their ability to prevent or disrupt it. Particular attention needs to be given to the types of platforms that this review has identified as being under-researched, most notably direct messaging services and gaming sites.
- The emergence of livestreaming as a significant location for TA-CSA highlights the need for research that examines how grooming strategies are used, particularly when there are multiple perpetrators involved.
- The development of more empirical evidence about children's exposure to sexual risk and harm in virtual reality spaces is also important given that the embodied and immersive experience of being in these spaces has the potential to intensify the experience of sexual harm for children. Such research will ensure that policy makers have a clear understanding of how children are potentially harmed in these spaces prior to their wider use.
- It is important for research to examine how generative AI and other technology that may be used to create deepfakes, synthetic CSAM and NPAI of children can impact on the production and distribution of CSAM. Research should explore how perpetrators engage with this material and how offenders interact with each other on the open and dark web.

Vulnerability, outcomes and impacts

- Further research is needed on how demographic, psychological and environmental factors combine to increase children's vulnerability to different types of online risk and harm, as well as how this can lead to the victimisation of the same child in multiple ways (polyvictimisation). Moreover, there is a need for greater understanding of how exposure to online risk and harm relates to offline victimisation (e.g., physical violence, sexual exploitation) and to polyvictimisation. Some children may experience harm across these different domains and experience post-traumatic stress disorder as a result. Developing further understanding of the pathways by which polyvictimisation occurs has implications for practitioners when supporting children who present with specific psychological difficulties or victimisation experiences. Greater understanding in these areas will inform the development of more effective strategies to prevent harm, build resilience and develop media literacy for children.

- There has also been less empirical examination of children engaging in harmful sexual behaviour online, and how this relates to involvement in offline offending. This is important given that it may be part of a wider pattern of anti-social behaviour for some children. Developing improved understanding in this area has implications for criminal justice and youth workers: examining the online lives and activities of children engaging with their services will be important in developing effective interventions and diversion strategies.

9.2 Recommendations on how technology companies can contribute to building the evidence base

- Platforms should increase the information they provide in transparency reports about the level of risk and harm to children facilitated by their products. This should include:
  - The number of reports and detections across different categories of risk and harm to children, including data on:
    - The amount of CSAM 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 TA-CSA.
  - Levels of usage of safety tools and reporting processes 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 content and contact that may be harmful.
  - The type of age-assurance processes implemented.
  - The level of effective liaison with law enforcement, helplines and organisations like the IWF and NCMEC.
  - Processes by which safety tools, reporting processes, content moderation, detection tools, and age-assurance processes are evaluated.
- Making this data available (as raw data in addition to aggregated data or summaries) will enable a more effective assessment of levels of risk and harm on different platforms, and the identification of problematic or risky design choices.
- Greater transparency will also enable policy makers and researchers to make a more effective assessment of the efficacy of platforms' actions in protecting children from online risk and harm.
- This information is likely to be required under the Online Safety Act as part of platform risk assessments, but providers should not wait until regulatory requirements are imposed to increase transparency.

9.3 Recommendations on how policy makers and regulators can contribute to building the evidence base

- Specific and long-term funding should be provided for research in this area, particularly for studies related to online sexual risk and harm. This will support the development of a robust evidence base that can be used to inform policy decisions and other strategies to increase the online safety of children.
- 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, as well as the development of related prevention and response strategies.
- A framework should be initiated 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. Having this in place will ensure that policy and regulatory systems are able to effectively respond to risks associated with emerging technologies.
- Regulators should create and enforce robust and effective regulatory standards that ensure platforms provide:
    - Detailed transparency reports that address the recommendations described in the previous section.
    - Comprehensive risk assessments, including information about recommendation algorithms and content moderation, as well as the efficacy of reporting and detection tools.
- Policymakers and regulators should ensure there is effective collaboration with platforms and the technology industry, as well as other stakeholders, and receive relevant input and guidance in these developments.

References

5Rights Foundation (2021a) *But how do they know it is a child? Age Assurance in the Digital World*. https://5rightsfoundation.com/uploads/But_How_Do_They_Know_It_is_a_Child.pdf

5Rights Foundation (2021b) *Pathways: How digital design puts children at risk.* https://5rightsfoundation.com/uploads/Pathways-how-digital-design-puts-children-at-risk.pdf

5Rights Foundation (2023) *Disrupted childhood: The cost of persuasive design.* https://5rightsfoundation.com/uploads/Disrupted-Childhood-2023-v2.pdf

All-Party Parliamentary Group on Social Media (2021) *Selfie Generation, What's behind the rise of self-generated indecent images of children online?* https://uploads-ssl.webflow.com/6109364ea51f0b14f7efeb5c/613b6ef4224bfeb57bfb0ca0_APPG%20on%20Social%20Media%20-%20Selfie%20Generation.pdf

Allen, C. & McIntosh, V. (2023) *Child safeguarding and immersive technologies: an outline of the risks*. London: NSPCC. https://learning.nspcc.org.uk/media/3333/child-safeguarding-immersive-technologies-key-concepts.pdf

Augusti, E.M., Saetren, S.S. & Hafstad, G.S. (2021) Violence and abuse experiences and associated risk factors during the COVID-19 outbreak in a population-based sample of Norwegian adolescents. *Child Abuse & Neglect*, 118, Article 105156. https://doi.org/10.1016/j.chiabu.2021.105156

Bada, M. & Clayton, R. (2020) Online suicide games: A form of digital self-harm or a myth? In B. Wiederhold, G. Riva, & S. Debb (Eds.), *Annual Review of Cybertherapy and Telemedicine (ARCTT) International Association of CyberPsychology, Training, and Rehabilitation (iACToR), 2019*. https://doi.org/10.48550/arXiv.2012.00530

Barber, C.S. & Bettez, S.C. (2021) Exposing patterns of adult solicitor behaviour: towards a theory of control within the cybersexual abuse of youth. *European Journal of Information Systems*, 30(6), 591–622. https://doi.org/10.1080/0960085X.2020.1816146

Barbovschi, M., Ní Bhroin, N., Chronaki, D., Ciboci, L., Farrugia, L., Lauri, M.A., Ševčíková, A., Staksrud, E., Tsaliki, L. & Velicu, A. (2021) *Young people's experiences with sexual messages online. Prevalence, types of sexting and emotional responses across European countries*. EU Kids Online: University of Oslo. www.duo.uio.no/bitstream/handle/10852/88679/2021_SexualMessages_Online.pdf?sequence=4&isAllowed=y

Barrense-Dias, Y., Berchtold, A., Surís, J. & Akre, C. (2017) Sexting and the definition issue. *Journal of Adolescent Health*, 61(5), 544–554. https://doi.org/10.1016/j.jadohealth.2017.05.009

Barroso, R., Ramião, E., Figueiredo, P. & Araújo, A.M. (2021) Abusive sexting in adolescence: Prevalence and characteristics of abusers and victims. *Frontiers in Psychology*, 12. https://doi.org/10.3389/fpsyg.2021.610474

Barter, C., Lanau, A., Stanley, N., Aghtaie, N. & Överlien, C. (2022) Factors associated with the perpetration of interpersonal violence and abuse in young people's intimate relationships. *Journal of Youth Studies*, 25(5), 547–563. https://doi.org/10.1080/13676261.2021.1910223

BBC (2023, March 1) *TikTok sets 60-minute daily screen time limit for under-18s*. www.bbc.co.uk/news/technology-64813981

Bianchi, D., Morelli, M., Baiocco, R. & Chirumbolo, A. (2017) Sexting as the mirror on the wall: Body-esteem attribution, media models, and objectified-body consciousness. *Journal of Adolescence*, 61, 164–172. https://doi.org/10.1016/j.adolescence.2017.10.006.

Boer, S., Erdem, O., de Graaf, H. & Goetz, H. (2021) Prevalence and correlates of sext-sharing among a representative sample of youth in the Netherlands. *Frontiers in Psychology*, 12. https://doi.org/10.3389/fpsyg.2021.655796

Bond, E. & Phippen, A. (2019) *Police response to youth offending around the generation and distribution of indecent images of children and its implications*. University of Suffolk.

Borj, P.R., Raja, K. & Bours, P. (2023) Online grooming detection: A comprehensive survey of child exploitation in chat logs. *Knowledge-Based Systems*, 259, Article 110039. https://doi.org/10.1016/j.knosys.2022.110039

British Board of Film Classification (2020) *Young people, pornography and age-verification*.

Brooks, F., Klemera, E., Chester, K., Magnusson, J. & Spencer, N. (2020) *HBSC England national report: Findings from the 2018 HBSC study for England*. University of Hertfordshire. https://hbscengland.org/wp-content/uploads/2020/01/HBSC-England-National-Report-2020.pdf

Brown, A. (2017) What is hate speech? Part 1: The myth of hate. *Law and Philosophy*, 36, 419–468. https://doi.org/10.1007/s10982-017-9297-1

Bryce, J. (2010) Online sexual exploitation of children and young people. In Y. Jewkes & M. Yar (Eds.), *Handbook of internet crime* (pp.320–342). Willan Publishing.

Budde, J., Witz, C. & Bohm, M. (2022) Sexual boundary violations via digital media among students. *Frontiers in Psychology*, 12, Article 755752. https://doi.org/10.3389/fpsyg.2021.755752

Burén, J. & Lunde, C. (2018) Sexting among adolescents: A nuanced and gendered online challenge for young people. *Computers in Human Behavior*, 85, 210–217. https://doi.org/10.1016/j.chb.2018.02.003

Calvete, E., Fernández-González, L., Royuela-Colomer, E., Morea, A., Larrucea-Iruretagoyena, M., Machimbarrena, J.M., Gónzalez-Cabrera, J. & Orue, I. (2021) Moderating factors of the association between being sexually solicited by adults and active online sexual behaviors in adolescents. *Computers in Human Behavior*, 124, Article 106935. https://doi.org/10.1016/j.chb.2021.106935

Cambridge Consultants (2019) *Use of AI in online content moderation*. www.ofcom.org.uk/__data/assets/pdf_file/0028/157249/cambridge-consultants-ai-content-moderation.pdf

Canadian Centre for Child Protection Inc (C3P) (2021) *Project Arachnid: Online availability of child sexual abuse material*. https://protectchildren.ca/pdfs/C3P_ProjectArachnidReport_en.pdf

Canadian Centre for Child Protection Inc (C3P) (2023) *What is Project Arachnid?* https://www.projectarachnid.ca/en/#what-is-project-arachnid

Center for Countering Digital Hate (2022) *Deadly by design*. https://counterhate.com/research/deadly-by-design/

Chiang, E. & Grant, T. (2019) Deceptive identity performance: Offender moves and multiple identities in online child abuse conversations. *Applied Linguistics*, 40(4), 675–698. https://doi.org/10.1093/applin/amy007

Children's Commissioner (2020) *Access denied. How end-to-end encryption threatens children's safety online*. https://assets.childrenscommissioner.gov.uk/wpuploads/2020/12/cco-access-denied.pdf

Children's Commissioner (2022) *Digital childhoods: a survey of children and parents*. www.childrenscommissioner.gov.uk/wp-content/uploads/2022/09/cc-digital-childhoods-a-survey-of-children-and-parents.pdf

Children's Commissioner (2023) *'A lot of it is actually just abuse': Young people and pornography*. https://assets.childrenscommissioner.gov.uk/wpuploads/2023/02/cc-a-lot-of-it-is-actually-just-abuse-young-people-and-pornography-updated.pdf

Chiu, J. & Quayle, E. (2022) Understanding online grooming: An interpretative phenomenological analysis of adolescents' offline meetings with adult perpetrators. *Child Abuse & Neglect*, 128, Article 105600. https://doi.org/10.1016/j.chiabu.2022.105600

College of Policing (2016) *Police action in response to youth produced sexual imagery ('Sexting')*. https://library.college.police.uk/docs/college-of-policing/briefing-note-sexting-2016.pdf

Copp, J.E., Mumford, E.A. & Taylor, B.G. (2021) Online sexual harassment and cyberbullying in a nationally representative sample of teens: Prevalence, predictors, and consequences. *Journal of Adolescence*, 93, 202–211. https://doi.org/10.1016/j.adolescence.2021.10.003

Currin, J.M. & Hubach, R.D. (2019) "Motivations for non university-based adults who sext their relationship partners." *Journal of Sex & Marital Therapy*, 45 (4), 317–327. https://doi.org/10.1080/0092623X.2018.1526837

Davidson, J., DeMarco, J., Bifulco, A., Bogaerts, S., Caretti, V., Aiken, M., Cheevers, C., Corbari, E., Scally, M., Schilder, J., Schimmenti, A. & Puccia, A. (2017) *Enhancing police and industry practice*. European Child Online Safety Project. www.mdx.ac.uk/__data/assets/pdf_file/0017/250163/ISEC-report-FINAL.pdf

Davidson, J., Livingstone, S., Jenkins, S., Gekoski, A., Choak, C., Ike, T. & Phillips, K. (2019) *Adult online hate harassment and abuse: A rapid evidence assessment*. UK Council for Child Internet Safety (UKCCIS). https://repository.uel.ac.uk/item/89615

Davies, P. (2003) *The magenta book: Guidance notes for policy evaluation and analysis*. Cabinet Office.

DeMarco, J., Sharrock, S., Crowther, T. & Barnard, M. (2018) *Behaviour and characteristics of perpetrators of online-facilitated child sexual abuse and exploitation: A rapid evidence assessment*. NatCen Social Research. www.iicsa.org.uk/document/rapid-evidence-assessment-behaviour-and-characteristics-perpetrators-online-facilitated

DeMarco, J., Cheevers, C., Davidson, J., Bogaerts, S., Pace, U., Aiken, M., Caretti, V., Schimmenti, A. & Bifulco, A. (2017) Digital dangers and cyber-victimisation: A study of European adolescent online risky behaviour for sexual exploitation. *Clinical Neuropsychiatry*, 14(1), 104–112.

Department for Digital, Culture, Media, and Sport (2020) *VoCO (Verification of children online). Phase 2 Report*. https://assets.publishing.service.gov.uk/media/5faa9cffd3bf7f03a841cfc2/November_VoCO_report_V4__pdf.pdf

Department of Science, Innovation and Technology (2023) *Online Safety Bill: Government amendments at Lords report stage*. https://www.gov.uk/government/publications/online-safety-bill-government-amendments-at-lords-report-stage/online-safety-bill-government-amendments-at-lords-report-stage#age-assurance-and-age-verification

de Santisteban, P. & Gámez-Guadix, M. (2018) Prevalence and risk factors among minors for online sexual solicitations and interactions with adults. *The Journal of Sex Research*, 55(7), 939–950. https://doi.org/10.1080/00224499.2017.1386763

Digital Regulation Cooperation Forum (2022) *The benefits and harms of algorithms: A shared perspective from the four digital regulators*. www.gov.uk/government/publications/findings-from-the-drcf-algorithmic-processing-workstream-spring-2022/the-benefits-and-harms-of-algorithms-a-shared-perspective-from-the-four-digital-regulators

Dincelli, E. & Yayla, A. (2022) Immersive virtual reality in the age of the metaverse: A hybrid-narrative review based on the technology affordance perspective. *The Journal of Strategic Information Systems*, 31(2), Article 101717. https://doi.org/10.1016/j.jsis.2022.101717

Dixon-Woods, M., Agarwal, S., Jones, D., Young, B. & Sutton, A. (2005) Synthesising qualitative and quantitative evidence: A review of possible methods. *Journal of Health Service Research Policy*, 10, 45–53. https://doi.org/10.1177/135581960501000110

Dolev-Cohen, M., Nezer, I. & Zumt, A.A. (2022) A qualitative examination of school counselors' experiences of sextortion cases of female students in Israel. *Sexual Abuse*. Advance online publication. https://doi.org/10.1177/10790632221145925

Döring, N. (2014) Consensual sexting among adolescents: Risk prevention through abstinence education or safer sexting? *Cyberpsychology: Journal of Psychosocial Research on Cyberspace*, 8, Article 9. https://doi.org/10.5817/CP2014-1-9

Draper, L. (2022) *Protecting children in the age of end-to-end encryption*. Joint PIJIP/TLS Research Paper Series. 80. https://digitalcommons.wcl.american.edu/research/80

Eelmaa, S. (2022) Sexualization of children in deepfakes and hentai. *Trames: A Journal of the Humanities and Social Sciences*, 26, 229–248. https://doi.org/10.3176/tr.2022.2.07

Englander, E. & McCoy, M. (2018) Sexting – prevalence, age, sex, and outcomes. *JAMA Pediatrics*, 172(4), 317–318. https://doi.org/10.1001/jamapediatrics.2017.5682

Europol (2020) *Internet organised crime threat assessment*. www.europol.europa.eu/cms/sites/default/files/documents/internet_organised_crime_threat_assessment_iocta_2020.pdf

Facebook (n.d.) *Community standards enforcement report: Child endangerment: Nudity and physical abuse and sexual exploitation*. https://transparency.fb.com/data/community-standards-enforcement/child-nudity-and-sexual-exploitation/facebook/#content-actioned

Finkelhor, D., Turner, H. & Colburn, D. (2022) Prevalence of online sexual offenses against children in the US. *JAMA Network Open*, 5(10), Article e2234471. https://doi.org/10.1001/jamanetworkopen.2022.34471

Finkelhor, D., Ormrod, R.K. & Turner, H.A. (2007) Poly-victimization: A neglected component in child victimization. *Child Abuse and Neglect*, 31, 7–26. https://doi.org/10.1016/j.chiabu.2006.06.008

Gámez-Guadix, M. & Mateos-Pérez, E. (2019) Longitudinal and reciprocal relationships between sexting, online sexual solicitations, and cyberbullying among minors. *Computers in Human Behavior*, 94, 70–76. https://doi.org/10.1016/j.chb.2019.01.004

Gámez-Guadix, M., Sorrel, M.A. & Martínez-Bacaicoa, J. (2022) Technology-facilitated sexual violence perpetration and victimization among adolescents: A network analysis. *Sexuality Research & Social Policy*, 20, 1,000–1,012. https://doi.org/10.1007/s13178-022-00775-y

Gámez-Guadix, M., Mateos-Pérez, E., Wachs, S., Wright, M., Martínez, J. & Íncera, D. (2022) Assessing image-based sexual abuse: Measurement, prevalence, and temporal stability of sextortion and nonconsensual sexting ("revenge porn") among adolescents. *Journal of Adolescence*, 94, 789–799. https://doi.org/10.1002/jad.12064

Gewirtz-Meydan, A., Mitchell, K.J. & Rothman, E.F. (2018a) What do kids think about sexting? *Computers in Human Behavior*, 86, 256–265. https://doi.org/10.1016/j.chb.2018.04.007

Gewirtz-Meydan, A., Walsh, W., Wolak, J. & Finkelhor, D. (2018b) The complex experience of child pornography survivors. *Child Abuse & Neglect*, 80, 238–248. https://doi.org/10.1016/j.chiabu.2018.03.031

Gill, V., Monk, L. & Day, L. (2022) *Qualitative research project to investigate the impact of online harms on children*. Ecorys. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1123422/Qualitative_Research_Project_to_Investigate_the_Impact_of_Online_Harms_on_Children_.pdf

Gillespie, T. (2022) Do not recommend? Reduction as a form of content moderation. *Social Media + Society*, 8(3), 1–13. https://doi.org/10.1177/20563051221117552

Google (2023) *YouTube Community Guidelines enforcement*. https://transparencyreport.google.com/youtube-policy/removals?total_removed_videos=period:2022Q4;exclude_automated:human_only&lu=total_removed_videos

Greene-Colozzi, E.A., Winters, G.M., Blasko, B. & Jeglic, E.L. (2020) Experiences and perceptions of online sexual solicitation and grooming of minors: A retrospective report. *Journal of Child Sexual Abuse*, 29(7), 836–854. https://doi.org/10.1080/10538712.2020.1801938

Guerra, C., Pinto-Cortez, C., Toro, E., Efthymiadou, E. & Quayle, E. (2021) Online sexual harassment and depression in Chilean adolescents: Variations based on gender and age of the offender. *Child Abuse & Neglect*, 120, Article 015219. https://doi.org/10.1016/j.chiabu.2021.105219

Guerra, C., Aguilea, G., Lippians, C., Navarro, M., Paz, M., Rebolledo, D., Silva, G. & Alaeddine, R. (2022) Online sexual abuse and symptomatology in Chilean adolescents: The role of peer support. *Journal of Interpersonal Violence*, 37(7–8), NP5805–NP5817. https://doi.org/10.1177/0886260520957685

Hamilton-Giachritsis, C., Hanson, E., Whittle, H., Alves-Costa, F. & Beech, A. (2020) Technology assisted child sexual abuse in the UK: Young people's views on the impact of online sexual abuse. *Children and Youth Services Review*, 119, Article 105451. https://doi.org/10.1016/j.childyouth.2020.105451

Hamilton-Giachritsis, C., Hanson, E., Whittle, H., & Beech, A. (2017). *"Everyone deserves to be happy and safe": a mixed methods study exploring how online and offline child sexual abuse impact young people and how professionals respond to it (PDF).* London: NSPCC. https://learning.nspcc.org.uk/media/1123/impact-online-offline-child-sexual-abuse.pdf

Henry, N., Flynn, A. & Powell, A. (2019) Image-based sexual abuse: Victims and perpetrators. *Trends and Issues in Crime and Criminal Justice*, 572, 1–19.

Hilton, Z., Brion-Meisels, G. & Graham, R. (2021) *Exploring effective prevention education responses to dangerous online challenges*. Praesidio Safeguarding. https://praesidiosafeguarding.co.uk/safe-guarding/uploads/2021/11/Exploring-effective-prevention-education-responses-to-dangerous-online-challenges-English-UK-compressed-1.pdf?x70166=

HM Government (2021) *Tackling violence against women and girls strategy*. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1033934/Tackling_Violence_Against_Women_and_Girls_Strategy_-_July_2021.pdf

Home Office (2018) *The Child Abuse Image Database (CAID)* https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/759328/CAID_Brochure_May_2018_for_gov_uk.pdf

Home Office (2019a) *Factsheet: Encryption*. https://homeofficemedia.blog.gov.uk/2019/11/05/factsheet-encryption/

Home Office (2019b) *Pioneering new tools to be rolled out in fight against child abusers*. www.gov.uk/government/news/pioneering-new-tools-to-be-rolled-out-in-fight-against-child-abusers

Home Office (2020) *Interim code of practice on online child sexual exploitation and abuse*. www.gov.uk/government/publications/online-harms-interim-codes-of-practice/interim-code-of-practice-on-online-child-sexual-exploitation-and-abuse-accessible-version

Home Office (2022) *Hate crime, England and Wales, 2021 to 2022*. www.gov.uk/government/statistics/hate-crime-england-and-wales-2021-to-2022/hate-crime-england-and-wales-2021-to-2022

Home Office (2023a) *Police recorded crime and outcomes open data tables*. https://assets.publishing.service.gov.uk/media/652eb312697260000dccf9ea/prc-pfa-mar2013-onwards-tables-191023.ods

Home Office (2023b) *International statement: End-to-end encryption and public safety (accessible version)*. www.gov.uk/government/publications/international-statement-end-to-end-encryption-and-public-safety/international-statement-end-to-end-encryption-and-public-safety-accessible-version

Hudson, N., Haux, T., Kersting, F., MacNaboe, L., David, M., McDonough, T., Phillips, N., Woolfe, E. & Myers, C.A. (2022) *Content and activity that is harmful to children within scope of the Online Safety Bill: A rapid evidence assessment*. NatCen. https://natcen.ac.uk/publications/content-and-activity-harmful-children-within-scope-online-safety-bill

Hunehäll Berndtsson, K.H. & Odenbring, Y. (2021) 'They don't even think about what the girl might think about it': students' views on sexting, gender inequalities and power relations in school. *Journal of Gender Studies*, 30(1), 91–101. https://doi.org/10.1080/09589236.2020.1825217

Intelliagg (2016) *Deeplight: Shining a light on the dark web*. Intelliagg. https://onyxcomms.com/wp-content/uploads/2017/01/intelliagg-deeplight-report.pdf

Internet Watch Foundation (2023) *Annual Report 2022*. https://annualreport2022.iwf.org.uk/

Internet Watch Foundation (2022) *Annual Report 2021*. https://annualreport2021.iwf.org.uk/

Internet Watch Foundation (2021) *Annual Report 2020*. https://annualreport2020.iwf.org.uk/

Irwin-Rogers, K. & Pinkney, C. (2017) *Social media as a catalyst and trigger for youth violence*. Catch22. www.ucb.ac.uk/media/0ndf5xgp/report-social-media-and-youth-violence.pdf

Jacob, N., Evans, R. & Scourfield, J. (2017) The influence of online images on self-harm: A qualitative study of young people aged 16–24. *Journal of Adolescence*, 60, 140–147. https://doi.org/10.1016/j.adolescence.2017.08.001

John, A., Lee, S.C., Puchades, A., Del Pozo-Baños, M., Morgan, K., Page, N., ... & Murphy, S. (2023) Self-harm, in-person bullying and cyberbullying in secondary school-aged children: A data linkage study in Wales. *Journal of Adolescence*, 95(1), 97–114. https://doi.org/10.1002/jad.12102

Joleby, M., Lunde, C., Landström, S. & Jonsson, L.S. (2021) Offender strategies for engaging children in online sexual activity. *Child Abuse & Neglect*, 120, Article 105214. https://doi.org/10.1016/j.chiabu.2021.105214

Jonsson, L.S., Fredlund, C., Priebe, G., Wadsby, M. & Svedin, C.G. (2019) Online sexual abuse of adolescents by a perpetrator met online: A cross-sectional study. *Child and Adolescent Psychiatry and Mental Health*, 11, 9. https://doi.org/10.1186/s13034-017-0146-7

Kansok-Dusche, J., Ballaschk, C., Krause, N., Zeißig, A., Seemann-Herz, L., Wachs, S. & Bilz, L. (2022) A systematic review on hate speech among children and adolescents: Definitions, prevalence, and overlap with related phenomena. *Trauma, Violence, & Abuse*, 24(4), 2,598–2,615. https://doi.org/10.1177/15248380221108070

Kardefelt-Winther, D., Day, E., Berman, G., Witting, S.K. and Bose, A. (2020) *Encryption, privacy and children's right to protection from harm* (WP-2020-14). UNICEF Office of Research – Innocenti. www.unicef-irc.org/publications/1152-encryption-privacy-and-childrens-right-to-protection-from-harm.html

Katz, A. & El-Asam, A. (2020) *Look at me. Teens, sexting and risks*. Internet Matters and Youthworks. https://www.internetmatters.org/wp-content/uploads/2020/06/Internet-Matters-Look-At-Me-Report-1.pdf

Katz, A., El-Asam, A. & Chahal, L.C. (2023) *Teens, tech and wellbeing during COVID Year 2*. The Cyber Survey. www.thecybersurvey.co.uk

Kloess, J.A., Hamilton-Giachritsis, C.E. & Beech, A.R. (2017) A descriptive account of victims' behaviour and responses in sexually exploitative interactions with offenders. *Psychology, Crime and Law*, 23(7), 621–632. https://doi.org/10.1080/1068316X.2017.1293052

Kloess, J.A., Hamilton-Giachritsis, C.E., & Beech, A.R. (2019) Offense processes of online sexual grooming and abuse of children via internet communication platforms. *Sexual Abuse: Journal of Research and Treatment*, 31(1), 73–96. https://doi.org/10.1177/1079063217720927

Kostyrka-Allchorne, K., Stoilova, M., Bourgaize, J., Rahali, M., Livingstone, S. & Sonuga-Barke, E. (2023) Digital experiences and their impact on the lives of adolescents with pre-existing anxiety, depression, eating and nonsuicidal self-injury conditions–a systematic review. *Child and Adolescent Mental Health*, 28(1), 22–32. https://doi.org/10.1111/camh.12619

Kowalski, R.M., Giumetti, G.W., Schroeder, A.N. & Lattanner, M.R. (2014) Bullying in the digital age: A critical review and meta-analysis of cyberbullying research among youth. *Psychological Bulletin*, 140(4), 1,073–1,137. https://doi.org/10.1037/a0035618

Langdon-Shreeve, S., Nickson, H. & Bright, C. (2021) *Safeguarding and radicalisation: Learning from children's social care*. Department for Education. https://dera.ioe.ac.uk/id/eprint/38118/

Livingstone, S., Davidson, J., Bryce, J., Millwood Hargrave, A. & Grove-Hills, J. (2012) *Children's online activities, risks and safety: The UK evidence base*. London: UK Council for Child Internet Safety. http://eprints.lse.ac.uk/id/eprint/69571

Livingstone, S. (2013) Online risk, harm and vulnerability: reflections on the evidence base for child Internet safety policy. *Journal of Communications studies*, 18(35), 13–28. http://eprints.lse.ac.uk/62278/

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

Livingstone, S. & Stoilova, M. (2021) *The 4Cs: Classifying online risk to children* (CO:RE Short Report Series on Key Topics) Leibniz-Institute for Media Research Hans-Bredow-Institut; CO:RE – Children Online: Research and Evidence. https://doi.org/10.21241/ssoar.71817

Livingstone, S., Stoilova, M., Stänicke, L.I., Jessen, R.S., Graham, R., Staksrud, E. & Jensen, T.K. (2022) *Young people experiencing internet related mental health difficulties: The benefits and risks of digital skills. An empirical study*. ySKILLS. https://eprints.lse.ac.uk/116407/

Lloyd, J. (2020) Abuse through sexual image sharing in schools: Response and responsibility. *Gender and Education*, 32(6), 784–802. https://doi.org/10.1080/09540253.2018.1513456

Longobardi, C., Fabris, M.A., Prino, L.E. & Settanni, M. (2021) The role of body image concerns in online sexual victimization among female adolescents: The mediating effect of risky online behaviors. *Journal of Child Adolescent Trauma*, 14, 51–60. https://doi.org/10.1007/s40653-020-00301-5

Lundy, L., Byrne, B., Templeton, M., & Lansdown, G. (2019) *Two clicks forward and one click back: report on children with disabilities in the digital environment*. Council of Europe. https://rm.coe.int/two-clicks-forward-and-one-click-back-report-on-children-with-disabili/168098bd0f

Madigan, S., Ly, A., Rash, C.L., Van Ouytsel, J. & Temple, J.R. (2018) Prevalence of multiple forms of sexting behavior among youth: A systematic review and meta-analysis. *Jama Pediatrics*, 172(4), 327–335. https://doi.org/10.1001/jamapediatrics.2017.5314

Mandau, M.B.H. (2021) 'Snaps', 'screenshots', and self-blame: A qualitative study of image-based sexual abuse victimization among adolescent Danish girls. *Journal of Children and Media*, 15(3), 431–447. https://doi.org/10.1080/17482798.2020.1848892

Martellozzo, E., Monaghan, A., Davidson, J. & Adler, J. (2020) Researching the affects that online pornography has on UK adolescents aged 11 to 16. *Sage Open*, 10(1). https://doi.org/10.1177/21582440198994

Mateu, A., Pascual-Sánchez, A., Martinez-Herves, M., Hickey, N., Nicholls, D. & Kramer, T. (2020) Cyberbullying and post-traumatic stress symptoms in UK adolescents. *Archives of Disease in Childhood*, 105(10), 951–956. http://dx.doi.org/10.1136/archdischild-2019-318716

May-Chahal, C.A. & Palmer, C.E. (2018) *Rapid evidence assessment: Characteristics and vulnerabilities of victims of online-facilitated child sexual abuse and exploitation*. Independent Inquiry into Child Sexual Abuse. https://webarchive.nationalarchives.gov.uk/ukgwa/20221216172020/https://www.iicsa.org.uk/key-documents/3719/view/rapid-evidence-assessment-characteristics-vulnerabilities-victims-online-facilitated-child-sexual-abuse-exploitation-.pdf

McGeeney, E. & Hanson, E. (2017) *Digital romance: A research project: Exploring young people's use of technology in their romantic relationships and love lives*. National Crime Agency; Brook. www.brook.org.uk/wp-content/uploads/2020/03/DR_REPORT_FINAL.pdf

McIntosh, V. & Allen, C. (2022) *Safeguarding the metaverse. Institution of Engineering and Technology (IET)* www.theiet.org/media/9836/safeguarding-the-metaverse.pdf

Mento, C., Silvestri, M.C., Muscatello, M.R.A., Rizzo, A., Celebre, L., Praticò, M., ... & Bruno, A. (2021) Psychological impact of pro-anorexia and pro-eating disorder websites on adolescent females: A systematic review. *International Journal of Environmental Research and Public Health*, 18(4), Article 2186. https://doi.org/10.3390/ijerph18042186

Merdian, H.L., Wefers, S., Dradshaw, H.K. & Perkins, D. (2022) *Non-photographic images of child sexual abuse: the risks and public policy responses*. British Board of Film Classification and onlinePROTECT report.

Mitchell, K. & Štulhofer, A. (2021) Online sexual harassment and negative mood in Croatian female adolescents. *European Child & Adolescent Psychiatry*, 30(2), 225–231. https://doi.org/10.1007/s00787-020-01506-7

Mordock, J. (2019, January 1) Darknet keeps exploding child-porn epidemic a step ahead of prosecutors. *The Washington Times*. www.washingtontimes.com/news/2019/jan/1/darknet-child-porn-users/

Napier, S., Teunissen, C. & Boxall, H. (2021a) *Live streaming of child sexual abuse: An analysis of offender chat logs*. (Trends & Issues in Crime and Criminal Justice no. 639). Australian Institute of Criminology. https://doi.org/10.52922/ti78375

Napier, S., Teunissen, C. & Boxall, H. (2021b) *How do child sexual abuse live streaming offenders access victims?* (Trends & Issues in Crime and Criminal Justice no. 642). Australian Institute of Criminology. https://doi.org/10.52922/ti78474

National Center for Missing & Exploited Children (2022a) *Office of Justice Programs, U.S. Department of Justice CY 2022 Report to the Committees on Appropriations*. www.missingkids.org/content/dam/missingkids/pdfs/OJJDP-NCMEC-Transparency_2022-Calendar-Year.pdf

National Center for Missing & Exploited Children (2022b) *CyberTipline 2021 Report*. www.missingkids.org/content/dam/missingkids/pdfs/2021-CyberTipline-Report.pdf

National Center for Missing & Exploited Children (2023) *2022 CyberTipline Reports by Electronic Service Providers (ESP)*. www.missingkids.org/content/dam/missingkids/pdfs/2022-reports-by-esp.pdf

National Crime Agency (2021) *National strategic assessment of serious and organised crime*. www.nationalcrimeagency.gov.uk/who-we-are/publications/533-national-strategic-assessment-of-serious-and-organised-crime-2021/file

National Society for the Prevention of Cruelty to Children (2021) *Private messaging and the rollout of end-to-end encryption. The implications for child protection*. www.nspcc.org.uk/globalassets/documents/news/nspcc-discussion-paper-private-messaging-and-the-roll-out-on-end-to-end-encryption.pdf

National Society for the Prevention of Cruelty to Children (2022a) *Online grooming crimes have risen by more than 80% in four years*. www.nspcc.org.uk/about-us/news-opinion/2022/online-grooming-crimes-rise/

National Society for the Prevention of Cruelty to Children (2022b) *Child sexual abuse crimes reach record levels – here's how the Online Safety Bill can effectively tackle grooming*. www.nspcc.org.uk/about-us/news-opinion/2022/child-sexual-abuse-crimes-reach-record-levels/

National Society for the Prevention of Cruelty to Children (2022c) *Children's experiences of legal but harmful content online*. Insight briefing. London: NSPCC. https://learning.nspcc.org.uk/media/2727/legal-but-harmful-content-online-helplines-insight-briefing.pdf

National Society for the Prevention of Cruelty to Children (2023a) *82% rise in online grooming crimes against children in the last 5 years*. www.nspcc.org.uk/about-us/news-opinion/2023/2023-08-14-82-rise-in-online-grooming-crimes-against-children-in-the-last-5-years/

National Society for the Prevention of Cruelty to Children (2023b) *We're calling for effective action in the Online Safety Bill as child abuse image crimes reach record levels*. www.nspcc.org.uk/about-us/news-opinion/2023/2023-02-22-were-calling-for-effective-action-in-the-online-safety-bill-as-child-abuse-image-crimes-reach-record-levels/

Nienierza, A., Reinemann, C., Fawzi, N., Riesmeyer, C. & Neumann, K. (2021) Too dark to see? Explaining adolescents' contact with online extremism and their ability to recognize it. *Information, Communication & Society*, 24(9), 1,229–1,246. https://doi.org/10.1080/1369118X.2019.1697339

Nilsson, M.G., Tzani-Pepelasis, C., Ioannou, M. & Lester, D. (2019) Understanding the link between sextortion and suicide. *International Journal of Cyber Criminology*, 13(1), 55–69. https://doi.org/10.5281/zenodo.3402357

Nominet (2022) *Digital Youth Index report*. https://digitalyouthindex.uk/wp-content/uploads/2022/10/Digital_Youth_Index_Year_2.pdf?utm_medium=referral&utm_source=Referral&utm_campaign=DYI_Report_2022&utm_content=DYI_Report_2022_referral

O'Connell, R. (2003) *A typology of child cybersexploitation and online grooming practices*. www.jisc.ac.uk/uploaded_documents/lis_PaperJPrice.pdf

Ofcom (2022a) *Research into risk factors that may lead children to harm online*. www.ofcom.org.uk/__data/assets/pdf_file/0021/245163/children-risk-factors-report.pdf

Ofcom (2022b) *Online experiences tracker waves 1 and 2 data tables*. www.ofcom.org.uk/__data/assets/excel_doc/0018/244161/online-experiences-tracker-wave-2-data-tables.xlsx

Ofcom (2022c) *Overview of perceptual hashing technology*. www.ofcom.org.uk/__data/assets/pdf_file/0036/247977/Perceptual-hashing-technology.pdf

Ofcom (2022d) *Children's online user ages. Quantitative research study*. https://www.ofcom.org.uk/__data/assets/pdf_file/0015/245004/children-user-ages-chart-pack.pdf

Ofcom (2023) *Children and parents: Media use and attitudes*. www.ofcom.org.uk/__data/assets/pdf_file/0027/255852/childrens-media-use-and-attitudes-report-2023.pdf

Office for National Statistics (2020) *Child sexual abuse in England and Wales: year ending March 2019*. www.ons.gov.uk/file?uri=/peoplepopulationandcommunity/crimeandjustice/datasets/childsexualabuseappendixtables/yearendingmarch2019/childsexualabuseappendixtablescorrectionfinal.xlsx

Office for National Statistics (2021) *Children's online behaviour in England and Wales: Year ending March 2020*. www.ons.gov.uk/peoplepopulationandcommunity/crimeandjustice/bulletins/childrensonlinebehaviourinenglandandwales/yearendingmarch2020

Office for National Statistics (2023a) *Sexual offences prevalence and trends, England and Wales: year ending March 2022*. www.ons.gov.uk/file?uri=/peoplepopulationandcommunity/crimeandjustice/datasets/crimeinenglandandwalesotherrelatedtables/yearendingmarch2022/otherrelatedtablesmar22.xlsx

Office for National Statistics (2023b) *Sexual offences prevalence and victim characteristics, England and Wales*. www.ons.gov.uk/file?uri=/peoplepopulationandcommunity/crimeandjustice/datasets/sexualoffencesprevalenceandvictimcharacteristicsenglandandwales/yearendingmarch2022/prevalenceandvictimcharacteristicsfinalv3.xlsx

Office for National Statistics (2023c) *Sexual offences in England and Wales overview: year ending March 2022*. www.ons.gov.uk/peoplepopulationandcommunity/crimeandjustice/bulletins/sexualoffencesinenglandandwalesoverview/march2022#police-recorded-crime

Office for National Statistics (2023d) *Crime in England and Wales: Year ending March 2023*. www.ons.gov.uk/file?uri=/peoplepopulationandcommunity/crimeandjustice/datasets/crimeinenglandandwalesotherrelatedtables/yearendingmarch2023/otherrelatedtablesyemar2023correction.xlsx

O'Malley, R.L. & Holt, K.M. (2022) Cyber sextortion: An exploratory analysis of different perpetrators engaging in a similar crime. *Journal of Interpersonal Violence*, 37(1–2), 258–283. https://doi.org/10.1177/0886260520909186

Ortega-Barón, J., Machimbarrena, J.M., Calvete, E., Orue, I., Pereda, N. & González-Cabrera, J. (2022) Epidemiology of online sexual solicitation and interaction of minors with adults: A longitudinal study. *Child Abuse & Neglect*, 131, Article 105759. https://doi.org/10.1016/j.chiabu.2022.105759

Owens, J.N., Clapp, K., Craun, S.W., van der Bruggen, M., van Balen, I., van Bunningen, A. & Talens, P. (2023) Analysis of topic popularity within a child sexual exploitation TOR hidden service. *Aggression and Violent Behavior*, 68, Article 101808. https://doi.org/10.1016/j.avb.2022.101808

Patchin, J.W. & Hinduja, S. (2019) The nature and extent of sexting among a national sample of middle and high school students in the US. *Archives of Sexual Behavior*, 48, 2,333–2,343. https://doi.org/10.1007/s10508-019-1449-y

Patchin, J.W. & Hinduja, S. (2020) Sextortion among adolescents: Results from a national survey of U.S. youth. *Sexual Abuse: A Journal of Research and Treatment*, 32(1), 30–54. https://doi.org/10.1177/1079063218800469

Patton, D.U., Pyrooz, D., Decker, S., Frey, W.R. & Leonard, P. (2019) When Twitter fingers turn to trigger fingers: A qualitative study of social media-related gang violence. *International Journal of Bullying Prevention*, 1, 205–217. https://doi.org/10.1007/s42380-019-00014-w

Pedersen, M.R., Schjørring Larsen, J. & Frederiksen, P. (2020) *Everyday pictures of children in sexualizing context*. A report by Red Barnet Denmark. https://redbarnet.dk/media/6274/everyday_pictures_scdk.pdf

Peersman, C., Llanos, J.T., May-Chahal, C., McConville, R., Das Chowdhury, P. and De Cristofaro, E. (2023) *Towards a framework for evaluating CSAM prevention and detection tools in the context of end-to-end encryption environments: A case study*. REPHRAIN. https://bpb-eu-w2.wpmucdn.com/blogs.bristol.ac.uk/dist/1/670/files/2023/02/Safety-Tech-Challenge-Fund-evaluation-framework-report.pdf

Pettifer, S., Barrett, E., Marsh, J., Hill, K., Turner, P. & Flynn, S. (2022) *The future of eXtended reality technologies, and implications for online child sexual exploitation and abuse*. WeProtect Global Alliance. www.weprotect.org/wp-content/uploads/2022-June-XR-OCSEA-FINAL-PUBLISHED.pdf

Project deSHAME (2017) *Young people's experiences of online sexual harassment. A cross-country report from Project deSHAME*. www.childnet.com/what-we-do/our-projects/project-deshame/research/

Przybylski, A.K. & Nash, V. (2018) Internet filtering and adolescent exposure to online sexual material. *Cyberpsychology, Behavior and Social Networking*, 21(7), 405–410. https://doi.org/10.1089/cyber.2017.0466

Quayle, E., Jonsson, L.S., Cooper, K., Traynor, J. & Svedin, C.G. (2018) Children in identified sexual images –Who are they? Self- and non-self-taken images in the International Child Sexual Exploitation Image Database 2006–2015. *Child Abuse Review*, 27(3), 223–238. https://doi.org/10.1002/car.2507

Quayle, E. & Cariola, L. (2019) Management of non-consensually shared youth-produced sexual images: A Delphi study with adolescents as experts. *Child Abuse & Neglect*, 95, Article 104064. https://doi.org/10.1016/j.chiabu.2019.104064

Quayle, E. (2022) Self-produced images, sexting, coercion and children's rights. *ERA Forum* 23, 237–251. https://doi.org/10.1007/s12027-022-00714-9

Razi A., Kim S., Alsoubai A., Stringhini G., Solorio T., De Choudhury M. & Wisniewski P.J. (2021) A human-centered systematic literature review of the computational approaches for online sexual risk detection. *Proceedings of the ACM on Human-Computer Interaction*, Vol. 5 (CSCW2), 1–38. https://doi.org/10.1145/3479609

Revealing Reality (2021a) *Not Just Flirting. The unequal experiences and consequences of nude image-sharing by young people*. www.revealingreality.co.uk/2022/06/23/not-just-flirting/

Revealing Reality (2021b) *Online harms feasibility study*. https://revealingreality.co.uk/report-launch-pathways-how-digital-design-puts-children-at-risk/

Revealing Reality (2023) *Anti-social media – what some vulnerable children are seeing on Snapchat*. https://revealingreality.co.uk/anti-social-media-what-some-vulnerable-children-are-seeing-on-snapchat/

Ricciardelli, R. & Adorjan, M. (2019) 'If a girl's photo gets sent around, that's a way bigger deal than if a guy's photo gets sent around': Gender, sexting, and the teenage years. *Journal of Gender Studies*, 28(5), 563–577. https://doi.org/10.1080/09589236.2018.1560245

Ringenberg, T.R., Seigfried-Spellar, K.C., Rayz, J.M., & Rogers, M.K. (2022) A scoping review of child grooming strategies: pre- and post-internet. *Child Abuse & Neglect*, 123, Article 105392. https://doi.org/10.1016/j.chiabu.2021.105392

Ringrose, J., Regehr, K. & Milne, B. (2021a) *Understanding and combatting youth experiences of image-based sexual harassment and abuse*. Department of Education, Practice and Society, UCL Institute of Education: London, UK. https://discovery.ucl.ac.uk/id/eprint/10139669

Ringrose, J., Regehr, K. & Whitehead, S. (2021b) Teen girls' experiences negotiating the ubiquitous dick pic: Sexual double standards and the normalization of image based sexual harassment. *Sex Roles*, 85(9), 558–576. https://link.springer.com/article/10.1007/s11199-021-01236-3

Ringrose, J., Regehr, K. & Whitehead, S. (2022) 'Wanna trade?': Cis-heteronormative homosocial masculinity and the normalization of abuse in youth digital sexual image exchange. *Journal of Gender Studies*, 31(2), 243–261. https://doi.org/10.1080/09589236.2021.1947206

Romanou, E. & Belton, E. (2020) *Isolated and struggling: social isolation and the risk of child maltreatment, in lockdown and beyond*. London: NSPCC. https://learning.nspcc.org.uk/media/2246/isolated-and-struggling-social-isolation-risk-child-maltreatment-lockdown-and-beyond.pdf

Sentencing Guidelines Council (2014) *Sexual Offences: Definitive Guideline*.

Setty, E. (2019) A rights-based approach to youth sexting: Challenging risk, shame, and the denial of rights to bodily and sexual expression within youth digital sexual culture. *International Journal of Bullying Prevention*, 1, 298–311. https://doi.org/10.1007/s42380-019-00050-6

Seymour-Smith, S. & Kloess, J.A. (2021) A discursive analysis of compliance, resistance and escalation to threats in sexually exploitative interactions between offenders and male children. *British Journal of Social Psychology*, 60(3), 988–1,011. https://doi.org/10.1111/bjso.12437

Smahel, D., Machackova, H., Mascheroni, G., Dedkova, L., Staksrud, E., Ólafsson, K., Livingstone, S. & Hasebrink, U. (2020) *EU Kids Online 2020: Survey results from 19 countries*. EU Kids Online. https://doi.org/10.21953/lse.47fdeqj01ofo

Smirnova, S., Livingstone, S. & Stoilova, M. (2021) *Understanding of user needs and problems: A rapid evidence review of age assurance and parental controls*. London School of Economics and Political Science (LSE). https://eprints.lse.ac.uk/112559/1/Stoilova_understanding_of_user_needs_and_problems_published.pdf

Smith, P.K., López-Castro, L., Robinson, S. & Görzig, A. (2019) Consistency of gender differences in bullying in cross-cultural surveys. *Aggression and Violent Behavior*, 45, 33–40. https://doi.org/10.1016/j.avb.2018.04.006

Spence, R., Harrison, A., Bradbury, P., Bleakley, P., Martellozzo, E. & DeMarco, J. (2023) Content moderators' strategies for coping with the stress of moderating content online. *Journal of Online Trust and Safety*, 1(5). https://doi.org/10.54501/jots.v1i5.91

Spielhofer, T. (2010). *Children's online risks and safety: a review of the available evidence*. Slough: NFER. https://www.nfer.ac.uk/media/d1hakhb0/coj01.pdf

Ståhl, S. & Dennhag, I. (2021) Online and offline sexual harassment associations of anxiety and depression in an adolescent sample. *Nordic Journal of Psychiatry*, 75(5), 330–335. https://doi.org/10.1080/08039488.2020.1856924

Steffen, J.H., Gaskin, J.E., Meservy, T.O., Jenkins, J.L. & Wolman, I. (2019) Framework of affordances for virtual reality and augmented reality. *Journal of Management Information Systems*, 36(3), 683–729. https://doi.org/10.1080/07421222.2019.1628877

Stoilova, M., Livingstone, S. & Khazbak, R. (2021) *Investigating risks and opportunities for children in a digital world: A rapid review of the evidence on children's internet use and outcomes*. Unicef. www.unicef-irc.org/publications/1183-investigating-risks-and-opportunities-for-children-in-a-digital-world.html

Strassberg, D.S., Cann, D. & Velarde, V. (2017) Sexting by high school students. *Archives of Sexual Behavior*, 46 (6), 1,667–1,672. https://doi.org/10.1007/s10508-016-0926-9

Tech Coalition (2021, June) *Multi-stakeholder forum: Charting a collective path forward*. https://paragonn-cdn.nyc3.cdn.digitaloceanspaces.com/technologycoalition.org/uploads/Multi-Stakeholder-Forum_R1.pdf

Tejeiro, R., Alison, L., Hendricks, E., Giles, S., Long, M. & Shipley, D. (2020) Sexual behaviours in indecent images of children: A content analysis. *International Journal of Cyber Criminology*, 14(1), 121–138. https://doi.org/10.5281/zenodo.3743390

Tener, D., Wolak, J. & Finkelhor, D. (2015) A Typology of offenders who use online communications to commit sex crimes against minors. *Journal of Aggression, Maltreatment & Trauma*, 24, 319–337. https://doi.org/10.1080/10926771.2015.1009602

The Centre of Expertise on Child Sexual Abuse & Centre for Abuse and Trauma Studies, Middlesex University (2020) *A new typology of child sexual abuse offending*. www.csacentre.org.uk/documents/new-typology-of-child-sexual-abuse-offending/

Thiel, D., Stroebel, M. & Portnoff, R. (2023) *Generative ML and CSAM: Implications and Mitigations*. A report by THORN and Stanford Internet Observatory Cyber Policy Center. https://stacks.stanford.edu/file/druid:jv206yg3793/20230624-sio-cg-csam-report.pdf

Thorn (2022) *Online Grooming: Examining risky encounters amid everyday digital socialization*. Thorn. https://info.thorn.org/hubfs/Research/2022_Online_Grooming_Report.pdf

Thorn (2021) *Responding to online threats: Minors' perspectives on disclosing, reporting, and blocking*. Thorn. https://info.thorn.org/hubfs/Research/Responding%20to%20Online%20Threats_2021-Full-Report.pdf

Trott, M., Driscoll, R., Iraldo, E. & Pardhan, S. (2022) Changes and correlates of screen time in adults and children during the COVID-19 pandemic: A systematic review and meta-analysis. *eClinicalMedicine*, 48, Article 101452. https://doi.org/10.1016/j.eclinm.2022.101452

UK Council for Internet Safety (2020) *Sharing nudes and semi-nudes: Advice for education settings working with children and young people*. www.gov.uk/government/publications/sharing-nudes-and-semi-nudes-advice-for-education-settings-working-with-children-and-young-people

United Nations (2019) *Guidelines regarding the implementation of the Optional Protocol to the Convention on the Rights of the Child on the sale of children, child prostitution and child pornography*. UN Doc. CRC/C/156. www.ohchr.org/sites/default/files/Documents/HRBodies/CRC/CRC.C.156_OPSC_Guidelines.pdf

Van Gijn-Grosvenor, E.L. & Lamb, M.E. (2021) Online groomer typology scheme. *Psychology, Crime & Law*, 27, 973–987. https://doi.org/10.1080/1068316X.2021.1876048

Van Ouytsel, J., Van Gool, E., Walrave, M., Ponnet, K. & Peeters, E. (2017) Sexting: Adolescents' perceptions of the applications used for, motives for, and consequences of sexting. *Journal of Youth Studies*, 20(4), 446–470. https://doi.org/10.1080/13676261.2016.1241865

Van Ouytsel, J., Walrave, M., De Marez, L., Vanhaelewyn, B. & Ponnet, K. (2021) Sexting, pressured sexting and image-based sexual abuse among a weighted-sample of heterosexual and LGB-youth. *Computers in Human Behavior*, 117, Article 106630. https://doi.org/10.1016/j.chb.2020.106630

Wachs, S., Bilz, L., Wettstein, A., Wright, M. F., Kansok-Dusche, J., Krause, N. & Ballaschk, C. (2022) Associations between witnessing and perpetrating online hate speech among adolescents: Testing moderation effects of moral disengagement and empathy. *Psychology of Violence*, 12(6), 371–381. https://doi.org/10.3390/ijerph16203992

Wager, N., Armitage, R., Christmann, K., Gallagher, B., Ioannou, M., Parkinson, S., Reevers, C., Rogerson, M. and J. Synnott (2018) *Rapid evidence assessment: quantifying the extent of online-facilitated child sexual abuse: Report for the Independent Inquiry into Child Sexual Abuse*. University of Huddersfield. www.iicsa.org.uk/key-documents/3722/view/rapid-evidence-assessment-quantifying-extent-online-facilitated-child-sexual-abuse-.pdf

Webster, S., Davidson, J., Bifulco, A., Gottschalk, P., Caretti, V., Pham, T., Grove-Hills, J., Turley, C., Tompkins, C., Ciulla, S., Milazzo, V., Schimmenti, A. & Craparo, G. (2012) *Final report. European Online Grooming Project*. European Online Grooming Project. https://europeanonlinegroomingproject.com/wp-content/file-uploads/European-Online-Grooming-Project-Executive-Summary.pdf

WeProtect Global Alliance (2018) *Global threat assessment 2018: Working together to end the sexual exploitation of children online*. www.weprotect.org/wp-content/uploads/Global-Threat-Assessment-2018-EN.pdf

WeProtect Global Alliance (2020) *Estimates of childhood exposure to online sexual harms and their risk factors. A global study of childhood experiences of 18 to 20 years olds*. https://www.weprotect.org/economist-impact-global-survey/#report

WeProtect Global Alliance (2021) *Global threat assessment*. www.weprotect.org/global-threat-assessment-21/#report

Whittle, H.C., Hamilton-Giachritsis, C. & Beech, A. (2013) Victims' voices: The impact of online grooming and sexual abuse. *Universal Journal of Psychology*, 1, 59–71. https://doi.org/10.13189/ujp.2013.010206

Wolak, J., Finkelhor, D. & Mitchell, K.J. (2012) How often are teens arrested for sexting? Data from a national sample of police cases. *Pediatrics*, 129(1), 4–12. https://doi.org/10.1542/peds.2011-2242

Wolak, J., Finkelhor, D., Walsh, W. & Treitman, L. (2018) Sextortion of minors: Characteristics and dynamics. *Journal of Adolescent Health*, 62(1), 72–79. https://doi.org/10.1016/j.jadohealth.2017.08.014

Woodhouse, J. (2022) *Regulating online harms*. House of commons library No 8743. https://researchbriefings.files.parliament.uk/documents/CBP-8743/CBP-8743.pdf

Zych, I., Farrington, D.P. & Ttofi, M.M. (2019) Protective factors against bullying and cyberbullying: A systematic review of meta-analyses. *Aggression and Violent Behavior*, 45, 4–19. https://doi.org/10.1016/j.avb.2018.06.008

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

Photography by Tom Hull and Abbie Trayler-Smith (page 34). The adult and children pictured are models.