Children's COMMISSIONER

Access denied

How end-to-end encryption threatens children's safety online

December 2020

Contents

Authors

Elizabeth Reeves, Simone Vibert
Contact: elizabeth.reeves@childrenscommissioner.gov.uk

Foreword by the Children's Commissioner, Anne Longfield

Today's generation of kids are born and raised in a world of instant connection. Digital technology now shapes all fundamental aspects of childhood – from play, friendship and family life, to education and the development of thoughts and ideas. While adults can be quick to panic about the long-term effects of a 'digital childhood', my time as Children's Commissioner has reassured me that kids' lives online are far from entirely negative. I've found today's generation of children to be hugely connected, engaged and informed – possibly more so than any generation before. Indeed, the Covid-19 lockdown demonstrated the huge advantages that children with access to a device and internet enjoy, in being able to stay connected with friends and family and access educational resources from home.

However, my experience in this role has also revealed to me the huge gulf between the risks to children of being online and the safeguards in place to protect them. While we rightfully demand protections where children may encounter harm in the physical world – in schools, hospitals and playgroups, for example – I have routinely found a dangerous absence of any similar safeguarding mechanisms in the digital sphere.

Previous reports from my Office have exposed the challenges experienced by children on social media and online gaming platforms. I have repeatedly called for products which are designed for use by children to be developed with their safety and wellbeing at heart – whether that means removing gambling elements from online games, simplifying terms and conditions on social media platforms or embedding greater transparency in tech firms' use of children's data.

The focus of this briefing turns to messaging services, like Whatsapp or Facebook Messenger, which have shot up in popularity in recent years. According to our survey, **nine in ten children aged 8-17 are using these services**. The vast majority are using messaging platforms to chat safely with friends and family. However, this is not always the case. **Over a third of children say that they have received something that made them feel uncomfortable on one of these messaging services**. 1 in 10 had received something from a stranger that made them feel uncomfortable. This was most likely to happen to teenage girls: around one in six girls aged 14-17 had received something distressing from a stranger. Because these services are, by their very nature, private, accessed on a child's own phone for example, it can be hard for parents, carers and teachers to keep track or understand who their children are contacting by direct message. **Almost one in ten children report using a messaging site to talk to people they don't already know**. Children, particularly older girls, are regularly sharing photos and videos that they have taken of themselves. While most images are shared between friends, 1 in 20 teenage girls say they are sharing photos with strangers.

Our findings also show that a significant proportion of children are using messaging platforms that they are not old enough to be accessing. **Over a third of 8-10-year-olds and over half of 11-13-year olds admit that they have said they were older than they were in order to sign up to an online messaging**

service. When asked which messaging sites they were regularly using, an even higher proportion of children – **60% of 8-year-olds and 90% of 12-year-olds** – reported using a messaging app with an age restriction of 13 or older. Three quarters of 12-year-olds reported using Whatsapp, despite the minimum age requirement for this platform being 16.

This is a concern because the privacy of direct messaging platforms can conceal some of the most serious crimes against children. In recent years there has been a sharp rise in the quantity of 'self-generated' child abuse material, in which children are groomed or bullied into recording their own abuse from a 'safe' space like their own bedroom. There has also been an exponential increase in the number of images and videos of child abuse flagged and reported by the largest tech firms to national clearinghouses like the Internet Watch Foundation. Many people assume that this kind of content is confined to the dark web – but official figures show the staggering rate at which child abuse material is distributed across major messaging sites like Facebook Messenger, which are also popular among children. Platforms are not yet legally required to scan for child abuse content, and not all do – meaning that the official figures are likely to be the tip of the iceberg.

Given these real risks to children, we would expect that tech firms would be planning to do more, not less, to aggressively root out and prevent the proliferation of child abuse on their platforms. However, last year we heard announcements from Facebook, and indications by other platforms such as Snap, that they plan to apply end-to-end encryption across all their messaging services. End to end encryption makes it impossible for the platform itself to read the contents of messages, and risks preventing police and prosecutors from gathering the evidence they need to prosecute perpetrators of child sexual exploitation and abuse. I worry that this could be a cynical attempt on the part of some tech firms to side-step sanctions and litigation, especially as the UK Government prepares to establish a new legal 'duty of care' on companies towards their users. If a platform is unable to read a message shared across their server, it follows that it would be hard for a Government to hold them accountable for its contents. I don't see this challenge as unassailable – but we must ensure that Government and the tech industry work together to embed child protection in end-to-end encryption, and that future technologies are designed with child safety in mind.

As Children's Commissioner, I have taken seriously my responsibilities to protect children from online harms. I was delighted that the Government heeded my calls and those of others in publishing the Online Harms White Paper and developing the Age Appropriate Design Code. However, the trend towards greater use of end-to-end encryption is just the latest sign that it is not enough for us to rely on technology companies to regulate themselves. The onus is now on Government to follow through on its promise to make the UK 'the safest place in the world to be online'. 18 months have now passed since the Online Harms White Paper was published and yet no legislation has been laid before Parliament and threats to children keep on growing. Now is a historic moment to reinforce our safety net for children online in order to ensure that all children are properly protected, now and for years to come.

Anne Longfield OBE
Children's Commissioner for England

Introduction

This briefing aims to understand which apps and sites children and teens are using to communicate, and to find out more about what they are sharing on these platforms. In March 2020 we polled 2,003 children aged 8-17 on their use of messaging platforms, in an effort to understand the risks that these services may pose to both children and teenagers. This briefing sets out our findings.

It is the latest research as part of the Commissioner's digital programme, which explores the full range of online platforms used by children – from social media, to gaming, and now messaging. Today's generation of children have grown up with digital technology as a central feature of their lives: an estimated one in three internet users around the world are children, [1] and half of ten year olds in the UK now own their own smartphone. [2] The biggest technology companies have a significant impact on all of our lives, including children's - as highlighted by plans to establish a new technology regulation unit at the Competition and Markets Authority.

The digital world has so much to offer children, from having fun and playing games, to connecting with friends and family – not to mention accessing vital educational resources during the Covid-19 crisis. But it was not designed with children in mind, and despite their overwhelming presence on these platforms, the digital world has not kept pace in keeping them from harm – as shown by successive reports by the Children's Commissioner's Office in recent years:

- Children often see content that is inappropriate for their age. For example, research commissioned by the CCO and the NSPCC in 2016 found that more than half of young people aged 11-16 have been exposed to online pornography.
- Children's confidence and self-esteem is ground down by what they see on social media. CCO's report "Life in Likes" found that while younger children use social media in a playful, creative way, for older children the pursuit of likes and comments is a way of seeking validation and cementing their place in their social circles. Children confided that social media can be a source of immense stress and pressure, with girls and boys as young as 10 telling us that they worried about their body image and the number of likes they were receiving on images and posts.
- Vast volumes of data are collected from and about children, with little transparency about where it goes and how it is used. CCO's report "Who Knows What About Me?" sets out how children aged 11-16 post on social media on average 26 times a day, which means by the age of 18 they are likely to have posted 70,000 times. By the age of 13, a child's parents will have posted on average 1,300 photos and videos of them on sites like Facebook or Instagram. [3]
- Children feel under pressure to spend money in online games and other platforms, in ways which can quickly get out of control. Our report "Gaming the System" found that some children are spending over £300 in a single year on in-game products, in addition to the upfront cost of the game itself. In some cases, this spending is done in order to receive a collection of unknown rewards, so-called "loot boxes", which children themselves described as "gambling".

This briefing sets out what needs to be done to protect children online so they can enjoy the benefits of the online world, whichever platforms they use.

[1] https://www.unicef-irc.org/publications/pdf/idp_2016_01.pdf
[2] https://www.ofcom.org.uk/__data/assets/pdf_file/0023/190616/children-media-use-attitudes-2019-report.pdf
[3] https://www.childrenscommissioner.gov.uk/report/who-knows-what-about-me/

Background

The age of instant messaging

In recent years, children's time and focus online has shifted towards more private methods of communication. As children's social media use has diversified, messaging platforms like Snapchat and Whatsapp have surged in popularity among younger users. Ofcom research shows that, over the last 5 years, Whatsapp and Snapchat, along with image-sharing platform Instagram, have eaten away at Facebook's dominance as the main online platform used by 12-15-year-olds: [4]

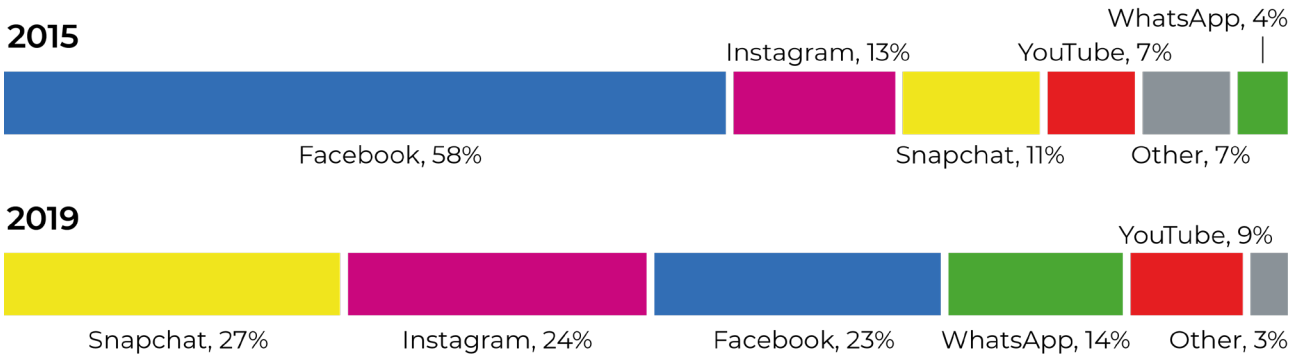


Figure 1: main social media/messaging site used by children aged 12-15, figures from Ofcom 'Children's media use and attitudes' 2015-2019

The scale of serious harm to children

The privacy of direct messaging platforms can conceal some of the most heinous crimes against children, including grooming, exploitation and the sharing of child sexual abuse material. An NSPCC investigation found that Facebook, Instagram and WhatsApp were used in child abuse image and online child sexual offences an average of 11 times a day in 2019. [5] The rate of grooming offences committed in the UK appears to have further accelerated over the course of lockdown, with 1,220 offences recorded in just the first three months of national lockdown – Facebook-owned apps (Facebook, Instagram, Whatsapp) accounted for 51% of these reports and Snapchat a further 20%. [6]

Recent figures have also exposed the staggering scale of the production and distribution of child sexual abuse material across the world's biggest social media and messaging platforms. In 2019 nearly 70 million pictures and videos were flagged as potentially containing images of child abuse, and sent to the US National Centre for Missing and Exploited Children (NCMEC) – a rate of over 100 images and videos a minute. [7] This figure has risen exponentially in recent years. [8] Facebook accounted for more than 85% of the total reports in 2019, reflecting both its significant share of the market and its aggressive approach to rooting out child abuse material.

Of particular concern is the rise in the volume of sexually explicit material shared by children. The Internet Watch Foundation (IWF) warned this year of the 'disturbing' rise in photos and videos of children who have been groomed or coerced into filming their own abuse. In the first 6 months of 2020,

[4] https://www.ofcom.org.uk/__data/assets/pdf_file/0023/190616/children-media-use-attitudes-2019-report.pdf
[5] https://www.nspcc.org.uk/about-us/news-opinion/2019/facebook-encryption-sexual-abuse/
[6] https://www.nspcc.org.uk/about-us/news-opinion/2020/instagram-grooming-crimes-children-lockdown/
[7] https://www.nytimes.com/2020/02/07/us/online-child-sexual-abuse.html
[8] http://www.cpcnetwork.org/wp-content/uploads/2019/07/Scope-of-CP-from-Cybertipline.pdf

44% of all the child sexual abuse content dealt with by IWF involved self-generated material, a rise of 15% on 2019. [9]

Children can also feel under pressure to share explicit images of themselves with friends or acquaintances they know in real life. [10] The sharing of sexually explicit material is an increasingly normal and accepted part of young people's social interactions and relationships. 2017-19 police figures show that children are being investigated for receiving intimate images from other children at increasingly younger ages. [11]

The move to end-to-end encryption

In 2019 **Facebook** [12] made public its intention to knit together its three major messaging platforms – Facebook Messenger, Instagram DM and Whatsapp – and to apply default end-to-end encryption across all three. **Snapchat** has indicated it is considering similar plans to apply end-to-end encryption to messages shared on Snapchat – snaps are already encrypted. [13]

End-to-end encryption ensures that the platform cannot decrypt messages as they flow through their servers. Tech companies suggest that this is to keep users 'safe' from hackers and spies; without the de-encryption key it would take longer than a lifetime to decode the contents of a message. Neither the platform itself nor the police are able to read the contents of end to end encrypted messages without either (a) physical access to the sender or recipient's device; (b) the supply of a 'backdoor key' [14]; or (c) by using malware to spy on the user's phone.

Some platforms currently give users the option to apply end-to-end encryption to their messages. Facebook's 'Secret Conversations' mode, for example, enables two friends to chat with end-to-end encryption. **Telegram** offers a similar optional 'Secret Conversation' mode. Others – notably **Whatsapp**, **Apple iMessage** and **Skype Instant Messenger** – apply end-to-end encryption as a default across all accounts.

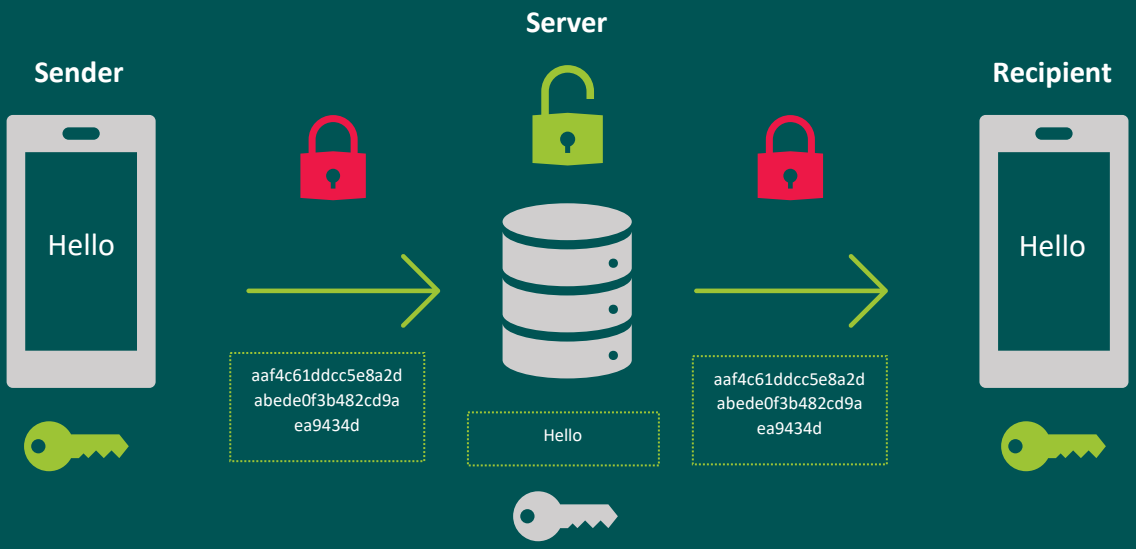
Facebook Messenger's 'Secret Conversation' mode

[9] https://www.iwf.org.uk/news/%E2%80%98disturbing%E2%80%99-rise-videos-of-children-who-have-been-groomed-into-filming-their-own-abuse
[10] https://www.nspcc.org.uk/globalassets/documents/online-safety/children-sending-receiving-sexual-messages.pdf
[11] https://www.theguardian.com/society/2019/dec/30/thousands-of-children-under-14-have-been-investigated-by-police-for-sexting
[12] https://www.facebook.com/notes/mark-zuckerberg/a-privacy-focused-vision-for-social-networking/10156700570096634/
[13] https://www.youtube.com/watch?v=9xePC0Tyeuc&feature=youtu.be&t=4547
[14] A 'backdoor key' is a security vulnerability built into the code of an end-to-end encrypted server. The key can be exploited by the platform to read the contents of a message. The UK issued a joint statement with six other governments in October 2020, calling for tech firms to introduce backdoor keys on E2EE messaging servers to aid law enforcement investigations.

A simple overview of end-to-end encryption

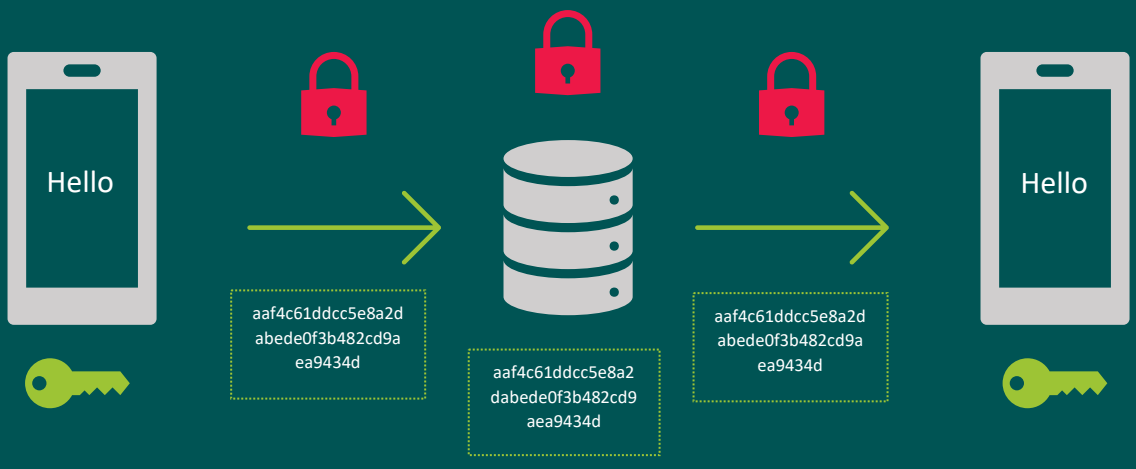
All messaging servers encrypt – or scramble into unreadable code – messages as they transfer between users' accounts. This ensures that a third party cannot intercept and read messages as they are transferred between sender, server and recipient.

Standard encryption



On a standard encrypted platform, the server acts as an intermediary. The server, e.g. Facebook Messenger, holds a 'key' to de-encrypt, or unlock the contents of a scrambled message, which it then re-encrypts to send on to the intended recipient.

End-to-end encryption



An end-to-end encrypted service operates on the same basic premise. However, now the server, e.g. Apple iMessage, does not hold a key with which to unlock the message. The message remains scrambled and unreadable until the point at which it reaches the recipient. The recipient's device holds the only copy of the key which is needed to unlock and read the contents of the message.

The CCO is concerned that a major shift to end-to-end encryption, if implemented rashly and irresponsibly, could provide a convenient loophole for tech companies to side-step their duty of care to young and vulnerable users. If the contents of a message are unreadable, it follows that it would be very difficult for the government to hold a platform to account for its contents. In short, there is a risk that the moves towards further regulation of the sector to protect people from online harms, might be acting as a perverse incentive for tech companies to actively turn a blind eye to the most grievous crimes against children committed on mainstream messaging platforms.

This concern is shared by child protection experts around the world: earlier this year the NSPCC, along with child protection organisations in 102 countries, wrote to Facebook calling for child safety measures to be embedded in encryption design. [15] Governments, including the UK Home Office have also called for major platforms to slow down plans to implement end-to-end encrypt without child protection safeguards. [16, 17]

In March 2020 a coalition of tech firms [18] – including Facebook, Snap, Apple, Google and Twitter – signed up to a series of principles put forth by five governments, including the UK, to counter online sexual exploitation and abuse. [19] While this is a step forward, the principles remain voluntary and steer clear of the live debate around end-to-end encryption and its impact on companies' abilities to continue combatting abuse.

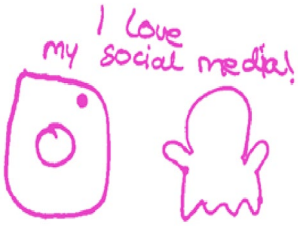
[15] https://www.nspcc.org.uk/globalassets/documents/policy/letter-to-mark-zuckerberg-february-2020.pdf
[16] https://www.gov.uk/government/publications/open-letter-to-mark-zuckerberg
[17] https://www.gov.uk/government/publications/international-statement-end-to-end-encryption-and-public-safety
[18] https://www.technologycoalition.org/
[19] https://www.gov.uk/government/publications/voluntary-principles-to-counter-online-child-sexual-exploitation-and-abuse

How are children communicating online?

Most children are using messaging platforms

Instant messaging is now an important part of most children's lives. Our survey found that **9 in 10 children in England (aged 8-17) use a messaging app or website** [20]. Usage of messaging platforms increases sharply with age, from 70% of children aged 8-10, 91% of children aged 11-13, to 97% of teenagers aged 14-17.

Children tell us they value the immediate connection with friends and family, and view messaging platforms as a positive part of childhood in 2020. [21]



8-year-old girl's response to "*What makes you happy?*" – CCO 2020/21 business plan

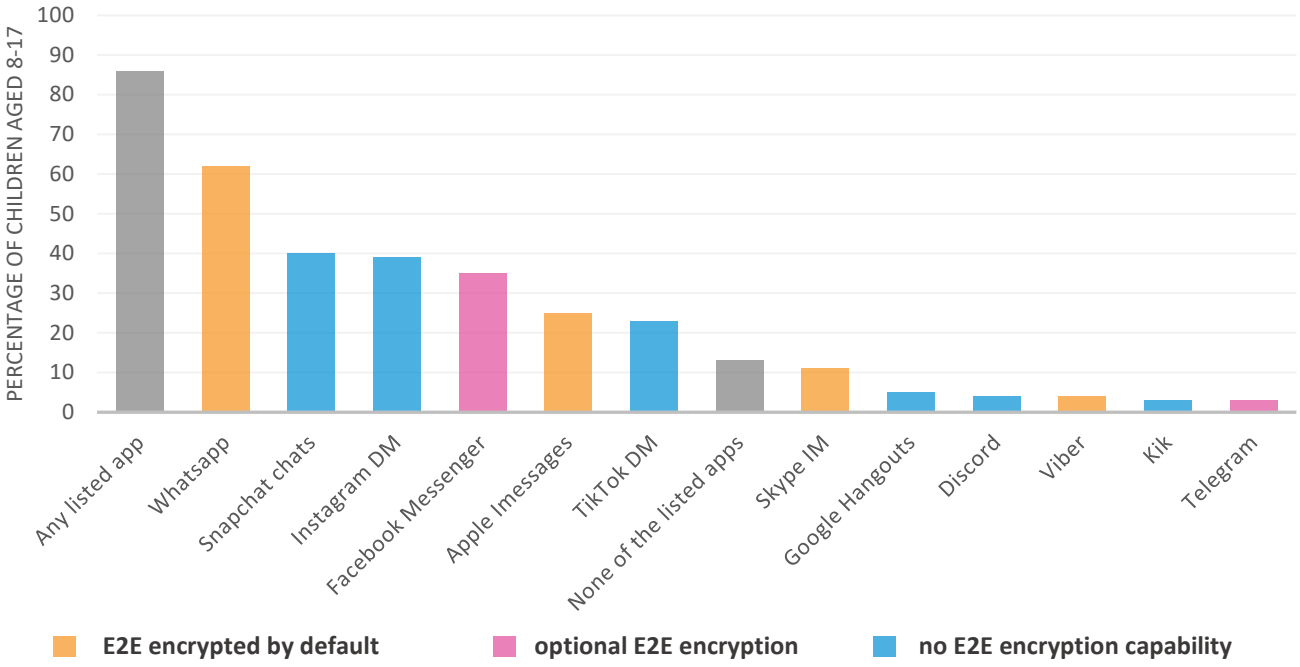


***Figure 2**: Use of different messaging apps/sites by children aged 8-17 (%)*

The most popular messaging apps are already (or are soon to be) end-to-end encrypted by default

The five most popular messaging services used by children – Whatsapp, Snapchat chats, Instagram DM, Facebook Messenger and Apple iMessage – are already fully end-to-end encrypted by default, have made public their plans to become so in the near future, or have indicated that they are considering the possibility.

Our survey found that Whatsapp – an end-to-end encrypted service owned by Facebook – is the most

[20] Survey of 2,000 carried out by Childwise on behalf of the Children's Commissioner for England from 14 – 26 March 2020. Percentages have been weighted to produce nationally representative estimates. The percentages shown have not been tested for statistical significance, differences between groups should be treated as descriptive or indicative only.

[21] https://www.childrenscommissioner.gov.uk/wp-content/uploads/2020/03/cco-childhood-in-2020.pdf

popular messaging app among all age groups, used by 62% of children surveyed.

Chat services attached to large social media sites, such as Snapchat, Instagram, Facebook and TikTok, are also popular, particularly among teenagers. All have age limits which children routinely ignore.

Most children are using platforms to send instant messages to friends and family

The overwhelming majority of children (89%) reported that they use messaging platforms to send messages to friends and family. The next most common activities reported were: video calling friends and family (50%); sharing pictures and videos with friends and family (38%); playing games (23%); and sharing location with friends and family (20%).

Girls, particularly teenage girls, are more likely to share pictures and videos of themselves

The number of children taking pictures or videos of themselves to share with friends and family rises with age, from 30% of 8-10s to 43% of 14-17-year-olds. Teenage girls are the most likely to share images and videos of themselves with friends and family: 50% of 14-17-year-old girls reported sending pictures and videos of themselves to people they know, compared to 35% of boys of the same age.

Boys and girls use messaging platforms in different ways

Image-sharing is not the only way in which girls' and boys' use of messaging platforms differ. Girls are increasingly more likely to share their location with friends and family as they reach their teen years. Girls are also more likely to report communicating with strangers as they get older, from 5% of 8-10 year-old girls to 12% of those aged 14-17; whereas among boys contact with strangers on messaging sites declines from 14% to 7%.

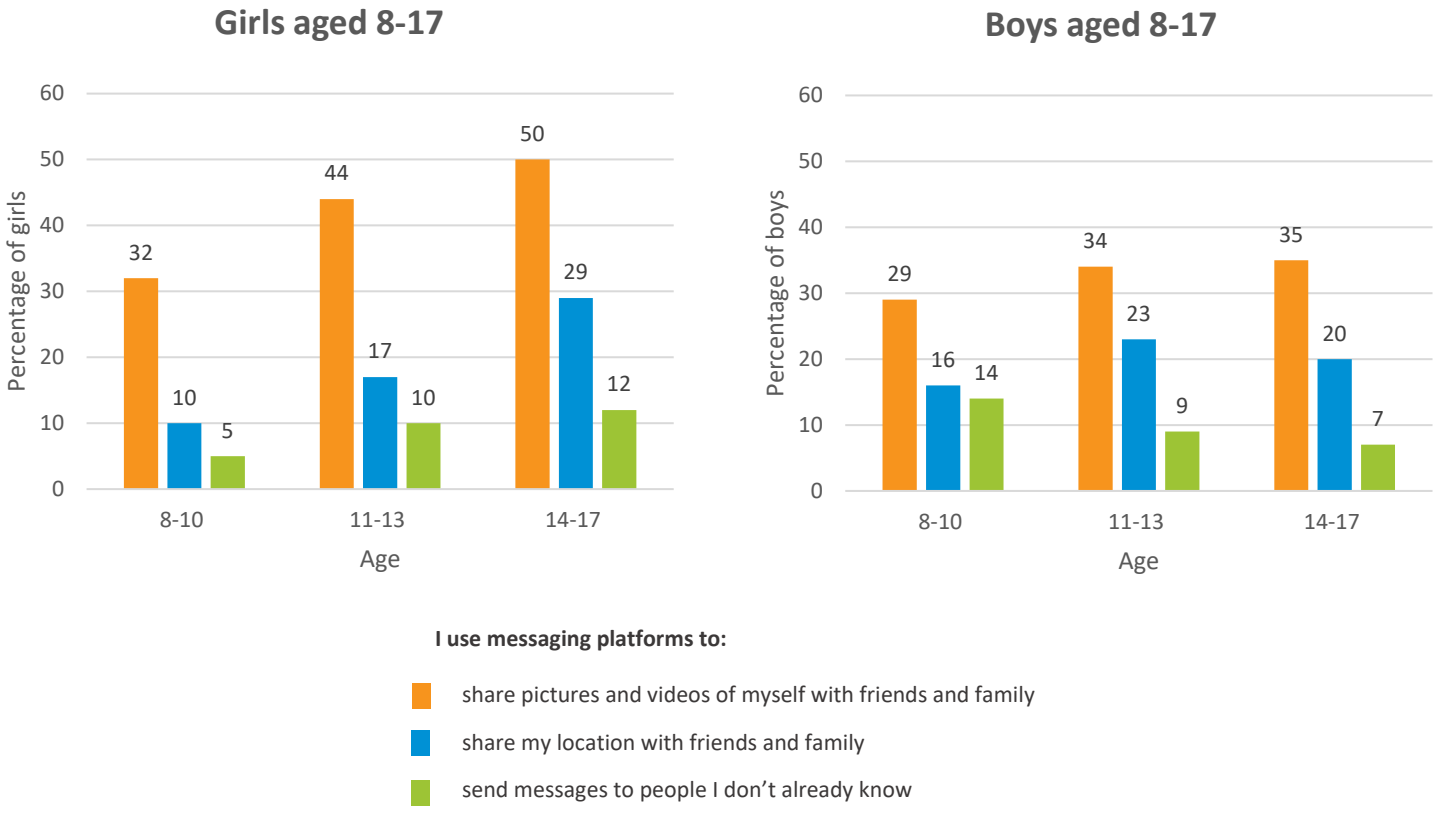


***Figure 3**: Use of messaging apps/sites by children, girls (L) and boys (R), aged 8-17 (%)*

A significant minority of children say they have contact with strangers on messaging platforms

The proportion of children who report contact with strangers on messaging platforms is small but not insignificant: around one in ten children are messaging people they don't know. The percentage of children does not increase with age – suggesting that a worrying number of younger children (aged as young as 8, 9 and 10) have contact with strangers on messaging sites.

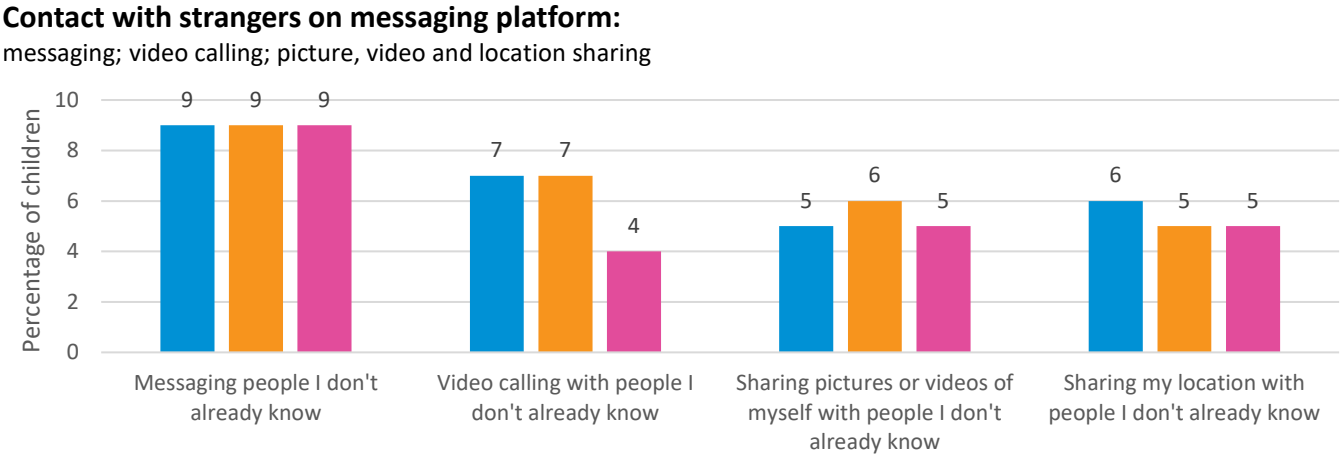


***Figure 4**: Use of messaging apps/sites by age*

Many children are receiving distressing content

Over a third (38%) of children aged 8-17 report having received content on a messaging platform which worried them or made them feel uncomfortable in the four weeks prior to our survey. This material was most likely to have been sent by friends (16%) but children also reported receiving content that made them feel uncomfortable from acquaintances (12%), family members (9%) and strangers (10%) in the previous four weeks.

The source of distressing content varies by age and by gender

Younger boys are more likely to have received worrying material from friends – 25% of 8-10-year-old boys compared to 15% of girls of the same age said that a friend had sent them something that made them feel uncomfortable in the last four weeks. The percentage of boys receiving distressing content from friends decreases with age. Older boys may be receiving less distressing content, or perhaps they are becoming desensitised and acquainted with worrying material as they reach their teenage years.

Conversely, **teenage girls are more likely than boys to have received something that made them uncomfortable from a stranger**: 16% of 14-17-year-old girls received a worrying message, picture or video from a stranger compared to 9% of boys of the same age. The proportion of girls stating they had received distressing content from strangers increased above the age of 14.

Underage use of online platforms

A key reason that children are experiencing harm online is that they are present on platforms that they are not old enough to be using, and platforms are not doing enough to stop this from happening. Many platforms have minimum age requirements in their terms and conditions for use, but do not have effective mechanisms to enforce these. Most online platforms require users to enter a date of birth in order to verify their age and get access to a site. This mechanism is easily circumvented by children, many of whom simply enter a fake date of birth.

We asked children whether they had said they were older than they were in order to sign up to an online messaging service – over half of 11-13 year olds and over a third of 8-10 year olds reported doing so.

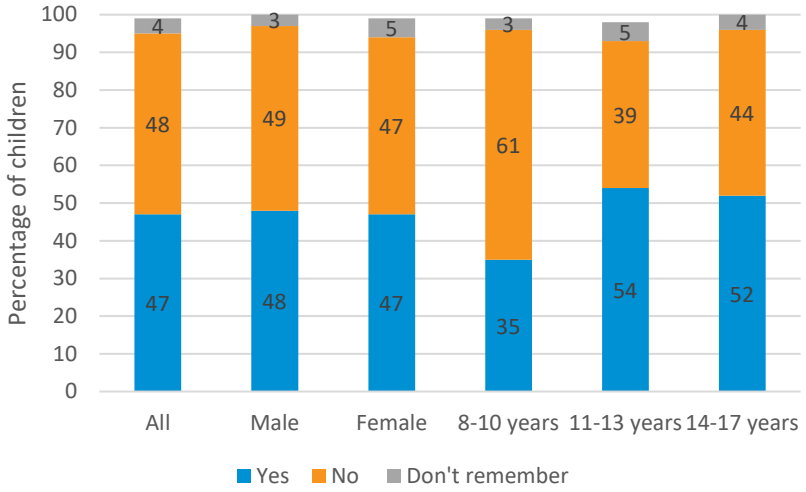


***Figure 5**: "Have you ever said you are older than you actually are to sign up for one of these social media or messaging apps/sites?" by gender and age group (%)*

We then asked children which platforms they had used in the four weeks prior to the survey, and compared this to the child's age and the minimum age of the platform. We found that nine in ten 12 year olds and six in ten 8 year olds said that they had used a messaging app or site with a minimum age of at least 13. Whatsapp was not only the most widely used platform across our sample, but it was also rife with underage users: 37% of 8 year olds and 72% of 12 year olds had used the platform in the four weeks prior to the survey – despite the fact that Whatsapp has a minimum age of 16.

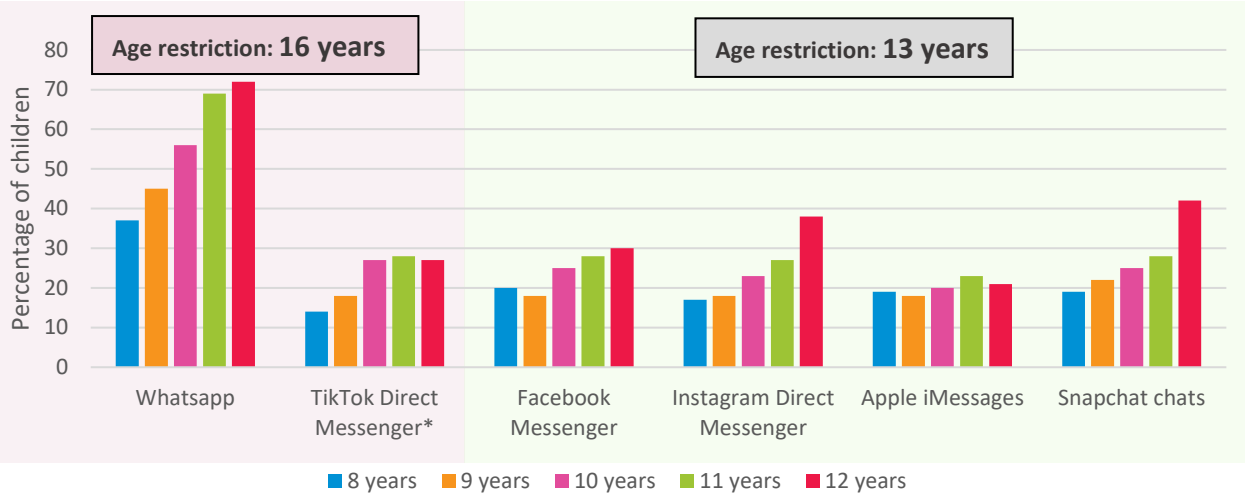


***Figure 6**: Use of messaging apps/sites by children below age restriction (ages 8-12)*

*The minimum age to sign up for a TikTok profile is 13, but under-16s are not allowed to send or receive private messages on the platform

Figure 6 shows our findings in relation to underage use for the most popular messaging platforms in our survey. A full table for all the messaging platforms we asked about can be found in Appendix 1.

Some children, therefore, are using services where they don't meet the minimum age, but deny having lied about their age to gain access. For some this could be because they weren't aware of the minimum age of the platforms they were using, or because they accessed the service on someone else's account (e.g. a parent's). But it is also likely that some deliberately lied about their age in order to gain access and did not want to admit to doing so.

In any case, these findings show how easily and frequently children get around the age restrictions of many key messaging platforms – knowingly or otherwise. This is not a problem restricted to messaging platforms, but extends to many other service types too, including social media and online gaming sites. For example, Ofcom report that 44% of 8-12 year olds in the UK use TikTok [22] (minimum age for platform 13, minimum age for private messaging function 16) and the CCO's report 'Gaming the System' found that popular game Fortnite was being played by many children under its minimum age of 12. [23]

[22] https://www.ofcom.org.uk/__data/assets/pdf_file/0027/196407/online-nation-2020-report.pdf
[23] https://www.childrenscommissioner.gov.uk/report/gaming-the-system/

Threats posed by end-to-end encryption

Our polling shows that almost 9 in 10 children (89%) use messaging platforms to chat with friends and family, and to share pictures and videos. For many children, instant messaging is a healthy and fun way to stay in touch with friends and loved ones.

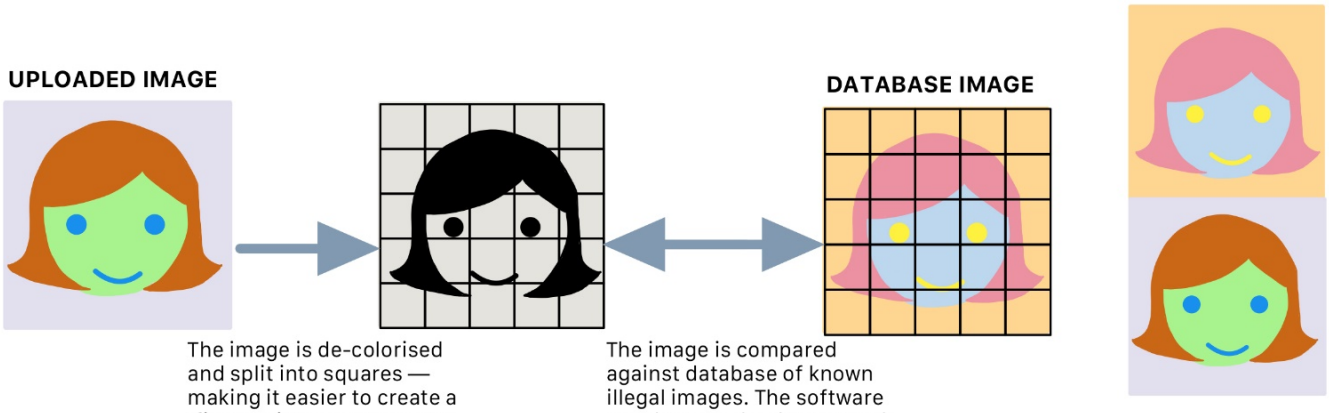
However, our polling also demonstrates that messaging platforms can pose risks to young users, both in terms of the content they are sending and the distressing material they are receiving. The survey did not ask children about the specific sorts of images and videos they were sharing, nor the nature of worrying material that they had received – but it is clear that there is risk inherent in both.

The Commissioner is concerned that end-to-end encryption, if implemented without robust child protection safeguards, could make it much harder for platforms to detect grooming, scan for child abuse materials and share reports with law enforcement agencies.

The capacity to scan for child sexual abuse material (CSAM) currently exists but is underused

Despite the fact that some 20 billion messages are transferred across major platforms like Facebook Messenger every month, [24] tech companies are currently able to employ photo- and video-scanning software to actively seek out child abuse perpetrated on their platforms.

PhotoDNA is software which automatically recognises known photos of child abuse, even if the image is altered or distorted. The algorithm works by creating a 'fingerprint' of a photo, cross-comparing it against a database of known CSAM and flagging to the server if the 'fingerprint' happens to be a close match of a known illegal image.



PhotoDNA was initially developed with the capacity only to recognise still images of child abuse. However the software has now been extended to enable the recognition of videos which have been labelled as CSAM. [25] Google has developed a similar tool – CSAI Match – which is used to scan for videos of child abuse on platforms including YouTube, Tumblr and Adobe. [26]

[24] https://messengernews.fb.com/2019/04/30/messenger-at-f8-2019-over-20b-messages-exchanged-between-people-and-businesses-every-month/
[25] https://news.microsoft.com/on-the-issues/2018/09/12/how-photodna-for-video-is-being-used-to-fight-online-child-exploitation/
[26] https://protectingchildren.google/intl/en_uk/

How the technology works

When a known piece of CSAM is detected by a platform, a report is automatically sent to national clearing houses for child abuse imagery – such as the Internet Watch Foundation (IWF) in the UK, or the National Centre for Missing and Exploited Children (NCMEC) in the US. In 2019, over 70 million images and videos were flagged and sent to NCMEC for investigation. The vast majority of these alerts, over 60 million or 85% of the total, were sent from Facebook. The Facebook Messenger app was responsible for nearly 80% of the company's total reporting last year. [27]

CSAM scanning happens automatically, in a matter of micro-seconds. Crucially, a user's privacy is preserved, unless they are sharing known images of child sexual abuse. CSA scanning software works on the same principle as an email spam filter or cybersecurity tools which detect viruses and malware: not by reading the contents of messages, but by detecting signs of suspicious and potentially harmful material and flagging this to the platform and user.

Implementing the technology in the real world

However, there is currently no legislation or statutory duty which *requires* platforms to scan for photos or videos of children being abused; those that do, do so voluntarily. While many industry experts are convinced that all content-sharing platforms are 'infested' with illegal images and videos of child abuse, figures released last year by NCMEC possible differences in detection and reporting processes. [28]

The move to end-to-end encryption could threaten platforms' abilities to scan for CSAM

Currently, active scanning for child sexual abuse material (CSAM) and grooming only occurs on servers which are not end-to-end encrypted. The software developer behind PhotoDNA, Professor Hany Farid, suggests that the search tool can be implemented at the moment an image is sent – rather than mid-transit – and is perfectly compatible with a secure, end-to-end encrypted messaging service. [29] However, it remains to be seen whether companies will be willing or able to implement photo- and video-scanning software like PhotoDNA and PhotoDNA for Video on end-to-end encrypted messaging platforms.

Facebook has not made public any intention to continue using PhotoDNA, or similar CSAM scanning software, on any content which is end-to-end encrypted once this is rolled out by default. The company says it will continue to use PhotoDNA software on unencrypted material, and on content flagged by user reports – but this is not where the vast majority of CSAM is detected. Facebook also says it has developed new tools, which are compatible with an end-to-end encrypted platform, to detect and combat grooming and sexual exploitation. These tools include gathering users' metadata (on actions such as the volume and frequency with which users send messages and friendship requests) and 'safety tips' to flag suspicious activity, in an effort to prevent unwanted contact on its platforms. [30] Facebook argue that these tools will help to prevent harmful interactions between predators and children on their social

[27] https://www.nytimes.com/2020/02/07/us/online-child-sexual-abuse.html – the very high proportion of CSAM detected on Facebook and Facebook Messenger may be a reflection of the company's substantial size (Facebook continues to hold its position as the largest social media site in the world, with 2.2 billion users globally). The company also takes an aggressive approach to detecting and removing CSAM from Facebook and Facebook Messenger.

[28] https://www.missingkids.org/content/dam/missingkids/gethelp/2019-reports-by-esp.pdf

[29] https://5rightsfoundation.com/uploads/5rights-briefing-on-e2e-encryption--csam.pdf

[30] https://messengernews.fb.com/2020/05/21/preventing-unwanted-contacts-and-scams-in-messenger/

media platforms, Instagram and Facebook. Both have recently been exposed as rife with grooming offences. [31]

However, child protection experts remain unconvinced that these kind of measures outlined by Facebook and other key industry players will be effective in preventing the activity of predators, nor in combatting the growing circulation of child sexual abuse material. Law enforcement agencies warn that Facebook's current plans effectively blind themselves and law enforcement to the most egregious forms of online child sexual abuse, which continues to grow in scope and severity. [32]

Snap have told us that they intend to ensure that any further implementation of end-to-end encryption on Snapchat will not impede the use of critical tools like PhotoDNA. However, the nature of the ephemeral messages shared on Snapchat mean it is less likely to be used by paedophiles as a platform to share images and videos of abuse. It remains unclear whether Snap are developing similar tools to combat grooming, coercion and sexual exploitation on Snapchat – which is a far more pressing issue on this platform. Snapchat has recently been described as a 'haven' for predators looking to make contact with children online and procure new, self-generated abuse content. [33]

[31] https://www.nspcc.org.uk/about-us/news-opinion/2020/instagram-grooming-crimes-children-lockdown/
[32] https://www.iwf.org.uk/sites/default/files/reports/2020-04/IWF_Annual_Report_2020_Low-res-Digital_AW_6mb.pdf p.48
[33] https://www.thetimes.co.uk/edition/news/predators-flock-to-snapchat-a-safe-space-for-child-abuse-9ztwf0tf9

Government plans to tackle online harms

It is no secret that many children are using online platforms they are not old enough to be using– nor is the fact that many children come to harm on popular sites. Yet despite these problems being discussed for several years, the online world remains a hostile environment for children, with key industry players doing little to protect their most vulnerable users. The move towards greater end-to-end encryption without appropriate protections in place is just further evidence that children's needs are an afterthought for many companies, rather than a primary consideration.

In 2017 the CCO published a report making the case for the UN Committee on the Rights of the Child to Comment on Children's Rights and the Digital Environment [34] – a piece of work that is now in train. But more needs to be done to ensure that children's rights are observed in practice. The failure of self-regulation has rightly led the Government to make firm commitments in recent years to intervene, raise standards and do what is needed to protect children, with some notable successes.

The Age Appropriate Design Code

The Age Appropriate Design Code, introduced into law through the Data Protection Act 2018, marks a significant step forward for children's rights to privacy online. Online platforms which are "likely to be accessed by children" will be required to meet 15 standards under the code from the end of the transition phase in September 2021. This is in order to protect children's data privacy, and platforms failing to do so will more than likely be in breach of the General Data Protection Regulation (GDPR), [35] risking fines of up to £18 million or 4% of global turnover.

The Online Harms White Paper

In addition, the Online Harms White Paper sets out the Government's proposal to create a new regulatory framework for online platforms which enable users to interact and share material. Central to this framework is the proposal of a ground-breaking new statutory duty of care, with compliance enforced by an independent regulator. Any online service found to be in breach of their duty of care would face sanctions. A range of online harms would be in scope of the regulation, including both legal and illegal harms. Boris Johnson's 2019 manifesto re-affirmed this commitment, promising to legislate to "make the UK the safest place in the world to be online – protecting children from online abuse and harms". [36]

Barriers to progress

But efforts to improve children's experiences online have also been undermined by setbacks and delays. It is now over 18 months since the publication of the Online Harms White Paper, and over 3 years since the publication of the Internet Safety Strategy green paper which preceded it. In this time the Government has released only its initial response to the White Paper consultation, with a full response promised by the end of this year. It has only committed to bringing forward legislation by the end of this Parliament, which could be as late as 2024. Bringing forward legislation is only the first step – it will take a significant amount of time for the regulator (most likely to be Ofcom) to consult and finalise the details of its regime, and then a further transition period to allow companies to prepare. This could mean many more years of children coming to harm online before any new regulation comes into effect. A child who was 8 years old when the Government's green paper was published in 2017 could be as old as 15 by the time legislation is passed, and an adult by the time the new regulation comes into effect.

[34] https://www.childrenscommissioner.gov.uk/2017/06/28/making-the-case-for-a-new-un-framework-for-childrens-digital-rights/
[35] GDPR is a regulation which requires businesses to protect the personal data and privacy of EU citizens
[36] https://assets-global.website-files.com/5da42e2cae7ebd3f8bde353c/5dda924905da587992a064ba_Conservative%202019%20Manifesto.pdf

Furthermore, the Government's initial response to the White Paper consultation leaves many key questions unanswered about how the new regulatory regime would work, which its success and effectiveness rest upon. For example, there has been no decision on what sanctions companies would face if they breach the duty of care, with the White Paper proposing everything from fines (which could be easily absorbed by the biggest platforms) up to blocking the sites from use by UK users. The Government's expectations around age verification remain unclear, as does its approach to messaging platforms. The White Paper suggests that the duty of care will not apply in the usual way to "private communications", and that requirements to scan or monitor content for tightly defined categories of illegal content will not apply to these channels. There is concern that end-to-end encrypted messaging services could be defined as "private communications" and could therefore not be subject to the duty of care in the same way as other platforms.

Moreover, the Government announced in October 2019 that it was scrapping plans to introduce age verification measures to prevent children from accessing pornography online. Under the plans, adults would have been required to prove their age in order to access commercial pornography, whether through traditional forms of ID or by buying an over-the-counter card from a shop. The Government has said that it will address the problem of children accessing pornography instead by bringing it into the scope of its forthcoming online harms legislation. While a single regulatory framework might have benefits, such as greater clarity and coherence, the ongoing delays to this legislation coupled with this move away from age verification measures mean that children are going for longer without protection and support. Meanwhile, the now well-documented harms associated with under-age pornography consumption – including negative body image, objectification, sexual aggression and heightened risk-taking – continue to proliferate among children who are freely able to access explicit content. [37]

[37] https://aifs.gov.au/publications/effects-pornography-children-and-young-people-snapshot

Policy recommendations

Online Harms

The long-awaited online harms legislation must be introduced into Parliament in 2021, at the very earliest opportunity.

It is not enough to promise legislation by the end of this Parliament. Now is the time to act to safeguard children's safety, privacy and wellbeing online. The new regulatory regime must:

- **Set a strong expectation on platforms to age verify their users.** Platforms likely to be accessed by children should be required to have robust age verification or age assurance measures in place – not only those which are deliberately targeted at children.
- **Allow strong sanctions for companies which breach the duty of care**, so that the behaviour of these platforms changes as a result. This should include GDPR-style fines, but extend to senior management liability and ISP-blocking in the most serious cases. Companies should also be required to issue notifications to their users when they are found to be in breach of the duty of care, outlining what went wrong and how they will ensure it does not happen again in the future, including in child-friendly language.
- **Bring the full range of services used by children in scope** – including social media, messaging services (end to end encrypted or not) gaming platforms, and more. The design of the regulatory standards must make sense across all of these platforms.

The new regulator needs the right funding to develop the details of the regime at pace.

Age Appropriate Design Code

The ICO should take robust action against platforms who do not conform to the requirements of the Age Appropriate Design Code when the transition ends in September 2021.

Platforms have been given a transition period of 12 months to understand and apply the standards outlined in the Code. There is no excuse if they are found to be in violation of the Code next September, and the ICO should not hesitate to use its enforcement powers in these cases.

In particular, the ICO should pay close attention to platforms where children come to harm and where there are not appropriate age verification/assurance measures in place.

Encryption

With intelligent design and co-ordinated effort from tech companies – alongside effective Government regulation – we can fight online child abuse and maintain secure messaging services.

Industry

Attempts to ensure users' privacy should not compromise children's safety. Tech firms should have to meet the four tests below **before** rolling out any new design feature.

1. **Demonstrate how the platform will ensure child safety by design.**
   Companies should be transparent about their plans to roll out new features, and allow for scrutiny by government, regulators and child protection experts before going ahead. If the platform is unable demonstrate that new software will not put younger users at heightened risk of harm, the feature should not be implemented.
2. **Do not apply end to end encryption to children's accounts, unless children will be afforded the same level of protection as before**
   End-to-end encryption has the potential to pose significant risks to children. It is critical that it is only rolled out to children's accounts on a given platform if doing so does not reduce children's safety. In the meantime, it should only be applied to adults' accounts. To this end, tech companies need to know which accounts belong to children and which belong to adults.
3. **Have mechanisms in place to proactively monitor for child sexual exploitation (CSE).**
   It is not enough to wait for children to report 'suspicious' and potentially illegal behaviour. Sex offenders use sophisticated techniques to groom children and will exploit any situation where a blind eye is turned. Platforms should routinely monitor and investigate suspicious behaviour, and continue to report any suspected exploitation to law enforcement agencies.
4. **Retain the ability to scan for child sexual abuse material (CSAM).**
   The tech industry has built an impressive arsenal of tools to routinely detect and eliminate images and videos of child abuse. However, these tools are currently under-used and it is not clear whether they would function as effectively within an end to end encrypted environment. These tools must be used more widely and must not be discarded in order to facilitate the roll out of end-to-end encrypted messaging.

Based on the detail made publicly available so far, the CCO does not believe that either Facebook or Snap's end-to-end encryption plans satisfy all four conditions yet. The onus is on these companies to demonstrate how they will continue to protect children's safety before rolling out default end-to-end encryption across their messaging platforms.

Government and regulators

Platforms which do not meet the 4 tests above should be judged to have breached their duty of care.

The Commissioner believes that these tests provide a robust framework with which to judge platforms' efforts to protect children online. The regulator should issue sanctions against those which fail the tests.

The duty of care should cover 'private communications', including those which are E2E encrypted.

In many ways, children face the greatest risk of harm on private messaging platforms – where illegal behaviour like sexual exploitation can occur beyond the supervision of parents, carers, teachers and other responsible adults. The Duty of Care should ensure that tech companies are the eyes and ears of children's safety on private messaging platforms. Law enforcement must also retain the ability to seek out criminal activity on private messaging platforms and to request evidence from tech firms where harm is detected.

Companies which do not tackle serious crimes committed on their private messaging platforms must be sanctioned by the regulator under the Duty of Care.

Schools and parents

The main responsibility for keeping children safe and happy online should lie with the platforms they use. Nevertheless, educating children about the benefits and risks of the digital world is also a priority.

Schools should begin to teach the new Relationships/Relationships and Sex education (RSE) curriculum as soon as they can.

With understandable delays due to Covid-19 closures, schools should now make it a priority to begin delivery of the RSE new curriculum. Schools should aim to equip all children with the knowledge that will help them stay safe and happy online – including how to manage their relationships with other people online, as well as recognising and reporting harmful content and contact.

Parents and carers should take an active interest in their children's online lives and keep an open line of communication about anything that is worrying their children.

The CCO has a number of resources to help parents play a positive role in their children's online lives:

- Our online toolkits for parents and children were developed in response to the Covid-19 crisis and provide a useful overview of how to help children stay safe and happy when many of us are spending more time online.
- Our Digital 5 a Day tool is a great way to begin a conversation with children about how to achieve a healthy and balanced digital diet.
- Our report *Who Knows What About Me?* includes five top tips for parents and children specifically looking at how children can manage their data online.

Appendix 1: underage use of messaging platforms in our survey

Messaging app	E2EE?	Age restriction	Age				
			8 (%)	9 (%)	10 (%)	11 (%)	12 (%)
Facebook Messenger	**Yes** – but not by default	13	20	18	25	28	30
Whatsapp	**Yes** – by default	16	37	45	56	69	72
Instagram direct messaging	**No**	13	17	18	23	27	38
Kik	**No** but retention period exceptionally short (i.e. deleted from server immediately so would be unable to share with authorities if requested)	13	2	3	1	4	4
Telegram	**Yes** – but not by default	16	3	4	3	4	2
Viber	**Yes**- by default	13	2	5	3	5	6
Skype instant messaging	**Yes** – by default	13	8	11	10	6	11
Google Hangouts	**No**	13	3	3	3	6	6
Apple iMessage	**Yes** – by default	13	19	18	20	23	21
Snapchat chats	- Snaps **yes** - Messages and group chats **no**	13	19	22	25	28	42
Discord	**No**	13	0.6	3	4	4	5
TikTok Direct Messaging	**No**	16 *	14	18	27	28	27

**The minimum age requirement to sign up for an account is 13, but access to private messaging platform is restricted to 16*

Children's COMMISSIONER

Children's Commissioner for England
Sanctuary Buildings
20 Great Smith Street
London
SW1P 3BT

Tel: 020 7783 8330
Email: info.request@childrenscommissioner.gov.uk
Visit: www.childrenscommissioner.gov.uk
Twitter: @ChildrensComm