House of Lords
House of Commons

Joint Committee on the Draft Online Safety Bill

Draft Online Safety Bill

Report of Session 2021–22

Report, together with formal minutes relating to the report

Ordered by the House of Lords to be printed on 10 December 2021

Ordered by the House of Commons to be printed on 10 December 2021

HL Paper 129
HC 609
Published on 14 December 2021
by authority of the House of Lords and House of Commons

Joint Committee on the draft Online Safety Bill

The Joint Committee on the draft Online Safety Bill was appointed by the House of Lords and the House of Commons to conduct pre-legislative scrutiny of the Government's draft Bill to establish a new regulatory framework to tackle harmful content online.

Membership

Damian Collins MP *(Conservative, Folkstone and Hythe)* (Chair)

Debbie Abrahams MP (*Labour, Oldham East and Saddleworth*)

Darren Jones MP (*Labour, Bristol Northwest*)

John Nicolson MP (*Scottish National Party, Ochil and South Perthshire*)

Dean Russell MP (*Conservative, Watford*)

Suzanne Webb MP (*Conservative, Stourbridge*)

Lord Black of Brentwood (*Conservative*)

Lord Clement Jones CBE (*Liberal Democrat*)

Lord Gilbert of Panteg (*Conservative*)

Baroness Kidron OBE (*Crossbench*)

Lord Knight of Weymouth (*Labour*)

Lord Stevenson of Balmacara (*Labour*)

Powers

The Committee had the power to send for persons, papers and records; to sit notwithstanding any adjournment of the House; to report from time to time; to appoint specialist advisers; and to adjourn from place to place.

Publication

Committee reports are published on the Committee's website at https://committees.parliament.uk/committee/534/draft-online-safety-bill-joint-committee/ and in print by Order of the House.

Evidence relating to this report is published on the inquiry publications page of the Committee's website.

Committee staff

The staff who worked on this committee were David Slater (Commons Clerk), Andrea Dowsett (Lords Clerk), Ellie Hassan (Lords Policy Analyst), Holly Woodhead (Second Lords Clerk), Zoe Hays (Committee Specialist), Hannah Stewart (Deputy Counsel), Jillian Luke (Committee Specialist), Ian Hook (Senior Executive Officer), Bruce Sinclair (Committee Operations Officer) and Monika Rubens (Committee Operations Officer), Lucy Dargahi (Senior Communication Officer).

Contacts

All correspondence should be addressed the Joint Committee on the draft Online Safety Bill Committee, Committee Office, House of Lords, London SW1A 0PW. Telephone 020 7219 0883. The Committee's email address is jconlinesafetybill@parliament.uk.

Contents

Summary

Self-regulation of online services has failed. Whilst the online world has revolutionised our lives and created many benefits, underlying systems designed to service business models based on data harvesting and microtargeted advertising shape the way we experience it. Algorithms, invisible to the public, decide what we see, hear and experience. For some service providers this means valuing the engagement of users at all costs, regardless of what holds their attention. This can result in amplifying the false over the true, the extreme over the considered, and the harmful over the benign. The human cost can be counted in mass murder in Myanmar, in intensive care beds full of unvaccinated Covid-19 patients, in insurrection at the US Capitol, and in teenagers sent down rabbit holes of content promoting self-harm, eating disorders and suicide.

This has happened because for too long the major online service providers have been allowed to regard themselves as neutral platforms which are not responsible for the content that is created and shared by their users. Yet it is these algorithms which have enabled behaviours which would be challenged by the law in the physical world to thrive on the internet. If we do nothing these problems will only get worse. Our children will pay the heaviest price. That is why the driving force behind the Online Safety Bill is the belief that these companies must be held liable for the systems they have created to make money for themselves.

The Online Safety Bill is a key step forward for democratic societies to bring accountability and responsibility to the internet. Our recommendations strengthen two core principles of responsible internet governance: that online services should be held accountable for the design and operation of their systems; and that regulation should be governed by a democratic legislature and an independent regulator—not Silicon Valley. We want the Online Safety Bill to be easy to understand for service providers and the public alike. We want it to have clear objectives, that lead into precise duties on the providers, with robust powers for the regulator to act when the platforms fail to meet those legal and regulatory requirements.

The most important thing this Bill will do, if our recommendations are accepted, is hold online services responsible for the risks created by their design and operation. To give just three examples: those which aim to maximise engagement will have to mitigate the risks of that engagement. A platform that recommends content using users' data will have to mitigate the risk it recommends dangerous content to vulnerable people. A platform that allows anonymous accounts will have to ensure those committing criminal acts can be traced in a timely way by UK law enforcement.

The criminal law relating to online communication pre-dates the age of social media and modern search engines. It needs updating. We welcome the Law Commission's recommendations to reform this. We want to see new offences on the statute book at the first opportunity for harmful, threatening and knowingly false communications, cyber-flashing, trying to induce seizures in people with photosensitive epilepsy, promoting self-harm and stirring up hatred against people on grounds of sex or gender, or disability. Service providers will be required to mitigate the risks presented by content and activity that society has deemed unacceptable, whether through the criminal law, through the Equality Act, or other established legal principles. Paid-for advertising can be used by

fraudsters and other criminals. Under our recommendations, providers will be held accountable for the risks created by adverts, like any other activity online.

Protecting children is a key objective of the draft Bill and our report. Our children have grown up with the internet and it can bring them many benefits. Too often, though, services are not designed with them in mind. We want all online services likely to be accessed by children to take proportionate steps to protect them. Extreme pornography is particularly prevalent online and far too many children encounter it—often unwittingly. Privacy-protecting age assurance technologies are part of the solution but are inadequate by themselves. They need to be accompanied by robust requirements to protect children, for example from cross-platform harm, and a mandatory Code of Practice that will set out what is expected. Age assurance, which can include age verification, should be used in a proportionate way and be subject to binding minimum standards to prevent it being used to collect unnecessary data.

If service providers fail to mitigate the risk of harm, the Bill will hold them accountable. We want Ofcom to have the powers to set minimum quality standards of risk assessment, under which service providers will be required to undertake independent audits of their systems, processes and algorithms. A radical transparency regime will empower people to take informed decisions about the online services they use. The Bill will introduce significant financial penalties for service providers that fail to comply, and we want to see criminal sanctions for executives who are grossly non-compliant in how they approach online safety.

Through our recommendations, the Bill will protect freedom of speech online. Service providers will no longer be able to ignore the abuse and hatred designed to silence women and minorities. They will be required to apply their terms and conditions consistently and transparently and, for the first time, be required to publish an accessible Online Safety Policy. Service providers will no longer be able to selectively censor without accountability. They will be told by Parliament and the Regulator what is illegal and unacceptable online, and how they should act against it. They will be required to protect speech that is vital to a democratic society—journalism, whistleblowing, political and societal debate, academic research, and more. If they fail, through our recommendations, individuals will have new rights of appeal and redress, through the service providers themselves, through an independent Ombudsman and, finally, through the civil courts.

We want this Bill to reset the relationship between citizens and online services, particularly the most risky. We should not have to rely on whistleblowers and court cases to get brief glimpses into how the online world is shaped. These recommendations offer a holistic and watertight regulatory regime that will make the sector accountable to UK citizens; they are not a pick and mix, but indivisible. We urge the Government to accept our recommendations and bring the Online Safety Bill to Parliament at the earliest opportunity.

1 Introduction

Background to the draft Bill

1. The safety of people online, particularly on social media, is one of the defining policy issues of our age. The major online services have become central to the way people around the world access news and information, do business, play games, and keep in touch with family and friends. These are highly profitable businesses, with a commercial model based on selling and targeting advertising. User data is collected and used to train their algorithms to maximise engagement and users' attention. The length of time, and the frequency with which users engage with the platforms increase their value: more time, means more advertising reaches the users, which leads to more revenue for the companies. However, actively seeking to increase engagement through personalisation also has the power to create more harmful user experiences. For example, vulnerable people are more likely to see content which will increase their vulnerabilities and the more people interact with conspiracy theories the more of them they will see. The grouping together of users with similar interests can create environments which normalise hate speech and extremism. Design features that favour spread of information over safety facilitate the targeting and amplification of abuse.

2. Despite concerns that have been repeatedly raised about these problems, the companies whose systems and processes distribute this content have been unable or unwilling to address them successfully. Whilst it is true that hate and harm existed before the internet, and still would without it, the evidence is that these systems and processes have actively made things worse. We have already seen the power of online media to undermine confidence in public health organisations during a pandemic, erode the protections of children, target people with abuse and even work to undermine democracy itself. We welcome the Government's decision to publish the draft Online Safety Bill and to open it up to pre-legislative scrutiny.

3. The draft Online Safety Bill is the result of an extensive public policy and parliamentary process, going back nearly half a decade. The draft Bill was published by the Government on 12 May 2021. It followed the Online Harms White Paper, published in April 2019 and the Government's interim (February 2020) and full (December 2020) responses to the consultation on it.[1] The White Paper itself was the result of a commitment made in the Internet Safety Strategy Green Paper, published in October 2017.[2]

4. The regulation of online platforms has been the subject of intense parliamentary scrutiny and inquiry in the UK and overseas. In many cases this has anticipated and driven Government action. In February 2019 the Commons' Digital, Culture, Media and Sport (DCMS) Committee recommended in its *Disinformation and Fake News* report that service providers should not be able to avoid liability for content identified as harmful on

1 Department for Digital, Culture, Media and Sport and the Home Office, *Online Harms White Paper: Full government response to the consultation*, CP 354, December 2020: https://www.gov.uk/government/consultations/online-harms-white-paper/outcome/online-harms-white-paper-full-government-response [accessed 12 November 2021]; Department for Digital, Culture, Media and Sport and the Home Office, *Online Harms White Paper - Initial White Paper Response*, December 2020: https://www.gov.uk/government/consultations/online-harms-white-paper/public-feedback/online-harms-white-paper-initial-consultation-response [accessed 12 November 2021]

2 Department for Digital, Culture, Media and Sport, *Internet Safety Strategy - Green Paper,* October 2017: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/650949/Internet_Safety_Strategy_green_paper.pdf [accessed 12 November 2021]

their platforms, and that a "code of ethics" and an independent regulator with statutory powers should be created to oversee this.[3] Later that year it published a further report on *Addictive and Immersive Technologies* which raised concerns about data driven online platforms which prioritise increasing user engagement with particular reference to online games, free to play games and extended reality.[4] In January and March of the same year, the House of Commons Science and Technology Committee and House of Lords Communications Committee recommended that social media service providers should have a duty of care to the people that use their platforms.[5]

5. To give just a few of the many examples of parliamentary work since then, the DCMS Committee has maintained its interest on the issue through the work of its Sub-committee on Online Harms and Disinformation. The House of Commons Petitions Committee has a long-standing interest in tackling online abuse, especially that directed against disabled people.[6] The All-Party Parliamentary Group on Social Media has examined the extent to which social media impacts young people's mental health and wellbeing.[7] The Treasury and Work and Pensions Committees have examined online fraud, whilst the Home Affairs Committee has taken evidence on racism and online harms more widely.[8] The House of Lords Democracy and Digital Technologies Committee's report *Digital Technologies and the Restoration of Trust*, published in June 2020, warned about the impact of online disinformation and misinformation, and called for electoral law and media literacy education to be brought up to date for the digital age.[9] The House of Lords Communications and Digital Committee recently published a wide-ranging report on *Freedom of Speech in the Digital Age* that looked closely at the draft Bill. The Joint Committee on Human Rights has also taken evidence on freedom of expression.[10]

6. During our inquiry, the revelations of former Facebook[11] employee Frances Haugen, in the Wall Street Journal and elsewhere, contributed to intensified international interest in regulation in this area. The US Senate has conducted hearings on the protection of

3 Digital, Culture, Media, and Sport Committee, *Disinformation and 'fake news': Final Report* (Eighth Report, Session 2017–19, HC 1791)

4 Digital, Culture, Media, and Sport Committee, *Immersive and Addictive Technologies* (Fifteenth Report, Session 2017–19, HC 1846)

5 House of Commons Science and Technology Committee, *Impact of social media and screen use on young people's health* (Fourteenth Report, Session 2017–19, HC 822); Communications Committee, *Regulating in a digital world* (2nd Report, Session 2017–19, HL Paper 299); Carnegie UK, *Internet Harm Reduction* (January 2019): https://d1ssu070pg2v9i.cloudfront.net/pex/carnegie_uk_trust/2019/01/27135118/Internet-Harm-Reduction-final.pdf [accessed 12 November 2021]

6 Petitions Committee, *Online Abuse and the Experience of Disabled People* (First Report, Session 2017–19, HC 759)

7 UK Safer Internet Centre: https://saferinternet.org.uk/appg-on-social-media

8 Letter from the Chairs of the Work and Pensions and Treasury Committees to the Prime Minister, 21 July 2021: https://committees.parliament.uk/publications/6956/documents/81066/default/; Written evidence from the Work and Pensions Select Committee (OSB0020); Oral evidence taken before the Home Affairs Committee, 20 January 2021 (Session 2020–21), https://committees.parliament.uk/oralevidence/1566/html/; letter from the acting Chair of the Home Affairs Committee, 1 December 2021, https://committees.parliament.uk/publications/8077/documents/83017/default/

9 Democracy and Digital Committee, *Digital Technology and the Resurrection of Trust* (Report, Session 2019–21, HL Paper 77)

10 Communications and Digital Committee, *Free for All? Freedom of Expression in the Digital Age* (First Report, Session 2021–22, HL Paper 54); Oral evidence taken before the Joint Committee on Human Rights, 9 October 2020, (Session 2020–21): https://committees.parliament.uk/oralevidence/1387/html/

11 Throughout this report, we use "Facebook" to refer to both the social media platform and the broader company which renamed itself during our inquiry to "Meta", as most of our evidence was heard before the renaming.

children on online platforms.[12] The European Parliament has continued to scrutinise the twin proposals of a Digital Markets Act and a Digital Services Act.[13] Parliaments across Europe have taken evidence from Ms Haugen. In awarding the Nobel Peace Prize jointly to journalists Maria Ressa, CEO of Rappler, and Dmitry Muratov, Editor-in-Chief of Novaya Gazeta, the Nobel Committee picked out their work exposing and combating disinformation and "trolling" online.[14]

7. What unites these pieces of work is a sense that self-regulation of large, online platforms has failed. Around the world, there has been a growing consensus that such platforms create a risk of harm against individuals and are taking decisions with societal impacts that should be taken by democratic governments and legislatures and by independent regulators. We talk more about this in Chapter 2.

Conduct of our inquiry

8. We were appointed on 22 July 2021 and met the following week to agree and publish a call for evidence. We received over 200 submissions of written evidence and held oral evidence hearings with over 50 witnesses. A full list of witnesses and evidence is at the end of the report. We are very grateful to everyone who contributed to our inquiry. We are also grateful to our specialist advisers: Jacqueline Hughes, Dr Charles Kriel and Dr Bertie Vidgen, for their support.

9. We are grateful to two academic institutions—the London School of Economics and Political Science (LSE) Department of Media and Communications and the Minderoo Centre for Technology & Democracy, University of Cambridge—for co-hosting roundtable discussions with us on safety by design (13 October 2021), age assurance and verification (27 October), and freedom of speech and effective regulation (3 November). We would like to thank everyone who assisted with and contributed to these events.[15]

10. Parliamentarians will be able to scrutinise the final Bill when it is introduced by the Government. At the same time, we recognised the extraordinary level of interest in the draft Bill and wanted to give an opportunity for colleagues to talk with us about their views. We are grateful to the All-Party Parliamentary Group (APPG) on Digital Responsibility and Regulation for agreeing to co-host a discussion event with us on 20 October, open to all parliamentarians. We would like to thank its secretariat for the logistical arrangements and colleagues from both Houses who attended the event.

11. The Committee also wanted to consider the draft Online Safety Bill alongside other proposed legislation with similar objectives. We undertook a short visit to Brussels on 8

12 US Senate Subcommittee on Consumer Protection, Product Safety, and Data Security, *Protecting Kids Online: Testimony from a Facebook Whistleblower* (October 5 2021): https://www.commerce.senate.gov/2021/10/protecting%20kids%20online:%20testimony%20from%20a%20facebook%20whistleblower [accessed 15 November 2021]; US Senate Subcommittee on Consumer Protection, Product Safety, and Data Security, *Protecting Kids Online: Snapchat, Tick Tock and YouTube* (October 26 2021): https://www.commerce.senate.gov/2021/10/protecting-kids-online-snapchat-tiktok-and-YouTube [accessed 15 November 2021]

13 European Commission, 'The Digital Services Act Package': https://digital-strategy.ec.europa.eu/en/policies/digital-services-act-package [accessed 15 November 2021]

14 The Nobel Prize, 'The Nobel Peace Prize 2021', 8 October 2021: https://www.nobelprize.org/prizes/peace/2021/press-release/ [accessed 15 November 2021]

15 Written evidence from Minderoo, Centre for Technology & Democracy—Safety by Design Roundtable (OSB0237); LSE Department of Media and Communications—Anonymity & Age Verification Roundtable (OSB0236); LSE, Department of Media & Communications—Freedom of Expression Roundtable (OSB0247)

November 2021 to meet with the European Commission, Alexandra Geese MEP, Interpol and others about international developments and the European Union's proposed Digital Services Act and Digital Markets Act. We would like to thank everyone who met us or helped facilitate meetings for us during the visit. On 17 November our Chair met with Christel Schaldemose MEP, rapporteur on the Digital Services Act for the inquiry of the European Parliament Committee on the Internal Market and Consumer Protection. On 22 November our Chair represented the Joint Committee in giving evidence to the French Senate Committee on European Affairs, at the invitation of Senator Catherine Morin-Desailly, as part of its inquiry on the EU Digital Services Act.

Structure of this report

12. In Chapter 2 we discuss the scale of harm being experienced online and the objectives and structure of the draft Bill. In Chapter 3 we look at the role of platform design, anonymity, and its relationship to harm, particularly societal harm. In Chapters 4 and 5 we examine the safety duties in respect of adults and children, online fraud and pornography. In Chapter 6 we discuss the scope of the draft Bill across types of providers and the exemption for paid-for advertising. In Chapter 7 we look at the draft Bill's provisions in respect of freedom of expression, journalistic content and content of democratic importance. Chapters 8, 9 and 10 focus on the role of the regulator in enforcement, ensuring transparency and redress for users. Finally, in Chapter 11, we have a brief conclusion.

2 Objectives of the Online Safety Bill

13. All the service providers we heard from were taking measures to reduce activity that creates a risk of harm and illegal activity on their platforms.[16] These measures were wide-ranging and included explicit content filters on search results;[17] manual curation of content on public-facing areas of the service;[18] user voting which affects the visibility of content;[19] and, following the introduction of the Age Appropriate Design Code, default privacy settings for children who use their platforms.[20]

14. Nevertheless, we heard that illegal and harmful activity remain prevalent online. Throughout our inquiry, we have heard about the failures of self-regulation by online service providers. Witnesses have told us that the current system of self-regulation is akin to allowing service providers to mark their own homework, and that this has made the online world more dangerous.[21] This has real-world implications—during the short timescale of our inquiry, illegal and harmful activity online has been linked to the suicide of 15 year old Frankie Thomas[22] and the kidnap, rape, and murder of Sarah Everard.[23] To give just a few examples of events that occurred in the months and years immediately preceding our inquiry:

- The Internet Watch Foundation (IWF) Annual Report 2020 reported record increases in self-generated child sexual abuse material.[24]
- 5Rights' "Pathways" research showed how the design and operation of major social media services led to children being exposed to extreme pro-suicide, eating disorder and pornographic content.[25]

16 Oral evidence taken on 28 October 2021 (Session 2021–2022), QQ 200-222, QQ 223-232, QQ 233-249

17 Written evidence from Google (OSB0175)

18 Written evidence from Snap Inc (OSB0012)

19 Written evidence from Reddit (OSB0058)

20 Written evidence from TikTok (OSB0181); examples of announcements of safety measures made following the introduction of the Code included: from Microsoft, 'Introducing Microsoft Edge Kids Mode, a safer space for your child to discover the web': https://blogs.windows.com/windowsexperience/2021/04/15/introducing-microsoft-edge-kids-mode-a-safer-space-for-your-child-to-discover-the-web/ [accessed 9 December 2021]; TikTok, 'Strengthening privacy and safety for youth on TikTok': https://newsroom.tiktok.com/en-us/strengthening-privacy-and-safety-for-youth [accessed 9 December 2021]; Instagram, 'Continuing to Make Instagram Safer for the Youngest Members of Our Community': https://about.instagram.com/blog/announcements/continuing-to-make-instagram-safer-for-the-youngest-members-of-our-community [accessed 9 December 2021]

21 Q 67; Q 14, Q 3, Q 16, Q 31, Q 59, Q 178, Q 186, Written evidence from: Centre for Countering Digital Hate (OSB0009); Compassion in Politics (OSB0050); Full Fact (OSB0056); Dr Elly Hanson (OSB0078); Association of British Insurers (OSB0079); Dame Margaret Hodge MP (OBS0201)

22 BBC News, 'Frankie Thomas: Coroner rules school failed teen who took own life': https://www.bbc.co.uk/news/uk-england-surrey-58817821 [accessed 30 November 2021]

23 BBC News, Sarah Everard: 'Gross misconduct probe into Couzens WhatsApp group': https://www.bbc.co.uk/news/uk-58760933; [accessed 15 November 2021]; Care, 'Everard Killer viewed 'brutal pornography': https://care.org.uk/news/2021/09/everard-killer-viewed-brutal-pornography [accessed 15 November 2021]

24 Internet Watch Foundation, *Face the Facts: Annual Report 2020* (2020): https://www.iwf.org.uk/sites/default/files/inline-files/PDF%20of%20IWF%20Annual%20Report%202020%20FINAL%20reduced%20file%20size.pdf [accessed 15 November 2021]

25 5Rights, *Pathways: How digital design puts children at risk (July 2021)*: https://5rightsfoundation.com/uploads/Pathways-how-digital-design-puts-children-at-risk.pdf [accessed 6 December 2021]

- 2000 abusive tweets were directed at four Black players following the England national football team's loss at the Euro 2020 final.[26]
- A record number of antisemitic incidents were reported in the UK in May-June 2021, such that the Community Security Trust termed this period "the month of hate".[27] Many of these incidents took place online.[28]
- Facebook[29] was implicated in the mass murder of Rohingya Muslims in Myanmar.[30]
- The House of Commons Department for Culture, Media, and Sport (DCMS) Committee inquiry into Disinformation and "fake news"[31] and the Intelligence and Security Committee of UK Parliament[32] both concluded that Russian agents had used social media to attempt to influence UK elections. The United States Senate Select Committee on Intelligence found similar attempts to influence the 2016 US Election.[33]
- Twitter was implicated in the 6 January 2021 riot at the US Capitol, after which Twitter permanently suspended former US President Donald Trump from the platform for violating policies around inciting violence.[34]

15. In this Chapter we give an overview of the content and activity that creates risks of harm experienced by different groups of people online. We then discuss the relationship between people's experiences of online risks, their prevalence online and the systems that underpin most large online platforms. Finally, we draw conclusions for the objectives that we believe the Bill should pursue.

Harms affecting children

16. Research by DCMS has shown that "80 per cent of six to 12 year-olds have experienced some kind of harmful content online", whilst half of 13 to 17 year-olds believe they have seen something in the last three months that constitutes illegal content.[35] Children can

26 Channel 4 News, 'Nearly 2,000 abusive tweets targeted Marcus Rashford, Jadon Sancho, Bukayo Saka and Raheem Stirling after Euro 2020 final, research shows': https://www.channel4.com/news/nearly-2000-abusive-tweets-targeted-marcus-rashford-jadon-sancho-bukayo-saka-and-raheem-sterling-after-euro-2020-final-research-shows [accessed 15 November 2021]

27 Community Security Trust, *The Month of Hate: Antisemitism and extremism during the Israel-Gaza conflict* (2021): https://cst.org.uk/data/file/4/a/The_Month_of_Hate.1626263072.pdf [accessed 15 November 2021]

28 *Ibid.*

29 Facebook renamed itself to "Meta" during our inquiry, in fact on the very that day that they gave oral evidence to us. We refer to the company as "Facebook" throughout this report as this is how they are referred to in most of the sources we cite.

30 BBC News, 'UN: Facebook has turned into a beast in Myanmar': https://www.bbc.co.uk/news/technology-43385677 [accessed 15 November 2021]

31 Digital, Culture, Media, and Sport Committee, *Disinformation and 'fake news': Final Report* (Eighth Report, Session 2017–19, HC 1791)

32 Intelligence and Security Committee, *Russia* (Report, Session 2021–22, HC 632)

33 United States Senate, Select Committee on Intelligence, *Russian Active Measures Campaigns and Interference in the 2016 US Election, Volume 5: Counter Intelligence Threats and Vulnerabilities* (2020): https://www.intelligence.senate.gov/sites/default/files/documents/report_volume5.pdf [accessed 15 November 2021]

34 Twitter 'Permanent Suspension of @realDonaldTrump': https://blog.twitter.com/en_us/topics/company/2020/suspension [accessed 15 November 2021]

35 Q 54

be vulnerable to a wide range of online harms.[36] Izzy Wick, Director of Policy at 5Rights, told us:

> "We know from speaking with children and young people that the harms they experience online are extensive and wide ranging. They can be extreme, from exposure to self-harm and suicide content, violent sexual pornography and unsolicited contact with adults they do not know, right the way through to more insidious harms that might build up over time."[37]

17. The National Society for the Prevention of Cruelty to Children (NSPCC) reported that 10,391 child sex crimes were recorded by police forces across the UK for 2019/20, an increase of 16 per cent.[38] Since 2017/18, Sexual Communication with a Child offences have increased by 70 per cent reaching a record high of 5,441 recorded crimes between April 2020 and March 2021. Three quarters of these offences involved the use of Instagram, WhatsApp, Facebook Messenger, and Snapchat.[39]

18. Intentional access and accidental exposure to pornography is increasing among children. The Office of the Children's Commissioner told us that over half of 11–13-year-olds have seen pornography online.[40] Witnesses explained that pornography can distort children's understanding of healthy relationships, sex, and consent by, for example, normalising violence during sexual activity.[41] It has also been linked to addiction.[42]

19. Ian Russell, founder of the Molly Rose Foundation, told us that, in 26 per cent of cases where young people present to hospital with self-harm injuries and suicide attempts, those young people have accessed related content online.[43] The Samaritans reported that children as young as 12 have accessed suicide and self-harm material online.[44] We have heard that, while children and young people are particularly at risk, adults can also be led to suicide and self-harm as a consequence of online content and activity.[45]

20. We heard that Ofsted's review of sexual abuse in schools and colleges found 88 per cent of girls and 49 per cent of boys surveyed said that being sent pictures that they did not want to see happened "a lot" or "sometimes".[46] We also heard that children feel pressured by what they see online. This makes them feel insecure about their body image and can have significant impacts on their health, confidence, and self-esteem.[47] Frances Haugen noted: "When kids describe their usage of Instagram, Facebook's own research describes

36 Written evidence from Parent Kind (OSB0207)
37 Q 54
38 NSPCC, 'Police record over 10,000 online sex crimes in a year for the first time': https://www.nspcc.org.uk/about-us/news-opinion/2020/2020–09-03-cybercrimes-during-lockdown/ [accessed 15 November 2021]
39 NSPCC, 'Record high number of reported grooming crimes lead to calls for stronger online safety legislation': https://www.nspcc.org.uk/about-us/news-opinion/2021/online-grooming-record-high/ [accessed 15 November 2021]
40 Written evidence from The Office of the Children's Commissioner (OSB0019)
41 Written evidence from Barnardo's (OSB0017); The Office of The Children's Commissioner (OSB0019); Care (OSB0085)
42 Written evidence from Premier Christian Communications Ltd (OSB0093); COST Action - European Network for Problematic Usage of the Internet (OSB0038); CEASE (Centre to End All Sexual Exploitation) (OSB0104); Dignify (OSB0196)
43 Q 66
44 Written evidence from The Samaritans (OSB0182)
45 Written evidence from SWGfL (OSB0054)
46 Ofsted, *Review of sexual abuse in schools and colleges* (June 2021): https://www.gov.uk/government/publications/review-of-sexual-abuse-in-schools-and-colleges/review-of-sexual-abuse-in-schools-and-colleges [accessed 15 November 2021]
47 Written evidence from Girlguiding (OSB0081)

it as an addict's narrative. The kids say, 'This makes me unhappy. I feel like I don't have the ability to control my usage of it. And I feel that if I left, I'd be ostracized.'"[48]

Harms affecting adults

21. In their pilot Online Harms Survey, Ofcom found that three quarters of adult respondents reported having been exposed to at least one incidence of content or activity that creates a risk of harm in the previous month.[49] A number of individuals, online dating services, LGBTQ+ and disability rights groups, and campaigners against racism and antisemitism gave details and statistics relating to this imbalance and we discuss them more below.[50] Adults can be harmed online in a range of different ways,[51] including by fraud and scams (discussed in Chapter 4).[52]

Racist abuse

22. In many cases, the harm that these individuals face is direct abuse exacerbated or amplified by system design. In professional football, for example, an analysis by Signify funded by the Professional Football Association found that there was a 48 per cent increase in racist online abuse in the 2020–21 football season, with racist abuse peaking in May 2021 (excluding the Euro 2020 final).[53] Rio Ferdinand, former professional footballer, told us about his experiences of receiving racist abuse online. He said that experiencing racist abuse online can affect mental health and self-esteem and said it can have severe impacts on an individual's friends and family. He had personal experience of family members "disintegrating" because of online abuse being targeted at him.[54] We were told that the prevalence of racist abuse directed at football players is so great that the Football Association (FA) have had to provide guidance to their players on how to filter it from their social media feeds.[55]

23. We were very aware that the experiences of such high-profile people reflect much wider patterns of abuse and harm. Imran Ahmed, CEO and Founder of the Center for Countering Digital Hate (CCDH), told us:

> "When it comes to racism against footballers, the point that I have made to their representatives and to others is that the abuse of Marcus Rashford matters not because he is a wealthy footballer, but because if they can call Marcus Rashford the N-word, imagine what they would call me or my mum or anyone else from a minority, a woman, a gay person, anyone else."[56]

48 Q 166
49 Ofcom, 'Online Nation 2021 Report': https://www.ofcom.org.uk/research-and-data/internet-and-on-demand-research/online-nation [accessed 30 November 2021]
50 Written evidence from: Glitch (OSB0097); Centenary Action Group, Glitch, Antisemitism Policy Trust, Stonewall, Women's Aid, Compassion in Politics, End Violence Against Women Coalition, Imkaan, Inclusion London, The Traveller Movement, Stonewall (OSB0047); Antisemitism Policy Trust (OSB0005); Mencap (OSB0075); and Royal Mencap Society oral evidence 13 September QQ 52–68 and Dame Margaret Hodge (Member of Parliament for Barking and Dagenham at House of Commons) (OSB0201)
51 The five most prevalent types of harms reported by adult users In Ofcom's pilot Online Harms Survey were: spam emails, scams/fraud/phishing, misinformation, content encouraging gambling, and "alternative viewpoints"
52 Q 110; Written evidence from: UK Finance (OSB0088); Match Group (OSB0053); Glitch (OSB0097)
53 Professional Footballers' Association, 'Online Abuse': (2021), https://www.thepfa.com/news/2021/8/4/online-abuse-ai-research-study-season-2020–21 [accessed 16 November 2021]
54 Q 22
55 Q 291
56 Q 4

Abuse against LGBTQ+ people

24. When asked by Stonewall about their experiences of online abuse, one in ten LGBTQ+ people had experienced online abuse directed specifically at them within the preceding month.[57] We heard about the serious real-world impacts that online harms can have for LGBTQ+ people who, for example, have been "outed", resulting in the loss of their homes and jobs.[58] The LGBT Foundation told us that LGBTQ+ people are also at risk of being harmed by the actions of platforms themselves, with LGBTQ+ content being erroneously blocked or removed at greater rates than other types of content.[59]

Misogynistic abuse and violence against women and girls

25. Women are disproportionately affected by online abuse and harassment.[60] They are 27 times more likely to be harassed online than men.[61] 36 per cent of women report having been a victim of online abuse and harassment, with this rising to 62 per cent in women aged 18–34.[62] Abuse and harassment are not only directed towards adults: in 2020–21, half of 11–16 year old girls experienced hate speech online and a quarter were harassed or threatened.[63] Nina Jankowicz, Author and Global Fellow at the Wilson Center, told us:

> "Being a woman online is an inherently dangerous act. That is the long and short of it. It does not matter what you do. You are opening yourself up to criticism from every angle … Many women are changing what they write, what they speak about, what careers they choose to pursue because of that understanding that it is part and parcel of existing as a woman on the internet."[64]

26. Violence against women and girls (VAWG) "is increasingly perpetrated online" and online VAWG "should be understood as part of a continuum of abuse which is often taking place offline too."[65] Professor Clare McGlynn QC, Durham Law School, described an "epidemic of online violence against women and girls".[66] Online VAWG "includes but is not limited to, intimate image abuse, online harassment, the sending of unsolicited explicit images, coercive 'sexting', and the creation and sharing of 'deepfake' pornography."[67]

27. Cyberflashing—the unsolicited sending of images of genitalia[68] is a particularly prevalent form of online VAWG. 76 per cent of girls aged 12–18 and 41 per cent of all women reported having been sent unsolicited penis images. Regardless of the intention(s) behind it, cyberflashing can violate, humiliate, and frighten victims, and limit women's participation in online spaces.[69] The use of deepfake pornography in online VAWG is also

57 Stonewall, *LGBT Hate Crime in Britain: Hate and Discrimination (2017):* https://www.stonewall.org.uk/system/files/lgbt_in_britain_hate_crime.pdf [accessed 16 November 2021]
58 Q 38
59 Written evidence submitted by the LGBT Foundation (OSB0045); LGBT Foundation (OSB0046)
60 Written evidence from Dr Kim Barker and Dr Olga Jurasz (OSB0071)
61 Written evidence from Glitch (OSB0097)
62 Written evidence from Refuge (OSB0084)
63 Written evidence from Girlguiding (OSB0081)
64 Q 55
65 Written evidence from Centenary Action Group (OSB0047)
66 Q 69
67 Written evidence from Centenary Action Group (OSB0047); Refuge (OSB0084)
68 The Law Commission, 'Modernising Communications Offences: A Final Report', Law Com No 399, HC 547, July 2021: https://www.lawcom.gov.uk/project/reform-of-the-communications-offences [accessed 22 November 2021]
69 Written evidence from Professor Clare McGlynn (OSB0014)

becoming increasingly prevalent and is of great concern, having been recently debated in the House of Commons on 2nd December 2021.[70]

Religious hate and antisemitism

28. Antisemitism online is a cause of great concern,[71] comprising approximately 40 per cent of all antisemitic incidents recorded in the UK.[72] The Community Security Trust recorded 355 incidents of online antisemitism in the first six months of 2021, primarily through Twitter (35 per cent) and instant messaging services (22 per cent).[73] Danny Stone MBE, Director of the Antisemitism Policy Trust, told us about the impacts of antisemitism online:

> "There are a range of impacts. I do not post pictures of my children online often, because … there is a chance that someone will try to hurt my children … That is an individual impact.
>
> … There was a video on BitChute about the Antisemitism Policy Trust, my organisation. That has impacts on my board and what they consider about their own safety and what that means. …
>
> Also, on Jews in public life, Luciana Berger was in this House and faced an onslaught of antisemitic abuse. …
>
> There are all these impacts. There are many different impacts."[74]

29. Hate crime offences against Muslims constituted 45 per cent of recorded religious hate crimes from 2020–21,[75] with reports of online Islamophobia rising by 40 per cent during the first UK COVID-19 lockdown.[76] Islamophobic online material has real consequences—the attackers in both the Finsbury Park Mosque attack in 2017 and the 2019 Christchurch Mosque attack were thought to have been at least in part radicalised online, with the Finsbury Park Mosque attacker said to have become "obsessed" with Muslims.[77] Reset told us that, currently, "widely debunked far-right conspiracy theories about Islam run rife on social media sites/blogs", ranging from "claims of 'No Go Zones' in Western nations which are run by Sharia Law and bar non-Muslims and police" to claims about "a plot by Islamic nations to take over Europe to create 'Eurabia'".[78]

70 HC Deb, 2 December 2021, col 1154–1162

71 Written evidence from The Antisemitism Policy Trust (OSB0005)

72 CST, 'Antisemitic Incidents Report 2019': https://cst.org.uk/news/blog/2020/02/06/antisemitic-incidents-report-2019 [accessed 22 November 2021]; Written evidence from the Board of Deputies of British Jews (OSB0043)

73 Community Security Trust, *Antisemitic incidents January-June 2021* (2021): https://cst.org.uk/data/file/f/c/Incidents%20Report%20Jan-Jun%202021.1627901074.pdf [accessed 15 November 2021]

74 Q 38

75 Home Office, *Official Statistics: Hate Crime, England and Wales 2020 to 2021* (October 2021): https://www.gov.uk/government/statistics/hate-crime-england-and-wales-2020-to-2021/hate-crime-england-and-wales-2020-to-2021 [accessed 9 December 2021]

76 Newsweek, 'Muslims Falsely Blamed for COVID-19 Spread as Hate Crime Increase': https://www.newsweek.com/islam-muslims-coronavirus-islamophobia-social-media-twitter-facebook-1523346 [accessed 9 December 2021]

77 Antisemitism Policy Trust, *Policy Briefing* (August 2020): https://antisemitism.org.uk/wp-content/uploads/2020/08/Online-Harms-Offline-Harms-August-2020-V4.pdf [accessed 9 December 2021]

78 Written evidence from Reset (OSB0138)

Abuse against disabled people

30. In its report *Online abuse and the experience of disabled people* in January 2019, the House of Commons Petitions Committee found that, despite the importance of social media for many disabled people's lives, many felt that the online environment was toxic for them. The harms faced by disabled people online include direct abuse, problems with accessibility, and exploitation by malicious actors.[79] Matt Harrison, Public Affairs and Parliamentary Manager at the Royal Mencap Society, told us that negative attitudes and stigma towards disabled people expressed online can "unravel those threads of work that lots of people with learning disabilities themselves have been doing on social media" to move in a positive direction.[80]

31. We have also heard about the unique risk that online platforms can present to individuals with photosensitive epilepsy. Clare Pelham, Chief Executive of the Epilepsy Society, told us that people with photosensitive epilepsy are "regularly" targeted with flashing images that are intended to cause a seizure.[81] She told us that, beyond the severe physical harm that can be caused by having a seizure, this can cause isolation as individuals are "driven off" social media.[82]

Impact on freedom of speech

32. Compassion in Politics described the "current climate of hostility, toxicity, and abuse online" and told us that this "prevents many people from joining social media sites". Their polling found that this can infringe on individuals' freedom of expression, with "1 in 4 … scared of voicing an opinion online because they expect to receive abuse if they do so." [83] Mr Ahmed illustrated this:

> "You do not have free speech if you are a black footballer and 100 racist people jump down your throat every time you post. In fact … this vital tool for promoting your brand and for transacting business is taken away from you."[84]

33. The freedom of social media and search engines to make their own decisions on censoring and recommending content without accountability or oversight was also raised. For example, DMG media told us:

> "We believe it is incompatible with freedom of expression and media plurality for legitimate, responsible news content to be subject to blocking and take-down by a commercial organisation which is open to business pressures such as advertising boycotts, operates without due process, and has no authority to make judgments about the value of journalism."[85]

79 House of Commons Petitions Committee, *Online abuse and the experience of disabled people* (First Report, Session 2017–19, HC 759)
80 Q 56
81 Q 53
82 Q 63
83 Written evidence from Compassion in Politics (OSB0050)
84 Q 6
85 Written evidence from DMG Media (OSB0133)

Societal harms

34. The harms resulting from activity online are not limited to individuals. For example, online disinformation—the intentional spreading of factually incorrect information—and online misinformation—the unknowing spreading of factually incorrect information—harm society more broadly.[86] We heard that the prevalence of disinformation during the COVID-19 pandemic has resulted in vaccine hesitancy and vaccine refusal.[87] This has been linked to higher death rates in certain groups.[88] Vaccine-hesitant individuals have had their health severely impacted by contracting COVID-19, or in the worst cases, died. In the UK, this has created pressure on the NHS.[89] COVID-19 misinformation has led individuals to engage in risky behaviour such as using ineffective drugs as home remedies,[90] or drinking poisonous disinfectant.[91]

35. We heard that disinformation has the potential to harm democracy and national security.[92] Ms Ressa told us that disinformation can affect the integrity of elections: "we will not have integrity of elections if we do not have integrity of facts."[93] Disinformation relating to democratic processes can affect social cohesion[94] with societal divides having been exploited by malicious foreign actors to undermine democratic processes in the US and the UK.[95] We have heard that inauthentic accounts created by real people can give fake legitimacy to political candidates or spread mistrust.[96] Meanwhile, the creation and sharing of manipulated videos and messages such as deepfakes can be used to target political candidates.[97] We received evidence that service providers are aware of these threats, including statements from service providers themselves.[98]

Factors exacerbating harms: business models and system design

36. Many service providers collect data about people who use their platforms for commercial benefit.[99] We heard that service providers are incentivised to maximise users' engagement so that they can collect more data about them and show them more and

86 Written evidence from: LSE Department of Media and Communications (OSB0001); Conscious Advertising Network (OSB0180)

87 Q 2

88 Brit Trogen and Liise-anne Pirofski, 'Understanding Vaccine Hesitancy in COVID-19' *Elsevier Public Health Emergency Collection*, vol.2, (2021), pp 498–501: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8030992/ [accessed 30 November 2021]

89 Q 2, Q 3, Q 10

90 Written evidence from the Center for Countering Digital Hate (OSB0009)

91 Digital, Culture, Media and Sport Committee, *Misinformation in the COVID-19 Infodemic* (Second Report, Session 2019–21, HC 234)

92 Written evidence from Full Fact (OSB0056)

93 Q 193

94 Written evidence from: IMPRESS (OSB0092); Polis Analysis (OSB0108); Mr Hadley Newman (OSB0125); Henry Jackson Society (OSB0028)

95 Q 56; Intelligence and Security Committee, *Russia* (Report, Session 2021–22, HC 632); United States Senate, Select Committee on Intelligence, *Russian Active Measures Campaigns and Interference in the 2016 US Election, Volume 5: Counter Intelligence Threats and Vulnerabilities* (2020): https://www.intelligence.senate.gov/sites/default/files/documents/report_volume5.pdf; written evidence from Reset (OSB0138)

96 Q 131, Q 128

97 Written evidence from Reset (OSB0138)

98 CBS News, 'Whistleblower: Facebook is misleading the public on progress against hate speech, violence, misinformation': https://www.cbsnews.com/news/facebook-whistleblower-frances-haugen-misinformation-public-60-minutes-2021-10-03/ [accessed 15 November 2021]; Twitter, 'Permanent suspension of @realDonaldTrump': https://blog.twitter.com/en_us/topics/company/2020/suspension [accessed 15 November 2021]; Q 105, Q 178, Q 207, Q 222, Q 249, written evidence from Reset (OSB0138)

99 Q 95, Q 61

better targeted adverts.[100] Facebook's most recent quarterly report showed 99 per cent of their income was from advertising.[101] Quarterly reports from Alphabet Inc.[102] and Twitter showed 92 per cent and 88 per cent of their respective profits were from advertising revenue.[103]

37. Metrics concerning time spent on the platform and interaction with content form the basis of key performance indicators (KPIs).[104] We heard that KPIs focused on engagement are maximised regardless of the nature of that engagement or quality of the content that is being engaged with.[105] This can be problematic, as Guillaume Chaslot, ex-YouTube employee and founder of AlgoTransparency, told us:

> "You have cases where engagement is good for the user. When I listen to music, the longer I listen, the better it is for me. [But] When there was a problem with paedophile content on YouTube, they spent a lot of time on the platform, so the algorithm was trying to maximise the amount of paedophile content that was shown to users."[106]

38. We heard evidence from a range of sources that content that creates a risk of harm or factually inaccurate content is many times more engaging than innocuous or accurate content.[107] By making design choices that maximise engagement, service providers therefore exacerbate the presence, spread, and effect of harms.[108] Algorithmic design choices have been heavily implicated in the evidence we have received. The Anti-Defamation League told us:

> "When a user interacts with a piece of content, algorithmic systems recognise signals, like popularity, and then amplify that content. If content is forwarded, commented on, or replied to, social media algorithms almost immediately show such content to more users, prompting increased user engagement, and thus increasing advertising revenue. Research shows that controversial, hateful, and polarizing information and misinformation are often more engaging than other types of content and, therefore, receive wider circulation."[109]

39. Multiple witnesses told us that people who are not searching for misinformation, conspiracist content, and extremism will be recommended such content if their behaviour indicates they may be interested in it.[110] For example, someone interested in wellness may be shown anti-vaccination content.[111] If they interact with this, they could be recommended

100 Q 95
101 Facebook, *Earnings Presentation Q2 2021* (2021): https://s21.q4cdn.com/399680738/files/doc_financials/2021/q2/Q2-2021_Earnings-Presentation.pdf [accessed 16 November 2021]
102 Owners of Google and YouTube
103 Alphabet, 'Alphabet Announces Second Quarter 2021 Results': https://abc.xyz/investor/static/pdf/2021Q2_alphabet_earnings_release.pdf; [accessed 22 November 2021]; Twitter, *Q2 2021: Letter to Shareholders* (July 2021): https://s22.q4cdn.com/826641620/files/doc_financials/2021/q2/Q2'21-Shareholder-Letter.pdf [accessed 22 November 2021]
104 KPIs are a type of performance measurement. KPIs evaluate the success of an organisation or of a particular activity (such as projects, programmes, products and other initiatives) in which it engages [from https://en.wikipedia.org/wiki/Performance_indicator]; Q 92
105 Q 92, Q 95, Q 101, Q 102
106 Q 92
107 Q 136, Q 95, Q 151
108 Q 80
109 Written evidence from Anti-Defamation League (ADL) (OSB0030)
110 Q 101, Q 9
111 Written evidence from the Center for Countering Digital Hate (OSB0009)

far-right conspiracist content or antisemitic content.[112] Ms Haugen told us that service providers' algorithms currently "[make] hate worse" because of the way they amplify and recommend hateful content.[113]

40. People, including children, can be vulnerable to being targeted with content that creates a risk of harm, as algorithms collect data about their interests and serve them with progressively more extreme content to keep them engaged.[114] For example, the Wall Street Journal investigated TikTok's algorithms. They found that within 40 minutes of using the platform, 93 per cent of videos recommended to a user who showed an interest in videos about depression and anxiety would be depression-related.[115] Targeting users with content in this way can reinforce addictive behaviour, where people feel compelled to use the platform even though they may not enjoy doing so.[116]

41. Some people are also served disproportionately high amounts of content that creates a risk of harm by algorithms due to their personal characteristics, as inferred by the platform's algorithms.[117] The Information Commissioner Elizabeth Denham CBE told us that "inferred data is personal data", and that she had concerns about the way platforms use inferred data to direct content to people using their platforms and questioned if this was compliant with data protection law.[118]

The "Prevalence Paradox"

42. All the evidence so far would suggest that a high proportion of online material is hateful, false or creates a risk of harm. Yet, academic research which has systematically examined the prevalence of online content that creates a risk of harm consistently finds that its prevalence is low. Abusive content, for example, made up less than one per cent of overall content online according to a 2019 study.[119] However, 13 per cent of adult respondents to Ofcom's pilot harms survey had experienced trolling in the previous month; six per cent of those respondents had experienced bullying, abusive behaviour, or threats;[120] and 46 per cent of women and non-binary people surveyed by Glitch reported experiencing online abuse during the COVID-19 pandemic.[121] In football, 71 per cent of fans reported having seen racist comments on social media[122] despite only 0.03 per cent

112 Q 8
113 Q 156
114 Q 98, Q 101, Q 102
115 'Inside TikTok's highly secretive algorithm', *Wall Street Journal* (21 July 2021): https://www.wsj.com/video/series/inside-tiktoks-highly-secretive-algorithm/investigation-how-tiktok-algorithm-figures-out-your-deepest-desires/6C0C2040-FF25-4827–8528-2BD6612E3796 [accessed 16 November 2021]
116 Q 150, Q 166, QQ 168–169; Written evidence from: COST Action CA16207 - European Network for Problematic Usage of the Internet (OSB0038); ITV (OSB0204); 5Rights, *Pathways,* (September 2021), p 12: 5Rights Foundation, *Key findings and recommendations from Pathways: How digital design puts children at risk* (September 2021): https://5rightsfoundation.com/uploads/PathwaysSummary.pdf [accessed 9 December 2021]
117 Q 165
118 Q 86
119 The Alan Turing Institute, *How much online abuse is there? A systematic review of evidence for the UK: Policy Briefing – Summary* (2019): https://www.turing.ac.uk/sites/default/files/2019-11/online_abuse_prevalence_summary_24.11.2019_-_formatted_0.pdf [accessed 16 November 2021]
120 Ofcom, *Pilot Online Harms Survey 2020/21* (2021): https://www.ofcom.org.uk/__data/assets/pdf_file/0014/220622/online-harms-survey-waves-1-4-2021.pdf [accessed 16 November 2021]
121 Glitch, *The Ripple Effect: COVID-19 And The Epidemic Of Online Abuse* (September 2020): https://glitchcharity.co.uk/wp-content/uploads/2021/04/Glitch-The-Ripple-Effect-Report-COVID-19-online-abuse.pdf [accessed 16 November 2021]
122 Kick It Out, 'Reporting Statistics': https://www.kickitout.org/Pages/FAQs/Category/reporting-statistics [accessed 16 November 2021]

of posts being identified as discriminatory abuse.[123] In other words, some abusive posts, which make up a minority of content, are seen by a vastly disproportionately number of people.

43. Some of these differences may result from inconsistency between methodological approaches. It is, however, improbable this would account for all, or even most, of the gap. Other explanations are:

a) It may be a consequence of the easy dissemination and algorithmic promotion of content that creates a risk of harm, the "boosting" effect we discuss above.

b) The studies may not be using comparable definitions of harm, for example some reports focus on abuse specifically,[124] whereas others may include content which is discriminatory but not directly abusive.[125]

c) There may be an element of reporting bias or self-selection bias in polling studies.

d) Some groups are more likely to receive online abuse than others, so that whilst overall prevalence of content that creates a risk of harm may be low, people in these groups will report experiencing proportionately more harmful material.[126]

e) Activity on engagement-based platforms can often "snowball", meaning that people can be targeted for abuse by large groups of other users. Where individuals are the focus of such a "pile-on" attack they are the singular target of vast quantities of abusive material.

The "black box"

44. One of the challenges of establishing exactly why content and activity that is abusive, false or creates a risk of harm is so overexposed is that the systems underlying platforms are like a "black box". Users, researchers, and regulators often have limited understanding of their internal workings or the risks posed by them.[127] Currently researchers do not have access to high-quality data from service providers which would allow them to conduct systematic, longitudinal, trustworthy research, despite, as we heard from many witnesses, requests for it.[128] We discuss this further in Chapter 9.

45. For people using service providers' platforms, a lack of transparency can lead to frustration with systems when they do not appear to be working—for example when activity that creates a risk of harm or is abusive is reported but not addressed.[129] For researchers and civil society, a lack of transparency around data and the algorithms that

123 Professional Footballers' Association, *Online Abuse: AI Research Study: Season 2021/21* (2021): https://www.thepfa.com/news/2021/8/4/online-abuse-ai-research-study-season-2020–21 [accessed 16 November 2021]

124 The Alan Turing Institute, *How much online abuse is there? A systematic review of evidence for the UK: Policy Briefing – Summary (2021):* https://www.turing.ac.uk/sites/default/files/2019-11/online_abuse_prevalence_summary_24.11.2019_-_formatted_0.pdf [accessed 16 November 2021]

125 Kick It Out, 'Reporting Statistics': https://www.kickitout.org/Pages/FAQs/Category/reporting-statistics [accessed 16 November 2021]

126 Glitch UK and End Violence Against Women Coalition , *The Ripple Effect: COVID-19 and the Epidemic of Online Abuse* (2020): https://glitchcharity.co.uk/wp-content/uploads/2021/04/Glitch-The-Ripple-Effect-Report-COVID-19-online-abuse.pdf [accessed 16 November 2021]

127 Q 136, Q 146, Written evidence from: the Ada Lovelace Institute (OSB0101); ITV (OSB0204); Q 72

128 Written evidence from Dr Amy Orben (College Research Fellow at Emmanuel College, University of Cambridge) (OSB0131)

129 Q 29, Q 37, Q 46, Q 53, Q 62, Q 60, Q 82

platforms use is a barrier to being able to understand and tackle content that creates a risk of harm and illegal content online.[130] A lack of transparency also means that service providers do not have any accountability, to quote one provider: "Without transparency, there can be no accountability."[131] We heard repeatedly that service providers' lack of transparency is a key issue for online harms and must be addressed.[132]

46. Some companies now regularly produce transparency reports detailing information about content and activity that is illegal, creates a risk of harm or is against their terms of service.[133] We heard, however, that the information provided in some of these reports can be misleading. Certain metrics can imply high rates of success or low levels of content and activity that create a risk of harm, when that may not be an accurate reflection of what is occurring on platforms. For example, knowing that 90 per cent of policy-violating content that is removed from a service is identified and removed by algorithms, rather than due to user reports, does not give an indication of the overall proportion of policy-violating content that is successfully identified and removed by those algorithms. If only 1 per cent of policy-violating content is ultimately identified, the algorithms would be removing 0.9 per cent of the total amount of policy-violating content that is present on the service.[134] These metrics are therefore insufficient for achieving the transparency and accountability that is needed to understand and mitigate the presence and spread of online content and activity that creates a risk of harm.

47. These metrics also hide the human impact of content and activity that creates a risk of harm. Statistics about the prevalence of policy-violating content do not capture the people in urgent conditions in hospital who have taken fake medical cures. They do not show the children who have harmed themselves and who suffer from severe mental health difficulties because of what they have experienced online, or the enduring impact on people who have lost their life savings to online scams.

The draft Bill

48. The draft Bill introduced by the Government aims make the UK "the safest place in the world to be online". To achieve this aim, it proposes a new regulatory regime with Ofcom as an independent regulator for providers of online user-to-user and search services.[135]

49. Online service providers are broadly supportive of the Government introducing regulation that aims to enhance online safety.[136] Facebook themselves have said that they

130 Written evidence from Dr Amy Orben (College Research Fellow at Emmanuel College, University of Cambridge) (OSB0131)

131 Q 178; Twitter, *Protecting The Open Internet: Regulatory principles for policy makers:* https://cdn.cms-twdigitalassets.com/content/dam/about-twitter/en/our-priorities/open-internet.pdf [accessed 16 November 2021]

132 Q 87

133 For example: Twitter, 'Transparency Reports': https://transparency.twitter.com/en/reports.html [accessed 16 November 2021]; Meta, 'Community Standards Enforcement Report': https://transparency.fb.com/data/community-standards-enforcement/ [accessed 16 November 2021]

134 Q 154, Q 14

135 Written evidence from the Department of Digital, Culture, Media and Sport and Home Office (OSB0011)

136 Written evidence from: Snap Inc. (OSB0012); Mumsnet (OSB0031); Match Group (OSB0053); Bumble Inc. (OSB0055); Twitter (OSB0072); Microsoft (OSB0076); Patreon Inc. (OSB0123); Facebook (OSB0147); Google (OSB0175); TikTok (OSB0181)

feel they are currently making societal decisions that are better made by Government and regulators.[137]

50. We have, however, heard numerous concerns about the Online Safety Bill as currently drafted.[138] Briefly:

a) The Bill is overly complex, which has the potential to create legislative gaps and loopholes.[139] The "Duty(ies) of Care" framework is particularly complex and confusing.[140]

b) The Bill lacks clarity around several key aspects, making it more at risk of legal challenge:

   i) What would constitute content that is harmful, and associated definitions such as a "person of ordinary sensibilities";[141]

   ii) The definitions of "journalistic content" and "content of democratic importance", and how service providers would be expected to identify these types of content;[142]

   iii) Which types of content would be designated "priority harms";[143]

   iv) Which service providers would be in scope of the "Category 1" requirements;[144] and

   v) Some of the requirements of the Bill will undermine, conflict with or are misaligned to the standards in the Age Appropriate Design Code/existing regulation.[145]

c) The provisions in the draft Bill on content that is harmful to adults could have a "chilling effect" on freedom of expression and give too much power to service providers.[146]

d) Ofcom's powers may not be sufficient for them to achieve success in their role as a regulator.[147]

e) The transparency requirements placed on service providers may not go far enough.[148]

f) The Secretary of State's powers are extensive and may undermine Ofcom's independence with no effective accountability to Parliament.[149]

g) The Bill does not provide sufficient protections for children, including failure to capture all pornography sites.[150]

137 'Opinion: Mark Zuckerberg: The Internet needs new rules. Let's start in these four areas.' *The Washington Post (30 March 2019):* https://www.washingtonpost.com/opinions/mark-zuckerberg-the-internet-needs-new-rules-lets-start-in-these-four-areas/2019/03/29/9e6f0504–521a-11e9-a3f7-78b7525a8d5f_story.html [accessed 16 November 2021]

138 Please note that this is not an exhaustive list

139 Q 69

140 Written evidence from Snap Inc. (OSB0012)

141 Written evidence from Gavin Millar QC (OSB0221)

142 Written evidence from Dr Martin Moore (Senior Lecturer at King's College London) (OSB0063)

143 Written evidence from Care (OSB0085)

144 Written evidence from Barnardo's (OSB0017)

145 Written evidence from: Common Sense Media (OSB0018), 5Rights Foundation (OSB0096)

146 Written evidence from Dr Edina Harbinja (Senior lecturer in law at Aston University, Aston Law School) (OSB0145)

147 Q 85

148 Q 14

149 Q 72; Written evidence from Ofcom (OSB0021)

150 Written evidence from: NSPCC (OSB0228), The Office of the Children's Commissioner (OSB0019)

51. **Self-regulation of online platforms has failed. Our recommendations will strengthen the Bill so that it can pass successfully into legislation. To achieve success, the Bill must be clear from the beginning about its objectives. These objectives must reflect the nature of the harm experienced online and the values of UK society. Online services are not neutral repositories for information. Most are advertising businesses. Service providers in scope of the Bill must be held liable for failure to take reasonable steps to combat reasonably foreseeable harm resulting from the operation of their services.**

52. ***We recommend the Bill is restructured. It should set out its core objectives clearly at the beginning. This will ensure clarity to users and regulators about what the Bill is trying to achieve and inform the detailed duties set out later in the legislation. These objectives should be that Ofcom should aim to improve online safety for UK citizens by ensuring that service providers:***

a) ***comply with UK law and do not endanger public health or national security;***

b) ***provide a higher level of protection for children than for adults;***

c) ***identify and mitigate the risk of reasonably foreseeable harm arising from the operation and design of their platforms;***

d) ***recognise and respond to the disproportionate level of harms experienced by people on the basis of protected characteristics;***

e) ***apply the overarching principle that systems should be safe by design whilst complying with the Bill;***

f) ***safeguard freedom of expression and privacy; and***

g) ***operate with transparency and accountability in respect of online safety.***

An overarching duty of care?

53. The 2019 White Paper promised to introduce a "new duty of care" for service providers towards the people using their platforms.[151] This language and proposal drew on the work of Professor Lorna Woods and William Perrin OBE at Carnegie UK Trust. They proposed that service providers should be held responsible for a public space in the same way that property owners are responsible for physical spaces, and that service providers should have a duty of care in respect of the people using their platforms. Prof Woods and Mr Perrin also argued that a statutory duty of care would be "simple, broadly based and largely future-proof", much like the long-enduring Health and Safety at Work Act 1974.[152] The language of a duty of care for service providers has persisted, with the draft Bill setting out several "duties of care" and "safety duties". These "duties of care" however, operate in a fundamentally different way to the duty of care laid out in the Health and Safety at Work Act 1974.

151 Department for Digital, Culture, Media and Sport and The Home Office, *Online Harms White Paper*, CP 57, April 2019, p 8: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/973939/Online_Harms_White_Paper_V2.pdf [accessed 7 December 2021]

152 Carnegie UK , *Online harm reduction – a statutory duty of care and regulator* (April 2019): https://d1ssu070pg2v9i.cloudfront.net/pex/pex_carnegie2021/2019/04/06084627/Online-harm-reduction-a-statutory-duty-of-care-and-regulator.pdf [accessed 9 December 2019]

54. Some submissions we received noted that the draft Bill had moved away from the White Paper in the duties it places on service providers. The draft Bill places duties on service providers to do particular things, such as undertake risk assessments, to comply with safety duties in respect of illegal content, content that is harmful to children and content that is harmful to adults and other duties, for example in respect of journalistic content. It does not propose "a [singular] new duty of care" as set out in the Government's response to its White Paper.[153] Nor do these new duties constitute a duty of care in the legal sense. They are things that providers are required to do to satisfy the regulator. They are not duties to people who use their platforms, and they are not designed to create new grounds for individuals to take providers to court.

55. For children's rights charities and Carnegie UK Trust themselves, this was a significant step backwards. They were concerned that the lack of an overarching duty to address "foreseeable risks", might lead to emerging issues falling between the cracks of the various duties in the legislation.[154] The complexity of the interlocking series of duties were also a common theme in evidence, cutting across many of the different groups we took evidence from.[155]

56. The Government explained that the structure adopted in the draft Bill seeks to cover the same scope as the duty of care envisaged in the White Paper, but the move to "more specific duties will give companies and Ofcom greater legal certainty and direction about the regime. In turn this will make it easier for Ofcom to effectively enforce against non-compliance."[156] The Secretary of State was more explicit. She told us that a single duty of care:

> " … does not work … The definitions within that duty of care are huge, onerous and difficult legally to make tight and applicable. I am not going to tell you what to do, but I would probably put your efforts into other parts of the Bill, because we have already been there and we know that it would be almost impossible to get that into the Bill. … That is why the Bill is so long. It is a technical, long Bill, but in order to meet the criteria of watertight it has to be."[157]

57. The Secretary of State's concerns about the workability of a duty of care approach aligned with that of a few of our witnesses. Mr Ahmed told us that he wanted to see as much as possible defined on the face of the Bill, because: "The less clarity there is, the harder it is on those companies to do the right thing, and the more wriggle room there is for them to escape from it."[158] Gavin Millar QC, specialist in media law at Matrix Chambers, told us that the draft Bill as it stood was open to legal challenge. He saw a fundamental problem with transposing a duty of care approach into the regulation of

153 Department for Digital, Culture, Media and Sport and The Home Office, *Online Harms White Paper: Full government response to the consultation*, CP 354, December 2020: https://www.gov.uk/government/consultations/online-harms-white-paper/outcome/online-harms-white-paper-full-government-response [accessed 12 November 2021]

154 For example, Q 66 (Izzy Wick), Q 69 (William Perrin), Q 70 (Professor Sonia Livingstone); Written evidence from: NSPCC (OSB0109); Mr John Carr (Secretary at Children's Charities' Coalition on Internet Safety) (OSB0167)

155 For example, written evidence from: Snap Inc. (OSB0012); Internet Watch Foundation (IWF) (OSB0110); Parent Zone (OSB0124); Dr Martin Moore (Senior Lecturer at King's College London) (OSB0063); Damian Tambini (Distinguished Policy Fellow and Associate Professor at London School of Economics and Political Science) (OSB0066); Twitter (OSB0072); BBC (OSB0074); Care (OSB0085); Carnegie UK (OSB0095); techUK (OSB0098); NSPCC (OSB0109); Parent Zone (OSB0124); Facebook (OSB0147); Google (OSB0175); Confederation of British Industry (CBI) (OSB0186); TalkTalk (OSB0200)

156 Written evidence from Department of Digital, Culture, Media and Sport and Home Office (OSB0011)

157 Q 286 (Rt Hon Nadine Dorries MP)

158 Q 17 (Imran Ahmed)

online platforms being that the law of negligence is an unqualified duty, whereas duties on service providers involve balancing the fundamental rights of different groups of people against each other.[159] He, along with other witnesses who had similar concerns, wanted to see the Bill go further in the direction of specifying exactly which risks of harm it intends to address and what service providers should be doing about it.[160]

58. Towards the end of our inquiry, Carnegie UK Trust produced a series of revisions to the draft Bill. They proposed a restructuring with an overarching set of objectives, underpinned by a "Foundation duty", in turn underpinned by specific duties, along the lines of those that can be found in the draft Bill.[161] A possible model that also offers a similar structure may be provided by the Financial Conduct Authority's (FCA's) current consultation on a "consumer principle". The FCA's model proposes an overarching principle—that firms act in the best interests of consumers—followed by a set of cross-cutting rules and then detailed rules and guidance. In the same way, we envisage a Bill with an overarching set of core safety objectives on Ofcom, and a series of statutory requirements on providers to implement detailed mandatory Codes of Practice.[162]

59. **The draft Bill creates an entirely new regulatory structure and deals with difficult issues around rights and safety. In seeking to regulate large multinational companies with the resources to undertake legal challenges, it has to be comprehensive and robust. At the same time, a common theme in the evidence we received is that the draft Bill is too complex, and this may harm public acceptance and make it harder for those service providers who are willing to comply to do so.**

60. ***We recommend that the Bill be restructured to contain a clear statement of its core safety objectives—as recommended in paragraph 52. Everything flows from these: the requirement for Ofcom to meet those objectives, its power to produce mandatory codes of practice and minimum quality standards for risk assessments in order to do so, and the requirements on service providers to address and mitigate reasonably foreseeable risks, follow those codes of practice and meet those minimum standards. Together, these measures amount to a robust framework of enforceable measures that can leave no doubt that the intentions of the Bill will be secured.***

61. ***We believe there is a need to clarify that providers are required to comply with all mandatory Codes of Practice as well as the requirement to include reasonably foreseeable risks in their risk assessments. Combined with the requirements for system design we discuss in the next chapter, these measures will ensure that regulated services continue to comply with the overall objectives of the Bill—and that the Regulator is afforded maximum flexibility to respond to a rapidly changing online world.***

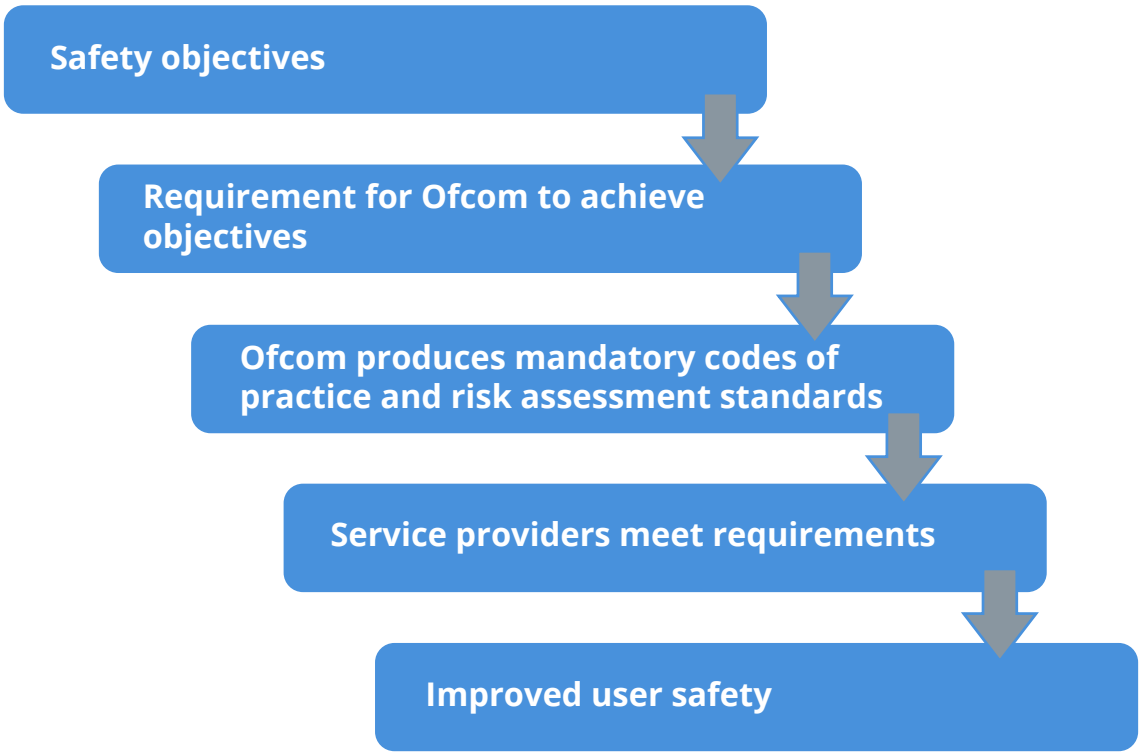
159 Q 143 (Gavin Millar QC)

160 Q 143 (Gavin Millar QC); for example, Q 60 (Dr Edina Harbinger).

161 Carnegie UK, 'Amendments and Explanatory Notes: Carnegie UK Revised Online Safety Bill - Nov 2021': https://www.carnegieuktrust.org.uk/publications/amendments-explanatory-notes-carnegie-uk-revised-online-safety-bill-nov-2021/ [accessed 18 November 2021]

162 Financial Conduct Authority, 'FCA proposes stronger protection for consumers in financial markets': https://www.fca.org.uk/news/press-releases/fca-proposes-stronger-protection-consumers-financial-markets [accessed 18 November 2021]

Figure 1: how the Online Safety Bill will work under our recommendations



3 Societal harm and the role of platform design

Content and activity

62. One of the most common criticisms we heard of the draft Bill was that it focused too heavily on content and not enough on system design or broader "activity".[163] The Government has said that the draft Bill is a "systems and processes" bill—aimed at addressing systemic issues with online platforms rather than seeking to regulate individual content. At the same time, it defines multiple different types of content and specifies how platforms should address them. The Bill should be clear on this.

63. Service providers can, and should, be held accountable for carelessly hosting content that creates a risk of harm. That requires clear definitions. At the same time, in many cases we heard it is the virality, the aggregation, and the frictionless nature of sharing that determines how much harm is caused by any individual piece of content. As Jimmy Wales, founder of Wikipedia, put it:

> "I do not have a crazy racist uncle, but we all know the stereotype, down at the pub spouting off nonsense to his mates. That is a problem, but it is not a problem requiring parliamentary scrutiny. When it becomes a problem is not that my crazy uncle posts his racist thoughts on Facebook, but that he ends up with 5,000 or 10,000 followers, because everyone in the family yells at him and the algorithm detects, "Ooh, engagement", and chases after that, and begins to promote it. That is a problem, and it is a really serious problem that is new and different."[164]

64. One of the changes that the Government made in the draft Bill, compared to the White Paper, was to replace references to "content and activity" with references solely to "content". This has reinforced the sense among many of our witnesses that the draft Bill is concerned solely with content moderation. 5Rights called for a return to the "content and activity" language of the White Paper, arguing that "content" alone does not reflect the full range of risks that children are exposed to online.[165] Interestingly the Government's own written evidence referred to "content and activity" when talking about provisions in the draft Bill.[166]

65. Activity that creates a risk of harm can take many forms and can originate from people using the platforms or from platforms themselves. Examples of people's activity that can create a risk of harm are the mass reporting of individuals to platforms for spurious breaches of terms and conditions as a form of harassment, adults initiating unsupervised contact with children, excluding individuals from online groups to harass them, and control of technology in domestic abuse cases.

163 For example, written evidence from: Reset (OSB0138); Dr Edina Harbinja (Senior lecturer in law at Aston University, Aston Law School) (OSB0145); LGBT Foundation (OSB0191)

164 Q 80 (Jimmy Wales)

165 Written evidence from 5Rights Foundation (OSB0096); although not explicitly discussed in the evidence we heard, many children's groups use the "four C's" of content, contact, contract and conduct to describe online risks. See for example UK Safer Internet Centre, 'What are the issues?': https://saferinternet.org.uk/guide-and-resource/what-are-the-issues [accessed 15 November 2021]

166 Written evidence from the Department of Digital, Culture, Media and Sport and Home Office (OSB0011)

66. As discussed in Chapter 3, platforms' activity can itself create a risk of harm, such as when unsafe content is promoted virally, people are automatically invited to join groups[167] which share extreme views or where recommendation tools prioritise content that creates a risk of harm.

67. We are also concerned that "content" may prove too limiting in a rapidly developing online world. We heard during our inquiry about the need to ensure that the Bill keeps up with changes in the online world, the increasing use of virtual and augmented reality and, of course, Facebook's launch of the "metaverse".[168]

68. ***We recommend that references to harmful "content" in the Bill should be amended to "regulated content and activity". This would better reflect the range of online risks people face and cover new forms of interaction that may emerge as technology advances. It also better reflects the fact that online safety is not just about moderating content. It is also about the design of platforms and the ways people interact with content and features on services and with one another online.***

Algorithmic design

69. Platform design is central to what people see and experience on social media. Platforms do not neutrally present content. For most user-to-user platforms, algorithms are used to curate a unique personalised environment for each user. [169] To create these environments, algorithms use detailed information about the user such as their behaviour on the platform (how long they have watched a certain video or what content they have interacted with), and their geographical location.[170] As Laura Edelson, a researcher at New York University, said:

> "In any Category 1 platform that I know of … there is no action that a user can take in the public news feed or in a public Twitter feed that will guarantee that another user will see that piece of content. Every action that you take in an interaction only feeds into the likelihood that a recommendation algorithm will then show that to another user."[171]

70. Designing curated environments for individual people can give them content that they are interested in and want to engage with, enhancing their experience on the platform. The commercial imperative behind this is to hold people's attention and

167 A 2016 report from Facebook showed that 64% of the time when Facebook users joined extremist groups, the groups had been recommended by the site's algorithms. Study: 'Facebook Allows And Recommends White Supremacist, Anti-Semitic And QAnon Groups With Thousands Of Members' *Forbes* (4 August 2020): https://www.forbes.com/sites/jemimamcevoy/2020/08/04/study-facebook-allows-and-recommends-white-supremacist-anti-semitic-and-qanon-groups-with-thousands-of-members [accessed 9 December 2021]

168 Q 77 (Dr Edina Harbinja); Q 271 (Dame Melanie Dawes)

169 Q 245

170 For example, Q93; Q92; For an example of the sorts of information used see written evidence from Elizabeth Kanter (Director of Government Relations at TikTok) (OSB0219)

171 Q 95

maximise engagement.[172] However, the choice to design platforms for engagement can be problematic:

> "Engagement is maximised by (1) strong emotion, (2) rabbit holes that lead to a warren of conspiracy, (3) misinformation that gets engagement from detractors and supporters, and (4) … algorithmic reinforcement of prior beliefs."[173]

71. Algorithms designed to maximise engagement can directly result in the amplification of content that creates a risk of harm.[174] For example, the CCDH found that 714 posts manually identified as antisemitic across five social media platforms reached 7.3 million impressions over a six-week period.[175] By maximising engagement, algorithms can also hyper-expose individual people to content which exposes them to a high risk of harm.[176] In showing people content that is engaging, algorithms can lead them down a "rabbit hole" whereby content that creates a risk of harm becomes normalised and they are exposed to progressively more extreme material.[177] As Mr Ahmed told us, people are more likely to believe things they see more often and news feeds and recommendation tools are a powerful way to influence a person's worldview.[178] ITV told us: "show an interest in a topic, even one that is potentially harmful, and their core business model and algorithms will find more of it for you."[179]

72. Ms Haugen explained how recommendation systems can be designed to continuously serve content to people: "Instead of you choosing what you want to engage with, [YouTube] Autoplay chooses for you, and it keeps you in … a flow, where it just keeps you going."[180] This can be dangerous[181] as "the AI (artificial intelligence) isn't built to help you get what you want—it's built to get you addicted … "[182] and "there is no conscious action of continuing or picking things, or whether or not to stop. That is where the rabbit holes come from."[183] In 2018 YouTube said that over 70 per cent of videos were viewed in response to recommendations.[184] In written evidence to us they said the figure "fluctuates" but remains a majority.[185] Continually serving content can create a risk of addictive behaviours in some people, and we heard particular concerns from witnesses about the susceptibility of children to addictive behaviours and "problematic use".[186]

172 Q 136; Q 92, Q 95, Q 101, Q 102; Written evidence from Global Action Plan (OSB0027)
173 Written evidence from Center for Countering Digital Hate (OSB0009)
174 Written evidence from: Center for Countering Digital Hate (OSB0009); 5Rights Foundation (OSB0096); Q156; Common Sense (OSB0018); Anti-Defamation League (ADL) (OSB0030); Glitch (OSB0097); 5Rights Foundation (OSB0206)
175 Center for Countering Digital Hate: *Failure to Protect: How tech giants fail to act on user reports of antisemitism* (2021): https://252f2edd-1c8b-49f5-9bb2-cb57bb47e4ba.filesusr.com/ugd/f4d9b9_cac47c87633247869bda54fb35399668.pdf [accessed 16 November 2021]
176 Q 165; Q 171
177 Q 98; QQ 101–102
178 Q 2
179 Written evidence from ITV (OSB0204)
180 Q 191
181 Written evidence from COST Action CA16207 - European Network for Problematic Usage of the Internet (OSB0038)
182 The Next Web News, '"YouTube recommendations are toxic" says dev who worked on the algorithm': https://thenextweb.com/news/youtube-recommendations-toxic-algorithm-google-ai [accessed 9 December 2021]; Written evidence from ITV (OSB0204)
183 Q 191
184 Cnet,'YouTube's AI is the puppet master over most of what you watch': https://www.cnet.com/news/YouTube-ces-2018-neal-mohan/ [accessed 30 November 2021]; Q 92
185 Written evidence from Google UK Limited (OSB0218)
186 Q 150

Frictionless activity

73. Platforms are often designed to minimise friction for users, maximising their ability to interact with one another and diversify their communications through multiple different services with minimal effort. Autoplay, discussed above, is an example of a friction-reducing design feature—making it easier for the person using the system to watch another piece of content chosen for them by the platform, rather than choosing their own content or switching off.[187]

74. We heard that " … safety measures frequently come into conflict with the 'maximise engagement, minimise friction' incentives of the surveillance advertising business model."[188] Mr Ahmed told us that platforms as currently designed allowed twelve individuals to produce two thirds of the anti-COVID vaccine disinformation that their organisation identified online.[189] Witnesses told us about one case where the ability to easily invite large volumes of people to join Facebook groups resulted in an individual user sending invites to 300,000 other users to a group which proliferated extreme views.[190]

75. Where platform design allows communication between adults and children, frictionless interaction and movement between platforms means there is a particular risk of facilitating child sexual exploitation and abuse (CSEA). The NSPCC was particularly concerned about cross-platform abuse: "abusers exploit the design features of social networks to make effortless contact with children, before the process of coercion and control over them is migrated to encrypted messaging or live streaming sites."[191] Private spaces such as chat rooms, closed groups, and encrypted messages were of particular concern to our witnesses.[192]

Safety by design as a mitigation measure

76. In June 2020, DCMS published guidance on how service providers can mitigate the risk of harmful and illegal activity by integrating safety into the design of platforms. They describe safety by design as: "the process of designing an online platform to reduce the risk of harm to those who use it … It considers user safety throughout the development of a service, rather than in response to harms that have occurred."[193]

77. Ms Haugen gave us an example of how non-content-based interventions can be effective in reducing the virality of content:

> "Let us imagine that Alice posts something and Bob reshares it and Carol reshares it, and it lands in Dan's news feed. If Dan had to copy and paste that to continue to share it, if the share button was greyed out, that is a two-

187 The Age Appropriate Design Code encourages services to introduce wellbeing enhancing behaviours such as taking breaks, many of which have now been introduced by some services.
188 Written evidence from Global Action Plan (OSB0027)
189 Q 2
190 Q 153; Written evidence from Google UK Limited (OSB0218)
191 Written evidence from NSPCC (OSB0109)
192 Written evidence from: NSPCC (OSB0109); Barnardo's (OSB0017); Mrs Gina Miller (OSB0112); Dame Margaret Hodge (Member of Parliament for Barking and Dagenham at House of Commons) (OSB0201); Information Commissioner's Office (OSB0211)
193 Department for Digital, Culture, Media and Sport, *Principles of safer online platform design* (June 2021): https://www.gov.uk/guidance/principles-of-safer-online-platform-design [accessed 19 November 2021]

> hop reshare chain, and it has the same impact as the entire third party fact-checking system … "[194]

78. We heard similar ideas from Renée DiResta, Technical Director at the Stanford Internet Observatory, who told us that you could implement a circuit-breaker whereby "when content reaches a particular threshold of velocity or virality" you could send it to "the relevant teams within the platform so that they can assess what is happening".[195] We heard that you could also "throttle its distribution while that is happening if it falls into a particular type of content that has the potential to harm."[196] Ms DiResta explained that this concept is like one already used in financial markets.[197]

79. Some service providers are already implementing some safety by design measures on their platforms. In response to the Age Appropriate Design Code, YouTube recently changed the settings for Autoplay so that it is turned off by default for people using their platform who are aged 13–17.[198] Twitter told us that they had introduced a "nudge" for people to read articles before they share them, [199] and Snap Inc. told us that Snapchat has no open news feeds where "unvetted publishers or individuals have an opportunity to broadcast hate or misinformation, and [that it doesn't] offer public comments that may amplify harmful behaviour."[200]

80. Reset argued that rather than asking companies to write rules for content, the Online Safety Bill "should require them to improve their systems and designs: mandating practical solutions to minimise the spread of harmful material by focusing on preventative measures such as reduced amplification, demonetisation and strict limits on targeting." They outlined how incorporating safety by design principles could affect people's experiences of using platforms:

> "… abusive tweets sent in the heat of the moment to a footballer who had a bad game aren't promoted to other disappointed fans, causing an abusive pile on. Before telling a Love Island heartthrob who has fallen from grace to kill themselves, users are asked to think twice. Having the option to delay when your comment is posted, becomes the norm. Being directed to authoritative, fact-checked sites about climate change or coronavirus before you watch a conspiracy theory video might give pause for thought." [201]

81. **We heard throughout our inquiry that there are design features specific to online services that create and exacerbate risks of harm. Those risks are always present, regardless of the content involved, but only materialise when the content concerned is harmful. For example, the same system that allows a joke to go viral in a matter of minutes also does the same for disinformation about drinking bleach as a cure for COVID-19. An algorithm that constantly recommends pictures of cats to a cat-lover**

194 Q 161
195 Q 105
196 QQ 105–106
197 Q106
198 Vox, 'YouTube's kids app has a rabbit hole problem',: https://www.vox.com/recode/22412232/youtube-kids-autoplay [accessed 22 November 2021]; Fatherly,'YouTube Finally Turns off Autoplay for Kids. Here's the Catch': https://www.fatherly.com/news/youtube-autoplay-kids [accessed 22 November 2021]; Alphabet Inc., '*YouTube Help: Autoplay Videos*': https://support.google.com/youtube/answer/6327615?hl=en [accessed 22 November 2021]
199 Q 247
200 Written evidence from Snap Inc. (OSB0012)
201 Written evidence from Reset (OSB0138)

is the same algorithm that might constantly recommend pictures of self-harm to a vulnerable teenager. Tackling these design risks is more effective than just trying to take down individual pieces of content (though that is necessary in the worst cases). Online services should be identifying these design risks and putting in place systems and process to mitigate them before people are harmed. The Bill should recognise this. Where online services are not tackling these design risks, the regulator should be able to take that into account in enforcement action.

82. ***We recommend that the Bill includes a specific responsibility on service providers to have in place systems and processes to identify reasonably foreseeable risks of harm arising from the design of their platforms and take proportionate steps to mitigate those risks of harm. The Bill should set out a non-exhaustive list of design features and risks associated with them to provide clarity to service providers and the regulator which could be amended by Parliament in response to the development of new technologies. Ofcom should be required to produce a mandatory Safety by Design Code of Practice, setting out the steps providers will need to take to properly consider and mitigate these risks. We envisage that the risks, features and mitigations might include (but not be limited to):***

a) ***Risks created by algorithms to create "rabbit holes", with possible mitigations including transparent information about the nature of recommendation algorithms and user control over the priorities they set, measures to introduce diversity of content and approach into recommendations and to allow people to deactivate recommendations from users they have not chosen to engage with;***

b) ***Risks created by auto-playing content, mitigated through limits on auto-play and auto-recommendation;***

c) ***Risks created by frictionless cross-platform activity, with mitigations including warnings before following a link to another platform and ensuring consistent minimum standards for age assurance;***

d) ***Risks created through data collection and the microtargeting of adverts, mitigated through minimum requirements for transparency around the placement and content of such adverts;***

e) ***Risks created by virality and the frictionless sharing of content at scale, mitigated by measures to create friction, slow down sharing whilst viral content is moderated, require active moderation in groups over a certain size, limit the number of times content can be shared on a "one click" basis, especially on encrypted platforms, have in place special arrangements during periods of heightened risk (such as elections, major sporting events or terrorist attacks); and***

f) ***Risks created by default settings on geolocation, photo identification/sharing and other functionality leading to victims of domestic violence or VAWG being locatable by their abusers, mitigated through default strong privacy settings and accessible guidance to victims of abuse on how to secure their devices and online services.***

83. ***We recommend that the Bill includes a requirement for service providers to co-operate to address cross-platform risks and on the regulator to facilitate such co-operation.***

Anonymity and traceability

84. A common design feature across many user-to-user services is allowing anonymous and pseudonymous accounts, where users are either publicly unidentifiable or partly identifiable (e.g. identifiable by their first name only). The role of anonymity in facilitating abuse was a key theme in the evidence we received from sporting bodies. Mr Ferdinand told us that "the fact that you can be anonymous online is an absolute problem for everybody in society."[202] We also heard about anonymous abuse or abuse from fake or disposable accounts directed against politicians, Jewish people, women, victims of domestic abuse, journalists in repressive regimes and in the UK, as well as to organise extremist activity and threats[203], and during our roundtable on age assurance and anonymity about the "disinhibition" effect associated with posting online and when posting anonymously. Most studies suggest that anonymity is likely to lead to more abusive or "uncivil" behaviour—though it is important to note there is at least one well-known study that points the other way.[204]

85. Ms Ressa noted that the ease of creation and disposal of anonymous online accounts made them key tools in the disinformation and harassment campaign against her. She told us that the exponential attacks she had experienced "came from anonymous accounts because they are easy to make and easy to throw out". She said that, as a consequence of activity from anonymous accounts, she has "watched [her] credibility get whittled away. You cannot respond. If you are the journalist or if you are a government official, your hands are tied. You are responding to a no-name account. You just do not do things like that. The attacks are horrendous."[205]

86. Ms Haugen and Mr Perrin both questioned whether ending anonymity would be effective or proportionate at achieving the desired outcome of ending online abuse.[206] We heard that being identifiable did not prevent abuse: much of the misogynistic abuse Ms Jancowicz and other prominent women received came from identifiable accounts, and she had received abuse on LinkedIn where the abuser's employer or prospective employer may see it.[207] We also heard of the importance of anonymity to marginalised groups, victims

202 Q 19 (Rio Ferdinand); see also for example written evidence from: The Football Association, The Premier League, EFL, Kick It Out (OSB0007); Sport and Recreation Alliance (OSB0090); 5 Sports: The Football Association, England and Wales Cricket Board, Rugby Football Union, Rugby Football League and Lawn Tennis Association, The FA (OSB0111)

203 For example, written evidence from: Compassion in Politics (OSB0050); Dame Margaret Hodge (Member of Parliament for Barking and Dagenham at House of Commons) (OSB0201); Antisemitism Policy Trust (OSB0005); Centenary Action Group, Glitch, Antisemitism Policy Trust, Stonewall, Women's Aid, Compassion in Politics, End Violence Against Women Coalition, Imkaan, Inclusion London, The Traveller Movement (OSB0047); Refuge (OSB0084); Q 194 (Maria Ressa); The National Union of Journalists (NUJ) (OSB0166); HOPE not hate (OSB0048); Mrs Gina Miller (OSB0112)

204 For a summary of some of the key research discussed see, Clean Up the Internet, 'Academic Research about online disinhibition, anonymity and online harms': https://www.cleanuptheinternet.org.uk/post/some-useful-scholarly-articles-about-online-disinhibition-anonymity-and-online-harms [accessed 18 November 2021]

205 Q 194 (Maria Ressa)

206 Q 73 (William Perrin); Q 171 (Frances Haugen)

207 Q 62 (Nina Jankowicz)

of violence, whistleblowers, and children.[208] Nancy Kelley, Chief Executive of Stonewall, explained:

> "I know people have suggested things like names being visible. Even in progressive countries that are accepting, we know that will expose LGBTQ people to harm. …
>
> If we look at that in the global context, we know from research that Article 19 has done that almost 90 per cent of LGBTQ users in Egypt, Lebanon and Iran said they are incredibly frightened of mentioning even their name in any kind of private messaging online. We know that over 50 per cent of the men charged in Egypt in recent years with homosexual 'offences'—because it is indeed illegal to be gay there, as it still is in 71 countries around the world—were the subject of online stings."[209]

87. One suggestion we heard to address the risks posed by anonymity would be to require people to provide an "anchor" to their real-world identity when creating an account, so that people can be held accountable. Such an approach has also been recommended by Siobhan Baillie MP in her Social Media Platforms (Identity Verification) Bill, and in her written evidence to the Committee.[210] This would provide traceability in the event of their posting illegal content or engaging in illegal activity, but without requiring them to post under their real names.[211] Crucially, we heard that traceability would have to meet minimum standards on quality in order to be effective. The FA told us that "the ability to trace back to an IP address or a location does not provide proof on the person operating behind the account" and that there are many tools that can be used "to cloud traceability".[212]

88. We heard, however, that service providers already have the ability to trace people online. Ms Haugen told us: "Platforms have far more information about accounts than I think people are aware of … It is a question of Facebook's willingness to act to protect people more than a question of whether those people are anonymous on Facebook."[213] Other witnesses, including ministers, agreed that platforms and law enforcement often do have the information and powers to identify people who act illegally online.[214] The House of Commons Petitions Committee noted that the capacity of law enforcement bodies to act was a major factor.[215]

89. Some raised the possibility that verification did not have to be a mandatory process. They suggested numerous system design features that could address the risks posed by anonymous accounts. Clean Up the Internet argued that all users should have the option to verify their account and the option to control the level of interaction they have with

208 For example, Q 73 (Dr Edina Harbinja); written evidence from: Demos (OSB0159); Glassdoor (OSB0033); HOPE not hate (OSB0048)

209 Q 44 (Nancy Kelley)

210 Social Media Platforms (Identity Verification) Bill; written evidence from Siobhan Baillie Member of Parliament for Stroud (OSB0242)

211 Q 194 (Maria Ressa); written evidence from: Antisemitism Policy Trust (OSB0005); Sport and Recreation Alliance (OSB0090)

212 Written evidence from The Football Association, Kick It Out (OSB0234)

213 Q 171 (Frances Haugen)

214 Q 219 (Rt Hon Nadine Dorries MP, Chris Philp MP, Rt Hon Damian Hinds MP)

215 Written evidence from Mr John Carr (Secretary at Children's Charities' Coalition for Internet Safety) (OSB0216); see for example Petitions Committee, *Online Abuse and the Experience of Disabled People*, (First Report, Session 2017–19, HC 759) paras 123–137, which detailed the problems disabled people often have in getting law enforcement to investigate potentially illegal online abuse.

unverified accounts on a sliding scale.[216] Hope not Hate argued that anonymity didn't have to mean a lack of accountability. They noted that anonymous accounts can be banned just the same as identifiable ones. They called for measures to introduce "friction" into the process of creating and removing accounts, requiring accounts to build up evidence of rules adherence and compliance before being able to access the full functionality of a platform.[217] Ms Ressa agreed that the mass creation of new accounts was a core part of the problem. She and Ms Haugen both noted that, for an engagement and advertising-based business model, there was a financial incentive on platforms to facilitate the mass creation of duplicate and disposable accounts and conceal the scale of it.[218]

90. Responding to the evidence we heard, the Secretary of State said the first priority of the draft Bill was to end all online abuse—not just that from anonymous accounts. She recognised the concerns and importance of anonymity to groups like whistleblowers and domestic abuse victims. She indicated she was looking into proposals along the lines of those proposed by Clean Up the Internet around giving people the option to limit their interaction with anonymous or non-verified accounts. Finally, she talked about the importance of traceability in the context of her own experiences of online abuse, noting as mentioned above that platforms often do have access to the information required by law enforcement.[219]

91. **Anonymous abuse online is a serious area of concern that the Bill needs to do more to address. The core safety objectives apply to anonymous accounts as much as identifiable ones. At the same time, anonymity and pseudonymity are crucial to online safety for marginalised groups, for whistleblowers, and for victims of domestic abuse and other forms of offline violence. Anonymity and pseudonymity themselves are not the problem and ending them would not be a proportionate response. The problems are a lack of traceability by law enforcement, the frictionless creation and disposal of accounts at scale, a lack of user control over the types of accounts they engage with and a failure of online platforms to deal comprehensively with abuse on their platforms.**

92. ***We recommend that platforms that allow anonymous and pseudonymous accounts should be required to include the resulting risks as a specific category in the risk assessment on safety by design. In particular, we would expect them to cover, where appropriate: the risk of regulated activity taking place on their platform without law enforcement being able to tie it to a perpetrator, the risk of 'disposable' accounts being created for the purpose of undertaking illegal or harmful activity, and the risk of increased online abuse due to the disinhibition effect.***

93. ***We recommend that Ofcom be required to include proportionate steps to mitigate these risks as part of the mandatory Code of Practice required to support the safety by design requirement we recommended in paragraph 82. It would be for them to decide what steps would be suitable for each of the risk profiles for online services. Options they could consider might include (but would not be limited to):***

a) ***Design measures to identify rapidly patterns of large quantities of identical content being posted from anonymous accounts or large numbers of posts being directed at a single account from anonymous accounts;***

216 Written evidence from Clean up the Internet (OSB0026)
217 Written evidence from HOPE not hate (OSB0048)
218 QQ 193–194 (Maria Ressa); Q 129 (Frances Haugen)
219 Q 291 (Rt Hon Nadine Dorries MP)

b) ***A clear governance process to ensure such patterns are quickly escalated to a human moderator and for swiftly resolving properly authorised requests from UK law enforcement for identifying information relating to suspected illegal activity conducted through the platform, within timescales agreed with the regulator;***

c) ***A requirement for the largest and highest risk platforms to offer the choice of verified or unverified status and user options on how they interact with accounts in either category;***

d) ***Measures to prevent individuals who have been previously banned or suspended for breaches of terms and conditions from creating new accounts; and***

e) ***Measures to limit the speed with which new accounts can be created and achieve full functionality on the platform.***

94. ***We recommend that the Code of Practice also sets out clear minimum standards to ensure identification processes used for verification protect people's privacy—including from repressive regimes or those that outlaw homosexuality. These should be developed in conjunction with the Information Commissioner's Office and following consultation with groups including representatives of the LGBTQ+ community, victims of domestic abuse, journalists, and freedom of expression organisations. Enforcement of people's data privacy and data rights would remain with the Information Commissioner's Office, with clarity on information sharing and responsibilities.***

Societal harm and the role of safety by design

95. As set out in Chapter 3, the spread of disinformation online has been associated with extensive real-world harm, from mass killings to riots and unnecessary deaths during the COVID-19 pandemic. Ms Ressa told us of research which illustrates the risk that inauthentic content poses to society: "cheap armies on social media are rolling back democracy in 81 countries around the world."[220]

96. Disinformation can inflict harm on individuals, as well as groups (called "collective harms") and wider society (called "societal harms"). Collective and societal harms were frequently discussed in relation to disinformation, but they can also refer to the cumulative effect of many instances of forms of undesirable online content and activity.

97. Examples include persistent racism or misogyny online, where the cumulative effect and frequency of attacks can make people feel less safe. Glitch told us that "the current status quo is driving women and particularly marginalised and racialised women and non-binary people to censor themselves online or remove themselves completely."[221] They also highlight how this abuse has implications for democracy, giving abuse as one reason why many women MPs choose not to run for re-election.[222] This intersects with race, with research by Amnesty International analysing tweets that mentioned women MPs in the run up to the 2017 General Election and finding the 20 Black, Asian and Minority Ethnic

220 Q 193
221 Written evidence from Glitch (OSB0097)
222 Written evidence from Glitch (OSB0097)

MPs received 41 per cent of the abuse, despite making up less than 12 per cent of the those in the study.[223] Ms Jankowicz told us:

> "[Misogynistic abuse online] is not just a democratic concern; it is a national security concern, which should make it of interest to everybody in government. It is not just about hurt feelings. It really affects the way our countries operate."[224]

98. The White Paper identified disinformation, misinformation and online manipulation as harms that were "threats to our way of life" and proposed a regulatory regime that focused on the "harms that have the greatest impact on individuals or wider society."[225] When the draft Bill was published however, the Government had noted concerns from stakeholders about the impact this might have on freedom of expression, and it was made clear that it would "cover content and activity that could cause harm to individuals rather than harms to society more broadly."[226] Calls remain to include collective, or societal harms in the scope of the Bill, with Reset noting that "the impact of disinformation is absolutely collective in nature."[227] Others have noted the difficulty in defining and attributing harm, particularly with misinformation, and supported the Government's decision to remove societal harm.[228]

99. BT Group described in their submission the impact on their staff and subcontractors of the 5G conspiracies which led to arson attacks on infrastructure.[229] While there are offline offences for the results of these attacks, it is less clear whether simply sharing an article hypothesising a link between 5G and COVID-19 would meet the threshold of harmful to an individual and it may be genuinely believed by the person sharing it, meaning it may not meet the threshold for the criminal offence. The Government gave the example of people with genuine concerns about vaccines, saying we "have good answers to those questions, and should educate people rather than silencing them, as some have called for in trying to legislate against vaccine misinformation."[230] We heard in one session that removing content could also have the unwanted effect of stoking conspiracies, adding a "censorship dynamic", when it may be better to reduce and inform.[231]

100. Later in this report we discuss new offences proposed by the Law Commission around harm-based or knowingly false communications. These may be helpful in some instances in tackling disinformation, but they also have limitations. The harm-based offence relates specifically to psychological harm, so may not be applicable to vaccine disinformation, and knowingly false means just that—the person sending the communication must know it is untrue. It is also unclear whether the latter offence would assist in cases of disinformation trying to disrupt elections, as the harm is based on psychological or physical harm,

223 Amnesty International, 'Black and Asian women MPs abused more online': https://www.amnesty.org.uk/online-violence-women-mps [accessed 30 November 2021]

224 Q 56

225 Department for Digital, Culture, Media and Sport and The Home Office, *Online Harms White Paper*, CP 59, April 2019, p 54: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/973939/Online_Harms_White_Paper_V2.pdf [accessed 22 November 2021]

226 Department for Digital, Culture, Media and Sport and The Home Office, *Impact Assessment*, April 2021, p 116, p 123: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/985283/Draft_Online_Safety_Bill_-_Impact_Assessment_Web_Accessible.pdf [accessed 22 November 2021]

227 Written evidence from Reset (OSB0138)

228 Written evidence from UKRI Trustworthy Autonomous Systems Hub (OSB0060)

229 Written evidence from BT Group (OSB0163)

230 Written evidence from Department of Digital, Culture, Media and Sport and Home Office (OSB0011)

231 Q 105

rather than harm to an institution, process, state, or society.[232] The Elections Bill, which is currently making its way through Parliament with the intention "to strengthen the integrity of the electoral process"[233], should address the issue of disinformation which aims to disrupt elections.

101. We asked the Government why they had chosen not to include societal harms and were told "this could incentivise excessive takedown of legal material due to the lack of consensus about what might result in societal harm."[234] In our final session the Secretary of State said:

> "If we put societal harms into the Bill, I am afraid we would not be able to make it work. We have looked at it. We have explored it. We have probed it. Legally, it is just a non-starter, I am afraid."[235]

102. The Government instead aims to tackle the problem of disinformation through strengthened media literacy which we consider in Chapter 8 on the role of the regulator. The draft Bill also includes a requirement for Ofcom to establish an advisory committee.[236] It will include representatives of service providers, experts, and platform users, and will provide advice and oversee Ofcom's exercise of their media literacy duties. The Government also established the Cross-Whitehall Counter Disinformation Unit (CDU) at the start of the pandemic and indicated in evidence that it would continue. Carnegie UK Trust expressed concerns about a lack of accountability for the CDU and called for it to be put on a statutory footing.[237]

103. Although we have been unable to see the Government's view on the draft Bill's compliance with the European Convention on Human Rights (ECHR) during our inquiry, we have heard in evidence that it may be open to legal challenge, including on the grounds of interference with people's right to freedom of expression.[238] The inclusion of societal, as well as individual harms, would likely raise further concerns about the extent of this interference. As Reset note in their evidence however, the European Union's Digital Services Act goes further than the draft Bill, "by recognising that the use of 'VLOPs' (Very Large Online Platforms) poses 'systemic risks' to individuals and to societies."[239]

104. As with other types of content, much of the risk of harm from disinformation lies not in the individual pieces of content but in their amplification, in the cumulative effect of large numbers of people seeing them, and in individual people being repeatedly exposed. Mr Chaslot told us that "algorithms create filter bubbles, where some people get to see the same type of content all the time", and that by being exposed to disinformation "over and over again" these people "[get] very disinformed" without realising it.[240]

105. Ms Haugen said that to tackle disinformation and the harm it risks causing, the systems that allow virality and encourage amplification need to be addressed, rather than

232 Law Commission, *Modernising Communications Offences* Law Com No 399, HC 547 (July 2021), pp 224–225: https://s3-eu-west-2.amazonaws.com/lawcom-prod-storage-11jsxou24uy7q/uploads/2021/07/Modernising-Communications-Offences-2021-Law-Com-No-399.pdf [accessed 22 November 2021]
233 Elections Bill [Bill 178 (2021–220]
234 Written evidence from Department of Digital, Culture, Media and Sport and Home Office (OSB0011)
235 Q 287
236 Draft Online Safety Bill, CP 405, May 2021, Clause 98
237 Q 287
238 Written evidence from Gavin Millar QC (OSB0221)
239 Written evidence from Reset (OSB0138)
240 Q 95

focusing on individual pieces of content, which is when you "run into freedom of speech issues".[241] Sophie Zhang, another former Facebook employee, also made a similar point, describing how fact-checking could only be of limited value as often it only took place once content had already been shared widely. She continued:

> "… fundamentally, companies cannot adjudicate every piece of content … that is why my proposals and suggestions have fallen more along the lines of reducing virality in general by reducing reshares—for instance, by requiring people to go to the initial post to reshare a piece of content rather than its being reshared and resharing it again and going to chronological news feed rankings. The problem at hand is not that the content is being made in the first place, but that it is being seen and widely distributed, and people have an incentive to make potentially sensationalist claims."[242]

106. **We recognise the difficulties with legislating for societal harms in the abstract. At the same time, the draft Bill's focus on individuals potentially means some content and activity that is illegal may not be regulated. We discuss this further in Chapter 4.**

107. **The viral spread of misinformation and disinformation poses a serious threat to societies around the world. Media literacy is not a standalone solution. We have heard how small numbers of people are able to leverage online services' functionality to spread disinformation virally and use recommendation tools to attract people to ever more extreme behaviour. This has resulted in large scale harm, including deaths from COVID-19, from fake medical cures, and from violence. We recommend content-neutral safety by design requirements, set out as minimum standards in mandatory codes of practice. These will be a vital part of tackling regulated content and activity that creates a risk of societal harm, especially the spread of disinformation. For example, we heard that a simple change, introducing more friction into sharing on Facebook, would have the same effect on the spread of mis- and disinformation as the entire third-party fact checking system.**

108. **Later in this report we also recommend far greater transparency around system design, and particularly automated content recommendation. This will ensure the regulator and researchers can see what the platforms are doing, assess the impact it has and, in the case of users, make informed decisions about how they use platforms. Online services being required to publish data on the most viral pieces of content on their platform would be a powerful transparency tool, as it will rapidly highlight platforms where misinformation and disinformation is drowning out other content.**

109. **Many online services have terms and conditions about disinformation, though they are often inconsistently applied. We recommend later a statutory requirement on service providers to apply their terms and conditions consistently, and to produce a clear and concise online safety policy. Later, we identify two areas of disinformation—public health and election administration—which are or will soon be covered in the criminal law and that we believe should be tackled directly by the Bill.**

241 Q 193
242 Q 130

110. ***As a result of recommendations made in this report, regulation by Ofcom should reduce misinformation and disinformation by:***

- ***Requiring a consistent enforcement of the providers' own terms and conditions to address user content that is in breach of those terms of service (see Chapter 11);***
- ***Working with the Advertising Standards Authority to address paid content that is in breach of ASA rules (see Chapter 6);***
- ***Use the Safety by Design Code of Practice set out in paragraph 82 to address the spread of misinformation by recommendation algorithms, frictionless sharing of content at scale, use of fake accounts and bots to share malign content and other features that make content viral;***
- ***Publishing codes of practice on Regulated Activity; and***
- ***Improvements to the responsiveness of the complaints processes operated by service providers.***

The Joint Committee that we recommend later in this report should take forward work to define and make recommendations on how to address other areas of disinformation and emerging threats.

111. ***Disinformation and misinformation surrounding elections are a risk to democracy. Disinformation which aims to disrupt elections must be addressed by legislation. If the Government decides that the Online Safety Bill is not the appropriate place to do so, then it should use the Elections Bill which is currently making its way through Parliament.***

112. ***The Information Commissioner, Elizabeth Denham, has stated that the use of inferred data relating to users' special characteristics as defined in data protection legislation, including data relating to sexual orientation, and religious and political beliefs, would not be compliant with the law. This would include, for example, where a social media company has decided to allow users to be targeted with content based on their data special characteristics without their knowledge or consent. Data profiling plays an important part in building audiences for disinformation, but also has legitimate and valuable uses. Ofcom should consult with the Information Commissioner's Office to determine the best course of action to be taken to investigate this and make recommendations on its legality.***

4 Safety duties relating to adults

Illegal content and activity

113. The Bill places duties on companies to tackle "illegal content" online. To do this, it has to define what is meant by "illegal content". The criminal law does not specify "this piece of content is illegal". Rather, it says that someone commits an offence if they (for example) publish an "obscene article".[243] The draft Bill therefore defines "illegal content" as being where the "the use of the words, images, speech or sounds" that make up the content can comprise a "relevant offence", either on their own or in the context of the rest of the site. It can also refer to content where it is the dissemination that is the offence. For example, under the Obscene Publications Act, cited above, it is publication that is the offence.[244]

114. Criminal guilt is determined by law enforcement and the courts. Platforms will not be able to rely on these findings when they draw up and apply their terms and conditions. The draft Bill therefore defines "illegal content" as content where the "service provider has reasonable grounds to believe" that use or dissemination amounts to a "relevant offence".[245] "Reasonable grounds" is a term recognised in law that requires objectivity. For example, a police officer must have reasonable grounds for suspecting someone is committing an offence or about to commit an offence before they use certain powers of arrest.[246]

115. The draft Bill is not concerned with all content whose creation or dissemination might be an offence. It is only concerned with "relevant offences". A "relevant offence" must be terrorism content,[247] CSEA content,[248] an existing offence specified by the Secretary of State in regulations[249] or an offence where the victim is an individual or individuals.[250] Designating an offence under the draft Bill cannot be used to create new offences. Equally, an offence that can be committed online is not considered "illegal content" under the draft Bill unless it falls under one of these categories above.

116. The draft Bill also has a category of "priority illegal content", defined by the Secretary of State in regulations under criteria that include the risk of harm, the severity of that harm and the prevalence of behaviour that could amount to a relevant offence. Defining "priority illegal content" distinguishes those forms of illegal content that providers are required to proactively seek out and "minimise" their presence on a platform, and those where they are only required to mitigate harm arising from it or to take down on report. It is unclear whether the Secretary of State could regulate for non-priority illegal content. The different categories are set out below.

243 Obscene Publications Act 1964
244 Draft Online Safety Bill, CP 405, May 2021, Clause 41(3)
245 Draft Online Safety Bill, CP 405, May 2021, Clause 41(3)
246 Police and Criminal Evidence Act 1984, sections 24(1)(c) and (d)
247 Draft Online Safety Bill, CP 405, May 2021, Clause 41(4)(a); Clause 42 and Schedule 2
248 Draft Online Safety Bill, CP 405, May 2021, Clause 41(4)(b); Clause 43 and Schedule 3
249 Draft Online Safety Bill, CP 405, May 2021, Clause 41(4)(c) also see clause 44
250 Draft Online Safety Bill, CP 405, May 2021, Clause 41(4)(d)

Table 1: Summary of definitions of "Illegal Content" in the draft Bill

Category	Defined by	Platforms required to
CSEA and Terrorism content	The face of the draft Bill	Mitigate and effectively manage risks to individuals. Ensure it is not persistent or prevalent. Proactively minimise its presence and dissemination, as well as swiftly take down when alerted to it.
Priority illegal content	Regulations from the Secretary of State	Mitigate and effectively manage risks to individuals. Proactively minimise its presence and dissemination, as well as swiftly take down when alerted to it.
Other illegal content	Clause 41(4)(d): another offence of which the intended victim is an individual	Mitigate and effectively manage risks to individuals. Swiftly take down such content when alerted to it.
Explicitly excluded	Offences concerning: a) the infringement of intellectual property rights; b) the safety and quality of goods; c) the performance of a service by someone not qualified to perform it.	Not covered by the draft Bill.
Implicitly excluded	Offences not covered above.	Do not appear to be considered illegal content by the draft Bill (but may be considered content harmful to adults or children).

Source: Clauses 3 and 41

Focus of the draft Bill

117. For many witnesses, including those who were generally critical of the draft Bill, tackling illegal content online was their priority. Professor Richard Wilson, Gladstein Distinguished Chair of Human Rights at the University of Connecticut, suggested that the Bill should "go after the low hanging fruit" and "suppress that which is already illegal."[251] Silkie Carlo, Director of Big Brother Watch, agreed saying "the first priority has to be getting a grip on the sheer amount of illegal content online, and criminal communications and criminal abuse …".[252] It was clear from evidence to us, as outlined in Chapter 2, that social media companies are failing to remove content that could amount to a criminal offence. To take just one specific example, Ms Jankowicz told us "technology companies

251 Q 135
252 Q 134

are not doing their due diligence. We receive rape threats; we are told that these do not constitute a violation of terms of service."[253]

118. **We have received a large amount of evidence in our inquiry but very little of it takes issue with the regulation of illegal content. This seems to us to point to a self-evident truth, that regulation of illegal content online is relatively uncontroversial and should be the starting point of the Bill.**

Scope of "illegal content"

119. We heard that what constitutes "illegal content" is not clear on the face of the draft Bill. As noted above, "illegal content" can be specified by the Secretary of State in regulations or is content that amounts to an offence against an individual. Barbora Bukovská, Senior Director of Law and Policy at Article 19, noted that the Bill does not exhaustively list illegal content meaning that the decision on what is covered is, partly, delegated to providers.[254] The British and Irish Law, Education and Technology Association made similar criticisms and noted the potential impact on fundamental rights of inconsistent decision making.[255]

120. The Crown Prosecution Service (CPS) lists at least fourteen criminal offences that can be committed or facilitated by online communication.[256] They include harassment or stalking,[257] making threats to kill,[258] disclosing sexual images without consent[259] and blackmail.[260]

121. We heard from the Department that the list of offences that would be designated priority illegal content under the Secretary of State's powers was already being developed:

> "Although the final list of offences is yet to be confirmed, priority categories of criminal offences are likely to include hate crime, revenge pornography, promoting or facilitating illegal immigration and the sale of illegal drugs and weapons."[261]

122. Examples of offences that are prevalent online and might not be covered under "illegal content" as it stands include some relating to extreme pornography, some elections offences (including the proposed offences relating to exercising undue influence through disinformation about election administration and failure to include information about origins of election material in the Elections Bill[262]) and some hate crime depending on how the regulations are drafted, as we discuss below.

123. The Minister's letter says that the Government intends to include hate crime as priority illegal content. The criminal law in England and Wales has two forms of hate crime. The

253 Q 53
254 Q 135
255 Written evidence from British & Irish Law, Education & Technology Association (OSB0073)
256 Crown Prosecution Service, 'Social Media - Guidelines on prosecuting cases involving communications sent via social media': https://www.cps.gov.uk/legal-guidance/social-media-guidelines-prosecuting-cases-involving-communications-sent-social-media [accessed 22 November 2021]
257 Protection form Harassment Act 1997, sections 2, 2A, 4 or 4A
258 Offences Against the Person Act 1861, section.16
259 Criminal Justice and Courts Act 2015, section 33
260 Theft Act 1968, section 21
261 Letter from Secretary of State 26 November 2021
262 Elections Bill as amended in Committee, Clause 7(1)(6)(f) and Clauses 38 to 43

first recognises hostility[263] to the victim as an aggravating factor where it is based on five characteristics or perceived characteristics: race, religion,[264] sexual orientation, disability, and transgender identity.[265] Where hostility is found the court can increase the sentence above the recommended guideline for the substantive offence. Similar provisions on hate crime apply in Scotland following the passing of the Hate Crime and Public Order (Scotland) Act 2021.

124. The second type of hate crime concerns specific offences such as Incitement to Racial Hatred[266] and Stirring Up Racial Hatred,[267] or Stirring Up Hatred on the Grounds of Religious Belief or Sexual Orientation.[268] Stirring up racial hatred may be particularly relevant to the online environment as it explicitly includes written material that is "threatening, abusive or insulting",[269] and does not necessarily require the offender to demonstrate intent.[270] It remains to be seen which forms of hate crime the Government intends to address.

125. We heard particular support for including offences that disproportionately affect women in the definition of illegal content. Bumble, a dating site, wanted to see the inclusion of offences "relating to image-based abuse, sexual harassment that takes place or is facilitated online, misogynistic content, gendered hate crime, and stalking" adding "it should not be left to the Secretary of State's discretion for services to be obliged to mitigate and tackle harms that disproportionately affect women."[271] Other witnesses agreed.[272] At the same time, not all of these issues are currently the subject of the criminal law. The Law Commission has recently published its review of the law on Hate Crime and is currently reviewing the law on intimate image abuse. We consider these issues below.

126. **We believe the scope of the Bill on illegal content is too dependent on the discretion of the Secretary of State. This downplays the fact that some content that creates a risk of harm online potentially amounts to criminal activity. The Government has said it is one of the key objectives of the Bill to remove this from the online world.**

127. ***We recommend that criminal offences which can be committed online appear on the face of the Bill as illegal content. This should include (but not be limited to) hate crime offences (including the offences of "stirring up" hatred), the offence of assisting or encouraging suicide, the new communications offences recommended by the Law Commission, offences relating to illegal, extreme pornography and, if agreed by Parliament, election material that is disinformation about election administration,***

263 Hostility is not defined in law and so the courts consider it as the ordinary English meaning of the word which the CPS website describes as "ill-will, ill-feeling, spite, prejudice, unfriendliness, antagonism, resentment, and dislike." CPS, 'Hate Crime': https://www.cps.gov.uk/crime-info/hate-crime [accessed 30 November 2021]

264 Crime and Disorder 1998, section 28-32 and Sentencing Act 2020, section 66; The criminal law interprets these characteristics broadly: race includes "race, colour, nationality (including citizenship) or ethnic or national origins" while religion includes a lack of religious belief. The law also covers erroneous assumptions as to a victim's race or religion by the offender, liability for abuse cannot be avoided simply because the offender picked an abusive term that did not actually apply to the victim. Case law has found that Gypsies, Irish Travellers, religious converts and apostates are all protected by the legislation.

265 Sentencing Act 2020, section 66; CPS, 'Hate Crime': https://www.cps.gov.uk/crime-info/hate-crime [accessed 30 November 2021]

266 Public Order Act 1986, part III, sections 18-23

267 Public Order Act 1986, sections 29B-29C

268 Public Order Act 1986, section 29B(1)

269 Public Order Act 1986, section 18(1)

270 Public Order Act 1986, section 18(1)(b)

271 Written evidence from Bumble (OSB0055).

272 For example, Carnegie UK (OSB0095), Refuge (OSB0084), Centenary Action Group (OSB0047).

has been funded by a foreign organisation targeting voters in the UK or fails to comply with the requirement to include information about the promoter of that material in the Elections Bill.

Reform of the Criminal Law

128. Whilst there is broad agreement that the Bill should regulate illegal content, the criminal law in relation to the online world has long been recognised as in need of reform. In 2019 the House of Commons Petitions Committee described the law relating to online abuse as "not fit for purpose".[273]

129. Communications offences have long existed in domestic law and have evolved as methods of communication have changed. The current law is based primarily on two offences: sending a communication with the intent to cause distress or anxiety contrary to Section 1 of the Malicious Communications Act 1988[274] and the improper use of the public communications network under Section 127 Communications Act 2003.[275] Section 1 of the Malicious Communications Act 1988 requires the communication be sent to another, meaning public posts on social media may not be covered. Both offences rely on the communication being "grossly offensive" or otherwise "indecent".[276]

130. Concerns over the utility of the current communications offences led the Government to ask the Law Commission to examine the law and make recommendations for reform. The Law Commission's report *Modernising Communications Offences* was published in July 2021.[277] The Commission concluded the current law potentially both over and under-criminalised social media users. Under-criminalisation occurred because "some abusive, stalking and bullying behaviours, despite causing substantial harm, simply fall through the cracks." Over-criminalisation happened because of the focus of the communications on the content of the message, not the harm it causes or is intended to cause. This, combined with the subjective nature of "grossly offensive" or "indecent", means:

> " … the law criminalises without regard to the potential for harm in a given context. Two consenting adults exchanging sexual text messages are committing a criminal offence, as would be the person saving sexual photographs of themselves to a 'cloud' drive."[278]

131. The Law Commission also concluded that the current criminal law does not adequately police behaviour such as the promotion of self-harm, cyber-flashing or the sending of flashing images to people with epilepsy. We heard compelling evidence on these issues, as outlined in Chapter 2. The Commission concluded that the following new

273 House of Commons Petitions Committee, *Online Abuse and the Experience of Disabled People* (First Report, Session 2017–19, HC 759), para 19

274 Malicious Communications Act 1988, section 1

275 Communications Act 2003, section 127

276 Malicious Communications Act 1988, section 1(1)(a)(i); Communications Act 2003, section 127(1)(a)

277 The Law Commission, 'Modernising Communications Offences: A Final Report': Law Com No 399, HC 547 [accessed 22 November 2021]

278 The Law Commission, 'Modernising Communications Offences: A Final Report', para 1.6: Law Com No 399, HC 547 [accessed 22 November 2021]

offences (which will apply to users rather than service providers or their senior managers) should be introduced:

(1) a new "harm-based" communications offence to replace the offences within Section 127(1) of the Communications Act 2003 and the Malicious Communications Act 1988 which would require intent to commit harm;

(2) a new offence of encouraging or assisting serious self-harm;

(3) a new offence of cyberflashing which would require either intent to cause harm or recklessness as to whether harm was caused and would be defined as a sexual offence;

(4) a new offence of intentionally sending flashing images to a person with epilepsy with the intention to cause that person to have a seizure; and

(5) new offences of sending knowingly false, persistent or threatening communications, to replace section 127(2) of the Communications Act 2003.[279]

132. The Law Commission's report was published after the publication of the draft Bill and the creation of our Committee. On 4 November, following press reports, the Secretary of State confirmed to us that she intends to adopt the Law Commission's recommendations on harm-based and false, persistent and threatening communications and anticipated that the other recommended offences would be taken forward by the relevant departments.[280]

133. The Law Commission's report paid particular attention to the need to protect freedom of expression online and the Commission made significant changes to its proposed offences in response to concerns about this raised during the consultation. These were noted by the House of Lords Communications and Digital Committee in their report.[281]

134. Days before we finalised our report, the Law Commission produced its report on reforming hate crime legislation. It recommended the creation of an offence of "stirring up" hatred or hostility on the grounds of sex, disability and transgender, or gender diverse, identity.[282] We had already considered how these forms of hate should be considered in online safety regulation and recommended that they should be covered, as set out below.

135. **Implementation of the Law Commission's recommendations on reforming the Communications Offences and Hate Crime will allow the behaviour covered by the new offences to be deemed illegal content. We believe this is a significant enhancement of the protections in the Bill, both for users online but also for freedom of expression by introducing greater certainty as to content that online users should be deterred from sharing. We discuss how to address concerns about ambiguity and the context-dependent nature of the proposed harm-based offence through a statutory public interest requirement in Chapter 7.**

136. ***We endorse the Law Commission's recommendations for new criminal offences in its reports, Modernising Communications Offences and Hate Crime Laws. The reports recommend the creation of new offences in relation to cyberflashing, the encouragement***

279 The Law Commission, 'Modernising Communications Offences: A Final Report', para 1.30, para 1.31: Law Com No 399, HC 547 [accessed 22 November 2021]

280 Q 278

281 Communications and Digital Committee, *Free for all? Freedom of expression in the digital age*, (1st Report, Session 2021–22, HL Paper 54)

282 Law Commission, 'Hate Crime Laws: Final Report', Law Com No 402, HC 942 [accessed 9 December 2021]

of serious self-harm, sending flashing images to people with photo-sensitive epilepsy with intent to induce a seizure, sending knowingly false communications which intentionally cause non-trivial emotional, psychological, or physical harm, communications which contain threats of serious harm and stirring up hatred on the grounds of sex or gender, and disability. We welcome the Secretary of State's intention to accept the Law Commission's recommendations on the Communications Offences. The creation of these new offences is absolutely essential to the effective system of online safety regulation which we propose in this report. We recommend that the Government bring in the Law Commission's proposed Communications and Hate Crime offences with the Online Safety Bill, if no faster legislative vehicle can be found. Specific concerns about the drafting of the offences can be addressed by Parliament during their passage.

137. New offences need enforcement resources to be addressed effectively. We heard from T/Commander Clinton Blackburn about the challenges the police face resourcing dealing with economic crime.[283] In Brussels we heard about the overwhelming amount of CSEA content that Interpol assist national enforcement services with. Adding new offences without increasing resources to enforce them will not help victims. When the Law Commission's new offences are brought into law, the police will need greater enforcement resources to ensure that perpetrators are brought to justice.

138. ***The Government must commit to providing the police and courts with adequate resources to tackle existing illegal content and any new offences which are introduced as a result of the Law Commission's recommendations.***

Identifying "illegal content"

139. The criminal law is designed to establish whether or not an individual is guilty of an offence to a high standard of proof following an extensive, and adversarial, legal process. Since an individual's liberty and good name may be at stake, the criminal law requires that all elements of an offence be proved so that jurors are "satisfied so you are sure" or convinced "beyond reasonable doubt" before an offender is found guilty.

140. As set out in paragraph 114, the test for illegal content as defined by the draft Bill[284] is substantially lower than the test applied in the criminal courts.[285]

141. The draft Bill requires the provider to operate "proportionate" systems and processes to mitigate risks of harm, minimise the presence of such content or take it down once reported. We heard that the application of the "reasonable grounds to believe" test by providers would be a challenging task given the complexity of much of the criminal law. We heard concerns from Mr Millar that:

> " … applying the statutory wording of most modern criminal offences to the facts is a difficult and technical exercise. It is one which police, CPS and courts often get wrong. This is both because of the flexibility of the language that is used and because of detailed nature of the drafting in most of our contemporary criminal offences. Criminal offences now, especially terrorism

283 Q 124
284 Draft Online Safety Bill, CP 405, May 2021, Clause 41(9)
285 Written evidence from Gavin Millar QC (OSB0221); 'Satisfied so that you are sure' in England and Wales; 'Beyond reasonable doubt' in Scotland and Northern Ireland

> and CSEA offences, are much more complex than they were 30 or 40 years ago."[286]

142. We heard concerns from some witnesses that this might lead to over-censorship by platforms, looking to avoid being subject to penalties under the draft Bill.[287]

143. The draft Bill addresses the problem of how some illegal content can be identified in practice by requiring Ofcom to publish a Code of Practice on terrorism content and CSEA content. It does not require such a Code of Practice for the wider duties around illegal content.

144. ***We recommend that Ofcom be required to issue a binding Code of Practice to assist providers in identifying, reporting on and acting on illegal content, in addition to those on terrorism and child sexual exploitation and abuse content. As a public body, Ofcom's Code of Practice will need to comply with human rights legislation (currently being reviewed by the Government) and this will provide an additional safeguard for freedom of expression in how providers fulfil this requirement. With this additional safeguard, and others we discuss elsewhere in this report, we consider that the test for illegal content in the Bill is compatible with an individual's right to free speech, given providers are required to apply the test in a proportionate manner that is set out in clear and accessible terms to users of the service.***

145. ***We recommend that the highest risk service providers are required to archive and securely store all evidence of removed content from online publication for a set period of time, unless to do so would in itself be unlawful. In the latter case, they should store records of having removed the content, its nature and any referrals made to law enforcement or the appropriate body.***

Power to designate priority illegal content

146. Given our recommendation that more offences should be listed on the face of the Bill, the question might arise as to whether the power of the Secretary of State to designate priority illegal content is still required.

147. Some of our witnesses expressed concern about the powers to designate priority content in the draft Bill. Ms Carlo described them as a "blank cheque"[288], whilst Prof Wilson said:

> "The question we should always ask of legislation is, 'Would I like this in the hands of my political opponents?' because one day they will come to power."[289]

148. ***We recommend that the Secretary of State's power to designate content relating to an offence as priority illegal content should be constrained. Given that illegal content will in most cases already be defined by statute, this power should be restricted to exceptional circumstances, and only after consultation with the Joint Committee of Parliament that we recommend in Chapter 9, and implemented through the affirmative procedure. The Regulator should also be able to publish recommendations on the***

286 Written evidence from Gavin Millar QC (OSB0221)
287 For example, written evidence from: Big Brother Watch (OSB0136); Global Partners Digital (OSB0194)
288 Q 138
289 Q 138

creation of new offences. We would expect the Government, in bringing forward future criminal offences, to consult with Ofcom and the Joint Committee as to whether they should be designated as priority illegal offences in the legislation that creates them.

Duties to protect adults' online safety

149. For some of our witnesses, the scope of the Bill should be confined to regulating content that is likely to be illegal. The campaign coalition, "Legal to Say, Legal to Type" argued:

> "If something is legal offline, it should be legal online. If the government believes that particular content should be criminalised online, they should address this through parliament and the courts, not big tech."[290]

150. We also heard that this would leave large areas of content and activity that causes risks of harm online unregulated. This would include content and activity that is legislated for offline in the criminal and civil law and potentially give scope for service providers to refuse to act even against content that has been or may be legislated for online.[291]

What lies outside "illegal content"

Characteristics

151. The criminal law in England and Wales covers hate crime arising from hostility to race, religion, disability, sexual orientation and transgender identity. This is significantly different from the protected characteristics under the Equality Act 2010 which prohibits discrimination on the grounds of age,[292] disability,[293] gender reassignment,[294] marriage and civil partnership,[295] race,[296] religion or belief,[297] sex,[298] and sexual orientation.[299] The Equality Act applies in workplaces and to the provision of public and private services.[300] However, the protections it provides against, for example abuse or harassment, do not apply in other offline settings or to private companies online or the users of social media platforms.

Misogyny

152. Much of the harmful online behaviour we heard about from witnesses would not be covered even under the expanded scope of illegal content we recommend. For example, hostility on the grounds of sex does not currently constitute a hate crime, even though

290 Written evidence from Legal to Say, Legal to Type (OSB0049)

291 For example, Q 69; Written evidence from: Sara Khan (Former Lead Commissioner at Commission for Countering Extremism); Sir Mark Rowley (Former Assistant Commissioner (2014–2018) at Metropolitan Police Service) (OSB0034); Centenary Action Group, Glitch, Antisemitism Policy Trust, Stonewall, Women's Aid, Compassion in Politics, End Violence Against Women Coalition, Imkaan, Inclusion London, The Traveller Movement (OSB0047)

292 Equality Act 2010, section 5

293 Equality Act 2010, section 6

294 Equality Act 2010, section 7

295 Equality Act 2010, section 8

296 Equality Act 2010, section 9

297 Equality Act 2010, section 10

298 Equality Act 2010, section 11

299 Equality Act 2010, section 12

300 Equality Act 2010, part 5 and section 29

we heard women face significant abuse online and offline. Edleen John, Director of International Relations and Corporate Affairs and Co-Partner for Equality, Diversity and Inclusion at the FA told us that misogynistic abuse was experienced by women players "from the top flight game—England players—down to the grass roots, so including young women and players in our impairment specific pathways."[301] Ms John told us the volume of recent misogynistic and racist abuse was so high that certain players were blocked from reporting it anymore to social media companies, apparently because the companies had assumed so many complaints from one person must be malicious.[302]

153. The Centenary Action Group, a coalition of groups campaigning to remove barriers to women's political representation, highlighted research that showed women are 27 times more likely than men to be harassed online.[303] Incels, who claim there is a conspiracy preventing some men from having sexual relationships with women, post "violent chatter including celebrating the murder of women and calling for the rights of women to be curtailed."[304]

154. Misogynistic abuse taking place in a workplace, on public transport or by the provider of a service, could lead to action under the Equality Act, a law intended to prevent people being disadvantaged by hostility to their personal characteristics. Many witnesses told us that women's experience of gendered abuse online leads to a "chilling" effect"[305] on their freedom of expression and professional careers, the fear of attracting abuse inducing self-censorship.[306] Prof McGlynn told us her research showed that woman can "experience a more general sense of threat of sexual harassment, violence and abuse from having been abused online which impacts their daily lives and decisions".[307] Other witnesses agreed.[308]

155. Several witnesses told us that gender-based abuse online deterred women from participating in public life. Ms Jankowicz noted as an example that the abuse received by Vice-President Kamala Harris was often sexualised. Ms Jankowicz highlighted the damaging impact on young women's participation in the democratic process.[309] Ms Wick agreed that misogynistic abuse to individuals led to societal harm: "We know that teenage girls are much less likely to speak up on social media for fear of being criticised. That has a very real effect on their ambitions to go into public life."[310] Mr Perrin highlighted

301 Q 29
302 Q 29
303 Written evidence from Centenary Action Group, Glitch, Antisemitism Policy Trust, Stonewall, Women's Aid, Compassion in Politics, End Violence Against Women Coalition, Imkaan, Inclusion London, The Traveller Movement (OSB0047)
304 Written evidence from Centenary Action Group, Glitch, Antisemitism Policy Trust, Stonewall, Women's Aid, Compassion in Politics, End Violence Against Women Coalition, Imkaan, Inclusion London, The Traveller Movement (OSB0047)
305 Written evidence from Advisory Committee For Scotland (OSB0067)
306 Among others Q 55; Written evidence from: Mumsnet (OSB0031); Bumble Inc. (OSB0055); Centenary Action Group, Glitch, Antisemitism Policy Trust, Stonewall, Women's Aid, Compassion in Politics, End Violence Against Women Coalition, Imkaan, Inclusion London, The Traveller Movement (OSB0047); Advisory Committee For Scotland. (OSB0067); HOPE not hate (OSB0048); Dr Kim Barker (Senior Lecturer in Law at Open University); Dr Olga Jurasz (Senior Lecturer in Law at Open University) (OSB0071)
307 Written evidence from Professor Clare McGlynn (Professor of Law at Durham University) (OSB0014)
308 Written evidence from HOPE not hate (OSB0048); Centenary Action Group, Glitch, Antisemitism Policy Trust, Stonewall, Women's Aid, Compassion in Politics, End Violence Against Women Coalition, Imkaan, Inclusion London, The Traveller Movement (OSB0047)
309 Q 56
310 Q 56

the impact of online intimidation on women's participation in political life in Northern Ireland.[311]

Other characteristics

156. We have not heard specific evidence on ageist abuse, abuse against non-religious belief or on the basis of maternity or marital status. At the same time, none of these would be covered by a Bill that focused only on illegal content. As another example, the laws covering the "stirring up of hatred" only apply to race, religion and sexual orientation, and would not apply in the case of other characteristics protected by the law on hate crime or the Equality Act.[312] The Law Commission has, however, recently recommended that the offences should be extended to cover sex, disability and transgender, or gender diverse, identity.[313]

Threshold

157. We discussed above the threshold of proof required to prove a criminal conviction and the test that the draft Bill applies. Although the threshold of proof is lower in the draft Bill, it remains the case that providers would need to have systems and processes in place to take a view as to whether all elements of the offence might reasonably have been committed. This presents particular problems with offences that require proof of "state of mind" (such as intent or malice) on the part of the guilty party, which would include the Law Commission's new harm-based communications offence.[314] For example, the offence of harassment requires that the person in question either knows or "ought to know" that their behaviour constitutes harassment.[315]

158. Much of the behaviour we heard creates risks of harm may not therefore fit easily into a regulatory regime solely focused on illegal content. For example, the Law Commission's new offence of cyberflashing requires the sender either intended to cause distress or sent the image for their personal sexual gratification and was reckless as to whether distress was caused.[316] We heard that the sending of unsolicited penis images was a particular problem for young women and girls, a concern borne out by the findings of Ofsted in its report on sexual abuse in schools.[317] Research suggests such images are frequently not sent with intent to distress or for sexual gratification but that a "large amount of it is a kind of male bonding among their peers. That is why they share unsolicited nude images as well; they want to share among their peer group, 'Oh, we've sent them'."[318]

311 Q 73
312 Public Order Act 1986 ss.
313 Law Commission, 'Hate Crime Laws: Final Report': Law Com No 402, HC 942 [accessed 9 December 2021]
314 The Law Commission, 'Modernising Communications Offences: A Final Report': Law Com No 399, HC 547 [accessed 22 November 2021]
315 Ofsted, 'Research and Analysis: Review of sexual abuse in schools and colleges': www.gov.uk/government/publications/review-of-sexual-abuse-in-schools-and-colleges/review-of-sexual-abuse-in-schools-and-colleges [accessed 10 December 2021]
316 The Law Commission, 'Modernising Communications Offences: A Final Report', para 6.133: Law Com No 399, HC 547 [accessed 22 November 2021]
317 Q 73; Ofsted, 'Research and Analysis: Review of sexual abuse in schools and colleges': www.gov.uk/government/publications/review-of-sexual-abuse-in-schools-and-colleges/review-of-sexual-abuse-in-schools-and-colleges [accessed 10 December 2021]
318 Q 73; for an alternative approach to the offence see written evidence from Professor Clare McGlynn (Professor of Law at Durham University)

159. Hope not Hate, among others, had particular concerns about the large volume of far-right or other extremist propaganda "that does not reach the legal threshold for prosecution."[319] The British Horseracing Authority and Professional Jockeys Association told us that people who had lost money betting on horseraces post huge volumes of vitriolic abuse of their members online. This included material that would fall short of a direct, criminal threat to kill but nonetheless is threatening: "A truly dodgy b*****d. Karma punish you, wish you break your neck and never ride again. A***hole. Idiot."[320]

Content that is harmful to adults

160. To address these cases, the Government introduced a third safety duty in Clause 11. Clause 11 introduces a duty on Category 1 providers to protect adults' online safety, covering content that is "harmful to adults".[321] Like the other safety duties, it requires a specific risk assessment and for service providers to state in the terms of service how they will deal with this type of content, where it is designated as priority content, or identified in the provider's risk assessment. It does not mandate specific outcomes, such as removing or minimising the presence of this content. As written, the draft Bill leaves such decisions to the service providers, although Ofcom has the power to issue a code of practice on compliance.[322]

Defining content that is harmful to adults - Clause 11

161. One of the problems this legislation must grapple with is defining what creates a risk of harm to adults. Clause 11 attempts this in a broad way, and we have heard throughout our inquiry that this will make it difficult to apply, as well as open to legal challenge.[323]

162. As with other categories of content, the Government aims to identify specific types of harmful content by designating them as "priority content that is harmful to adults". These are not listed on the face of the Bill but DCMS suggested they may include the "most prevalent forms of online abuse, together with other harmful material which might disproportionately impact vulnerable users, such as self-harm or suicide content."[324] The draft Bill requires service providers to set out priority content that is harmful to adults that will be "dealt with" by the service in their terms and conditions.[325] It does not specify what is meant by "dealt with".

163. Beyond those designated specifically as "priority content that is harmful to adults", content is considered to be harmful to adults if: "the service has reasonable grounds to believe that the nature of the content is such that there is a material risk of the content having, or indirectly having, a significant adverse physical or psychological impact on an adult of ordinary sensibilities."[326] This uses the same terminology as the definition for content that is harmful to children. The service provider only has to specify how non-

319 Written evidence from: HOPE not hate (OSB0048); Sara Khan (Former Lead Commissioner at Commission for Countering Extremism); Sir Mark Rowley (Former Assistant Commissioner (2014–2018) at Metropolitan Police Service) (OSB0034)
320 Written evidence from British Horseracing Authority (OSB0061)
321 Draft Online Safety Bill, CP 405, May 2021, Clause 11
322 Draft Online Safety Bill, CP 405, May 2021, Clause 29(3)
323 Q 143
324 Written evidence from Department of Digital, Culture, Media and Sport and Home Office (OSB0011)
325 Draft Online Safety Bill, CP 405, May 2021, Clause 11(2)(a)
326 Draft Online Safety Bill, CP 405, May 2021, Clause 46 (3)

priority content that is harmful to adults will be "dealt with" in their terms of service, if it is identified in the service provider's risk assessment.[327]

Power to designate priority harm

164. The difference between priority content that is harmful to adults and non-priority is whether a service provider has to include it in their terms and conditions regardless of whether it is identified in their risk assessment. The Government's justification for the power to designate priority content that is harmful to adults was to allow it to respond to upcoming risks of harm.[328] However, we heard widespread concern about the breadth of this power. Unlike the powers granted in relation to illegal content and content harmful to children, this power isn't bound by any definition or legislation, nor is it restricted to a particular group. The Secretary of State is required to consult Ofcom, but not follow their recommendations.[329] Whilst the initial use of the power requires an affirmative vote in both Houses, subsequent amendment is exercisable by negative statutory instrument, meaning there is no guarantee of parliamentary scrutiny.[330]

Delegation of decision making

165. Another aspect of Clause 11 that concerned witnesses was that it effectively delegates to service providers responsibilities for deciding what is 'harmful' and gives them the authority of the state in doing so. Prof Wilson said that the broad definition of harm "may contribute to the misapplication of the regulatory powers of the Bill"[331]. Journalist Matthew d'Ancona, Editor at Tortoise Media, added that it would involve "handing over the definition of harm" to tech companies, and that the draft Bill "allows huge latitude around what constitutes harm."[332] Ms Bukovská said vagueness would contribute to over removal and the suppression of minority voices.[333] The delegation to service providers of deciding what may be harmful was one of the most frequent concerns we heard. The British and Irish Law, Education and Technology Association said:

> "In being asked to make determinations of legal speech, commercial platforms are being trusted with decisions on what is–or is not–permitted speech. The model proposed therefore rests on trust, placing the operators of platforms in a position where they are directly controlling the speech of an individual."[334]

327 Draft Online Safety Bill, CP 405, May 2021, Clause 11(2)(b)

328 Department for Digital, Culture, Media and Sport, and the Home Office, 'Memorandum to the Delegated Powers and Regulatory Reform Committee', para 158: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/985030/Delegated_Powers_Memorandum_Web_Accessible.pdf [accessed 9 December 2021]

329 Draft Online Safety Bill, CP 405, May 2021, Clause 42(6)

330 Department for Digital, Culture, Media and Sport, and the Home Office, 'Memorandum to the Delegated Powers and Regulatory Reform Committee', para 161–162: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/985030/Delegated_Powers_Memorandum_Web_Accessible.pdf [accessed 9 December 2021]

331 Q 135

332 Q 135

333 Q 137

334 Written evidence from British & Irish Law, Education & Technology Association (OSB0073)

166. On the other side of the argument, there were concerns that Clause 11 as drafted was simply ineffective. Dr Francesca Sobande, Cardiff University, said:

> "Before the Bill is finalised it should include a more detailed explanation of online abuse and harms to appropriately contextualise how online safety is understood, and to ensure that a broad range of forms of online abuse are acknowledged (e.g. including, but not limited to, ableism, ageism, racism, sexism, misogyny, xenophobia, Islamophobia, homophobia, and transphobia)."[335]

167. Mr Ahmed agreed: "You are asking the companies to mark their own homework, but you are also, in one respect, asking them to set their own rules and set the test itself."[336]

168. We were told by several witnesses, including representatives of Facebook and Twitter, as well as the Football Association, that it should be for Parliament to decide what should and should not be covered by regulation.[337] The House of Lords Communications and Digital Committee came to a similar conclusion:

> "If a type of content is seriously harmful, it should be defined and criminalised through primary legislation. It would be more effective—and more consistent with the value which has historically been attached to freedom of expression in the UK—to address content which is legal but some may find distressing through strong regulation of the design of platforms, digital citizenship education, and competition regulation."[338]

169. As we discussed above, the conclusion of the Law Commission's work into reform of the communications offences creates a new harm-based offence for online communication.[339] This may allow for the provisions on content that is harmful to adults to be refined further, as some of this will in future be caught by the duties to act on regulated activity if the Government accepts our recommendations. As we discussed in the previous section however, there are challenges in setting thresholds for service providers to use, which may not be the same as those used by law enforcement. Some content that is harmful to adults will remain outside of scope of legislation.

Replacing Clause 11

170. We heard from those primarily concerned with freedom of expression that the definition of content that is harmful to adults is unsuitably broad, but also from many who welcome it or think it may mean leaving significant causes of harm unaddressed or in the hands of the service providers. While narrowly defining every type of content that could be harmful is appealing in some ways, we have also heard that the Bill needs to be agile and flexible.[340] The legislation will need to be able to adapt to an ever changing societal and technological landscape, but also not be subject to undue political influence.

335 Written evidence from Dr Francesca Sobande (Lecturer in Digital Media Studies at Cardiff University) (OSB0144)
336 Q 16
337 Q 206, Q 249, Q 27
338 Communications and Digital Committee, *Free for all? Freedom of expression in the digital age*, (1st Report, Session 2021–22, HL Paper 54), para 182
339 The Law Commission, 'Modernising Communications Offences: A Final Report': Law Com No 399, HC 547 [accessed 22 November 2021]
340 Written evidence from Department of Digital, Culture, Media and Sport and Home Office (OSB0011)

171. While a tighter definition of content that is harmful to adults may make the statutory requirements easier to fulfil and reduce some of the risk around overzealous moderation, a definitive list also carries risks that it will become out of date (if it is difficult to update) and that it is open to undue political influence (if it is too easy to add to). There is however an existing body of law that taken together may provide a reasonable estimation of what types of activity society may agree are potentially harmful, even if the relevant offences are not always directly applicable online. For example, while hate crime legislation only protects a limited number of groups, the characteristics named by the Equality Act and hate crime legislation together reflect those considered to warrant protection in civil society. It may be possible therefore to refine the definition of content that is harmful, in relation to existing law, even if not linking to specific offences. Mr d'Ancona agreed that there may already be a basis that could be used:

> "Except for free speech absolutists, there are plenty of perfectly legitimate, legislated, in precedent or in common law, restrictions on speech that now really need to be put into action. The problem is a legislative structure that matches the technological revolution rather than identifying speech that is harmful."[341]

172. Sanjay Bhandari, Chair of Kick It Out, said:

> "Sometimes people think that the legal part feels like a big grey area, and how do you legislate for that? Actually, we have some jurisprudence from elsewhere. There is a civil law cause of action in conspiracy, and conspiracy has two limbs: if lawful means conspiracy, or unlawful means conspiracy. You can conspire by lawful means and be held to be civilly responsible for that. That goes back to the 1940s and was clarified in the Lonrho v Fayed litigation in the late 1980s/early 1990s, and there has been a rich history of that economic tort.
>
> There are two key defining characteristics. Was harm experienced in this case? Yes, tick, harm was experienced. Was it intended? Was it aimed? If you send a monkey emoji to a footballer, that is pretty clearly intended to cause harm.
>
> We have precedents, we have jurisprudence. We just need to look at that jurisprudence from elsewhere and bring that under harmful content, because I think it is achievable."[342]

173. As well as the body of criminal law that may not have been designed to be applied online but that represents well understood causes of harm, this may also draw on the Equality Act, electoral law and legitimate reasons for interference in freedom of expression as defined by the ECHR, as well as its protections for free and fair elections. The BBFC, who regulate film and video in the UK, described in evidence to us the risks of harm they consider in detail when classifying video content, and how those standards are not applied online.[343] Similarly, the Communications Act 2003 references the characteristics described in the Charter of Fundamental Rights of the European Union when defining harmful content for video sharing platforms, albeit with the higher threshold of incitement. Using this extends beyond hate crime legislation and covers characteristics more similar to those in the Equality Act.

341 Q 141
342 Q 27
343 Written evidence from BBFC (OSB0006)

174. **Clause 11 of the draft Bill has been widely criticised for its breadth and for delegating the authority of the state to service providers over the definition of content that is harmful and what they should do about it. We understand its aims and that the Government intended it primarily as a transparency measure over something companies are already doing. As drafted, however, it has profound implications for freedom of speech, is likely to be subject to legal challenge and yet may also allow companies to continue as they have been in failing to tackle online harm.**

175. **We agree that the criminal law should be the starting point for regulation of potentially harmful online activity, and that safety by design is critical to reduce its prevalence and reach. At the same time, some of the key risks of harm identified in our evidence are legislated for in parts of the offline world, but not online, where the criminal law is recognised as needing reform, or where drafting that makes sense in the context of determining individual guilt would allow companies to challenge attempts to make them act. A law aimed at online safety that does not require companies to act on misogynistic abuse or stirring up hatred against disabled people, to give two examples, would not be credible. Leaving such abuse unregulated would itself be deeply damaging to freedom of speech online.**

176. ***We recommend that Clause 11 of the draft Bill is removed. We recommend that it is replaced by a statutory requirement on providers to have in place proportionate systems and processes to identify and mitigate reasonably foreseeable risks of harm arising from regulated activities defined under the Bill. These definitions should reference specific areas of law that are recognised in the offline world, or are specifically recognised as legitimate grounds for interference in freedom of expression. For example, we envisage it would include:***

- ***Abuse, harassment or stirring up of violence or hatred based on the protected characteristics in the Equality Act 2010 or the characteristics for which hatred may be an aggravating factor under Crime and Disorder Act 1998 and section 66 of the Sentencing Act 2020;***[344]
- ***Content or activity likely to cause harm amounting to significant psychological distress to a likely audience (defined in line with the Law Commission offence);***
- ***Threatening communications that would lead a reasonable person to fear that the threat might be carried out;***
- ***Knowingly false communications likely to cause significant physical or psychological harm to a reasonable person;***
- ***Unsolicited sending of pictures of genitalia;***
- ***Disinformation that is likely to endanger public health (which may include anti-vaccination disinformation);***
- ***Content and activity that promotes eating disorders and self-harm;***
- ***Disinformation that is likely to undermine the integrity and probity of electoral systems.***

344 Age, disability, gender reassignment, marriage and civil partnership, pregnancy and maternity, race, religion or belief, sex, sexual orientation and transgender status.

177. ***As with the other safety duties, we recommend that Ofcom be required to issue a mandatory code of practice to service providers on how they should comply with this duty. In doing so they must identify features and processes that facilitate sharing and spread of material in these named areas and set out clear expectations of mitigation and management strategies that will form part of their risk assessment, moderation processes and transparency requirements. While the code may be informed by particular events and content, it should be focused on the systems and processes of the regulated service that facilitates or promotes such activity rather than any individual piece of content. We envisage that this code would include (but not be limited to):***

- ***the moderation of user generated content to cover the use of AI for moderation;***
- ***the appropriate thresholds for human oversight;***
- ***the level of expertise needed for human moderation;***
- ***dedicated teams for election periods and involve relevant bodies–with planned circuit breakers;***
- ***the use of fact checking in proportion to reach and risk;***
- ***a transparency requirement on the top 20 viral messages, published on a monthly basis;***
- ***user control over their curation, including being joined to groups without permission; and***
- ***targeting through protected characteristics and or political affiliation.***

178. **Accepting these recommendations would create a narrower, but stronger, regulatory requirement for service providers to identify and mitigate risks of harm in the online world that may not necessarily meet the criminal thresholds, but which are based on the same criteria as those thresholds, indicating that society has recognised they are legitimate reasons to interfere with freedom of speech rights. It would place these areas on the face of the Bill and remove the broad delegation of decisions on what is harmful from service providers.**

179. **We recognise that the broad power to define new types of content that is harmful to adults in secondary legislation was a key concern with Clause 11. We recognise that there will need to be the ability to amend what is covered by this proposal to ensure that the Bill is futureproofed. At the same time, it needs to be tightly proscribed and subject to active parliamentary scrutiny and review.**

180. ***We recommend that additions to the list of content that is harmful should be by statutory instrument from the Secretary of State. The statutory instrument should be subject to approval by both Houses, following a report from the Joint Committee we propose in Chapter 9. Ofcom, when making recommendations, will be required by its existing legal obligations to consider proportionality and freedom of speech rights. The Joint Committee should be specifically asked to report on whether the proposed addition is a justified interference with freedom of speech rights.***

Accessibility and consistency of terms and conditions

181. We have heard throughout this inquiry that when it comes to the types of content and activity that present a risk of harm we have been discussing, many service providers already have terms and conditions, or terms of service, that prohibit them but they are poorly understood and inconsistently applied and enforced. Mr Ahmed told us that service providers already have policies banning vaccine disinformation but do not enforce them.[345] Mr Russell described a "veneer of useability about most platforms". He continued:

> "As soon as you get beneath that veneer to the … pages of the average terms and conditions, or whatever it is, it is a mystery to most people. They are a great example of how the user experience needs to be simplified for all, so that it is better understood and more readily understood by those who need to understand it."[346]

Others agreed that the size of most service providers' terms and conditions meant they were unlikely to be read. Mr Harrison told us accessibility would be the "key point for those with learning disabilities".[347] Ms Pelham suggested a one-page synopsis, accessible to the average reader, should be required.[348] As well as their complexity, there have also been cases of companies applying their terms and conditions inconsistently deliberately. For instance, Facebook has been reported to "whitelist" high profile accounts, such as celebrities and politicians, essentially exempting them from their terms and conditions altogether.[349] We discuss this further in the later chapter on transparency (Chapter 9).

182. Dame Melanie Dawes DCB, Chief Executive of Ofcom talked about the important role clear terms and conditions could play:

> "This is about terms and conditions that make sense to the user and are not just about your assent; they are about you being given information that helps you to manage your life online and manage risks to you. They are a commitment from the company to you as to what you can expect."[350]

183. **The original Clause 11 in the draft Bill, in common with the other safety duties, required providers to produce clear and accessible terms of service and enforce them consistently in relation to content harmful to adults. While we have recommended a narrower but stronger regulatory requirement for service providers to identify and mitigate risks of harm, the requirements for transparency, clarity and consistency are vital to ensuring users are well informed about how platforms promote content to them and what protections they can expect. Clear, concise and fully accessible terms will allow users to make informed choices.**

184. ***We recommend that the Bill mandates service providers to produce and publish an Online Safety Policy, which is referenced in their terms and conditions, made accessible for existing users and made prominent in the registration process for new users. This***

345 Q 16
346 Q 68
347 Q 68
348 Q 68
349 'Facebook Says Its Rules Apply to All. Company Documents Reveal a Secret Elite That's Exempt' *Washington Post (13 September 2021):* https://www.wsj.com/articles/facebook-files-xcheck-zuckerberg-elite-rules-11631541353 [accessed 30 November 2021]
350 Q 258

Online Safety Policy should: explain how content is promoted and recommended to users, remind users of the types of activity and content that can be illegal online and provide advice on what to do if targeted by content that may be criminal and/or in breach of the service providers' terms and conditions and other related guidelines.

185. ***The Online Safety Policy should be produced in an accessible way and should be sent to all users at the point of sign up and, as good practice suggests, at relevant future points. "Accessible" should include accessible to children (in line with the Children's Code), where service providers allow child users, and accessible to people with additional needs, including physical and learning disabilities. Ofcom should produce a Code of Practice for service providers about producing accessible and compliant online safety policies and on how they should make them available to users to read at appropriate intervals in line with best practice (for example, when the user is about to undertake an activity for the first time or change a safety-relevant setting).***

Online fraud

186. Some of the most prevalent illegal content risking harm to adults that we heard about was fraud, which was reported to be the single biggest single crime in the UK last year[351]—an estimated £2.3bn was lost by victims to fraud over the past year alone.[352] As well as financial detriment, victims suffer psychological harms. For example, 28 per cent of adults reported feeling depressed after being scammed.[353] According to Action Fraud, 85 per cent of scams rely on the internet in some way.[354]

187. We received evidence on a number of methods of online fraud which were of varying sophistication. For example, we heard from insurance companies that fraudsters can make copies of legitimate insurance providers' websites and pay for them to appear at the top of search results. Customers can inadvertently enter the copy website and disclose their details to criminals.[355]

188. Martin Lewis, founder of Money Saving Expert and the Money and Mental Health Policy Institute, highlighted investment scams, which promote fake investment opportunities promising high returns. He raised the case of a man who had lost £19,000 to such a scheme,[356] and a grandmother who put the money their grandchild inherited from a deceased parent into such a scam.[357] We also heard about romance scams, where fraudsters create fake accounts on dating sites and develop relationships with victims. Once victims are emotionally invested, fraudsters pretend to be in urgent need of money and request assistance. We heard that a total of £21.2 million was lost to romance scams in 2020 (an increase of 17 per cent from 2019), affecting nearly 9,000 reported victims.[358]

189. The Office of the City Remembrancer, City of London Corporation, told us that compared to 2019, there has been a significant increase in reports of fraud facilitated

351 Oral evidence taken before the Work and Pensions Committee, 6 January 2021 (Session 2019–2021), Q 223 (Graeme Biggar)
352 Written evidence from Which? (OSB0115)
353 Written evidence from Money and Mental Health Policy Institute (OSB0036)
354 Written evidence from Paul Davis (Director of Fraud at TSB Bank Plc) (OSB0164)
355 See written evidence from: Keoghs LLP (OSB0003), Somerset Bridge Group Ltd (OSB0004), Quilter (OSB0024), the Association of British Insurers (OSB0079), and M&G PLC (OSB0176)
356 Q 110 (Martin Lewis)
357 Q 112 (Martin Lewis)
358 Written evidence from UK Finance (OSB0088)

through online channels: online shopping and auction fraud (43 per cent increase), romance scams (15 per cent increase), and investment fraud (16 per cent increase).[359]

190. In the White Paper the Government had initially said fraud would not be covered by online safety regulation However, on publication of the draft Bill the Government announced it would be included[360] and the Prime Minister confirmed in his July 2021 appearance before the Commons Liaison Committee that "one of the key objectives of the Online Safety Bill is to tackle online fraud".[361]

191. Nonetheless, concerns remain over whether the provisions in the draft Bill will tackle fraud effectively. Notably, the draft Bill considers fraud "illegal content" rather than "priority illegal content" or explicitly mentioning it in the same vein as CSEA and terrorism content. This means that providers will have a duty to remove the content, but only on being notified of it by users.[362] We heard that this reactive rather than proactive approach is likely to be ineffective at dealing with fraud, given that people may only become aware and make a report after a crime has taken place. Designating fraud as "priority illegal content" would place a duty on providers to minimise the risk that the content would appear on their service in the first place—several witnesses argued that this would be a more effective provision in the fight against online fraud.[363]

192. The CMA was particularly concerned to see fraud designated as "priority illegal content" to ensure that the Bill does not undermine existing consumer legislation. The CMA outlined its interpretation of the Consumer Protection from Unfair Trading Regulations 2008 as requiring platform operators take proactive steps to minimise economically harmful content on their platforms, rather than simply responding to it when it is reported.[364] The CMA therefore expressed concern that, were fraud not designated "priority illegal content", "people will see that slightly narrower duty [reactive rather than proactive] and think that it supersedes the existing law, supplants it and therefore weakens it."[365] Guy Parker, the Chief Executive of the Advertising Standards Authority also told us: "I think there are good arguments for extending the scope of the Online Safety Bill to cover financial scams."[366]

193. UK Finance argued that fraud could be explicitly mentioned in the Bill in the same way CSEA and terrorism offences are and suggested amendments to achieve this.[367] This would mean that the Bill itself, rather than secondary legislation, would ensure that platforms were required to proactively prevent fraudulent content from appearing.

359 Written evidence from Office of the City Remembrancer, City of London Corporation (OSB0148)
360 HC Deb, 12 May 2021, UIN HCWS12
361 Oral evidence taken before the Liaison Committee, 7 July 2021 (Session 2019–2021), Q 79 (The Prime Minister)
362 Draft Online Safety Bill, CP 405, May 2021, Clause 9(3)
363 See written evidence from: the Financial Conduct Authority (OSB0044), Office of the City Remembrancer, City of London Corporation (OSB0148), Association of British Insurers (OSB0079), Barclays Bank (OSB0106), Which? (OSB0115) and Money and Mental Health Policy Institute (OSB0036). TSB Bank do not call for fraud to be made "priority illegal content", but do stress the importance of measures preventing fraud from appearing on platforms at all (OSB0164)
364 See written evidence from: the Competition and Markets Authority (OSB0160) and Q 119. The CMA interpret 'due diligence' in the Consumer Protection from Unfair Trading Regulations 2008 as placing a proactive duty on providers, but makes clear that this is its own interpretation and is subject to challenge.
365 Q 119
366 Q 118
367 Written evidence from UK Finance (OSB0088)

194. ***We welcome the inclusion of fraud and scams within the draft Bill. Prevention must be prioritised and this requires platform operators to be proactive in stopping fraudulent material from appearing in the first instance, not simply removing it when reported. We recommend that clause 41(4) is amended to add "a fraud offence" under terrorism and child sexual exploitation and abuse offences and that related clauses are similarly introduced or amended so that companies are required to proactively address it. The Government should consult with the regulatory authorities on the appropriate offences to designate under this section. The Government should ensure that this does not compromise existing consumer protection regulation.***

195. ***The Bill must make clear that ultimate responsibility for taking action against criminal content remains with the relevant regulators and enforcement bodies, with Ofcom reporting systemic issues relating to platform design and operation—including in response to "super complaints" from other regulators. The Bill should contain provisions requiring information-sharing and regulatory cooperation to facilitate this.***

5 Protection of Children

Definition of content harmful to children

196. As set out in Chapter 2, one of the key objectives of the draft legislation is to ensure a higher level of protection for children than adults. The draft Bill's definition of content that is harmful to children has three elements set out in Table 1 below. The broad category of "undesignated" content harmful to children is set out in Clause 45. It concerns content that the provider has reasonable grounds to believe may have or indirectly have "a significant adverse physical or psychological impact on a child of ordinary sensibilities."[368] The remainder of the clause requires providers to consider factors such as children's characteristics, the means of dissemination of the content, the impact on children of different age groups and so forth.[369]

Table 2: Categories of content harmful to children in the draft Bill

Category	Defined by	Duty on provider
Primary priority content	Regulations made by the Secretary of State	Use proportionate systems and processes to prevent children of any age from encountering it.
Priority content	Regulations made by the Secretary of State	Use proportionate systems and processes to protect children from age groups judged to be at risk of harm from encountering it
Undesignated content	Clause 45 (see above)	Use proportionate systems and processes to protect children from age groups judged to be at risk of harm from encountering it, if such a risk of harm has been identified in the most recent children's risk assessment.

197. The provisions relating to content that is harmful to children have been criticised by service providers and rights groups as requiring providers to make fine judgements about individual pieces of content. Some described the provisions as "overly broad", expressing concern about a possible impact on educational material and children's right to access information.[370] There was also some criticism for the lack of clarity in what will be covered by the "primary priority" and "priority" content duties.[371]

198. As we set out above, we have heard arguments that it is important for the Bill to include specifics about the types of harm and content that it will cover. At the same time, as set out in Chapters 2 and 3, we heard compelling evidence about the breadth of content and activity that creates a risk of harm that children are exposed to and concerns that too specific a definition will not keep pace with the changing online world. The draft

368 Draft Online Safety Bill, CP 405, May 2021, Clause 10(3)

369 Draft Online Safety Bill, CP 405, May 2021, Clause 10(5–8)

370 Written evidence from Google: (OSB0175)(OSB0175); TikTok (OSB0181); Global Partners Digital (OSB0194); Microsoft (OSB0076); Wikimedia UK (OSB0169)

371 Written evidence from Care (OSB0085)

Bill requires providers to have in place proportionate systems and processes to protect children from encountering such material, where there is a risk of harm resulting from them doing so as identified in the provider's risk assessment. That could include effective content moderation, it could also include design features like content warnings, safe search features, algorithmic tweaks, or age assurance or verification.[372] There is already a direct precedent in regulation for a similar requirement, the Video Sharing Platform Regulations, which require providers to prevent children accessing material "that might impair the physical, mental or moral development of under 18s."[373]

199. We heard that the definition of harm to children would benefit from being tightened. In particular, the Government has decided not to use the established formula of a "reasonable person", instead going with a relatively novel formula of a "child of ordinary sensibilities". In devising their proposed reforms to the Communications Offences, the Law Commission rejected "universal standards" such as a reasonable person test for establishing harm to an individual. They took the view that there are too many characteristics that may be relevant to whether communications may be harmful to apply such a test, as Mr Millar put it: "In reality there is no such person. Some people are more robust and resilient than others."[374]

200. The draft Bill attempts to recognise this in Clause 10(4) which requires providers to assume that a person encountering content has characteristics or is a member of a group that might reasonably be expected to mean they are particularly affected by content. The intent here appears to be to cover for example targeted abuse on the basis of someone's appearance. However, it could also be read as requiring providers to consider all possible audiences and act on behalf of the most vulnerable, even where that group might be very small, such as those with an uncommon phobia.

201. **The test the Law Commission arrived at for their harm-based offence was "likely to cause harm to a likely audience". We believe this is a better way of ensuring that service providers consider those who may be harmed or impacted by content or activity on a platform than the "person of ordinary sensibilities" test in the draft Bill. Having a single test for a key category of illegal content and for regulated content and activity harmful to children reduces regulatory burden and improves consistency. Online providers generally have a good understanding of their audience. Where their platform allows users to target content at particular people it would require service providers to consider how the design of their systems might be used to create or mitigate harm.**

202. ***Recognising the key objective of offering a higher level of protection for children than adults, we support the inclusion of a broad definition of content that is harmful to children. At the same time, we believe the definition should be tightened. We recommend that Clauses 10(3) to (8) are revised. Content and activity should be within this section if it is specified on the face of the Bill, in regulations or there is a reasonably foreseeably risk that it would be likely to cause significant physical or psychological distress to children who are likely to encounter it on the platform.***

372 Age assurance refers to any system of age checking and estimation. The Age Verification Providers Association (AVPA) makes the distinction between "age assurance" and "age verification": age assurance is a broad term for different methods of discerning the age or age-range of an online user; age verification is a subset of that with more stringent methods and a higher level of accuracy and confidence in the age or age-range of that user.

373 Ofcom, *Video Sharing Platform Guidance*, (6 October 2021): https://www.ofcom.org.uk/__data/assets/pdf_file/0015/226302/vsp-harms-guidance.pdf [accessed 30 November 2021]

374 Written evidence from Gavin Millar QC (OSB0221)

203. ***As with other duties, we recommend that key, known risks of harm to children are set out on the face of the Bill. We would expect these to include (but not be limited to) access to or promotion of age-inappropriate material such as pornography, gambling and violence material that is instructive in or promotes self-harm, eating disorders or suicide, and features such as functionality that allows adults to make unsupervised contact with children who do not know them, endless scroll, visible popularity metrics, live location, and being added to groups without user permission.***

204. ***We recognise the concerns that, without proper guidance, service providers might seek to place disproportionate age assurance measures in place, impacting the rights of both children and adults. We recommend that Ofcom be required to develop a mandatory Code of Practice for complying with the safety duties in respect of children. Ofcom should be required to have regard to the UN Convention on the Rights of the Child (in particular, General Comment No. 25 on children's rights in relation to the digital environment), the Information Commissioner's Office's Age Appropriate Design Code, and children's right to receive information under the ECHR when drawing up that Code.***

Alignment with the Age Appropriate Design Code

205. The Age Appropriate Design Code (AADC) came into force in September. It sets out 15 standards that online services need to follow to meet their obligations to protect children's data online. The AADC covers all Internet Society Services that collect personal data and are likely to be accessed by children.[375]

206. There are synergies and overlaps between the AADC and the draft Bill. Both cover social media platforms and search engines. Both have the aim of protecting children online. Both include a test of whether a service is "likely to be accessed by children". Given the importance of data to the design-driven risks of harm we identify in Chapter 2, the protections required for children's data by the AADC are a key part of online safety. Our witnesses did not consider it a coincidence that Google and other service providers introduced a raft of new safety features aimed at children, including a widely welcomed decision to turn off auto-play by default on YouTube for Kids, just a few weeks before the AADC came into force.[376] TikTok and Snap Inc. both told the US Senate that they welcomed the ADDC.[377]

207. The Information Commissioner's Office (ICO) says that it applies the test in section 123 of the Data Protection Act 2018 on the balance of probabilities—is it more likely than not that children will access an Internet Society Service?[378] The test in the draft Bill is more complex. It must be possible for children to access all or a part of a service and "the child

375 Information Commissioner's Office (ICO), 'Introduction to the Age Appropriate Design Code': https://ico.org.uk/for-organisations/guide-to-data-protection/ico-codes-of-practice/age-appropriate-design-code/. [accessed 18 November 2021]. ISSs are defined as "any service normally provided for remuneration, at a distance, by electronic means and at the individual request of a recipient of services."

376 Q 53 (Ian Russell); Q 62 (Izzy Wick); Google, 'Giving kids and teens a safer experience online': https://blog.google/intl/en-in/company-news/technology/giving-kids-and-teens-safer-experience-online/ [accessed 18 November 2021];

377 U.S. Senate Committee on Commerce, Science and Transportation: Hearings, 'Protecting Kids Online: Snapchat, TikTock and YouTube': https://www.commerce.senate.gov/2021/10/protecting-kids-online-snapchat-tiktok-and-youtube [accessed 9 December 2021]

378 ICO, 'Age Appropriate Design Code, A Code of Practice for Online Services': https://ico.org.uk/for-organisations/guide-to-data-protection/ico-codes-of-practice/age-appropriate-design-a-code-of-practice-for-online-services/services-covered-by-this-code/ [accessed 18 November 2021]

user condition" is met in respect of that service or part of it. The "child user condition" is whether the service (or part of a service) attracts or is likely to attract an undefined "significant number of child users".[379]

208. Common Sense argued in their written evidence that the scope of the Online Safety Bill in respect of children should not be more restricted than the AADC and the "likely" test should be the same in both. They drew on the experience of the US Children's Online Privacy Protection Act to argue that the provision in the draft Bill is too weak. They felt it gives too much scope for providers to argue that there is insufficient evidence of a "significant number" of child users to fall within the provisions relating to children—an argument Microsoft used in their written evidence.[380] They also argued that the distinction would be problematic for businesses, seeking to comply with two different standards.[381]

209. Greater alignment between the two sets of regulation may also be welcomed by providers and aid compliance. TechUK, for example, were concerned on the burdens for smaller businesses of navigating two different sets of regulation. They called on the regulators to issue a joint statement to address any inconsistencies.[382] Google and Microsoft, in their written evidence, urged us to consider how the two regulations could be best aligned.[383]

210. Ofcom told us that the scope of the Bill "did not need to be identical" for them to cooperate with the ICO but agreed with the principle of alignment: "What Elizabeth Denham [the Information Commissioner] and I would both like to do over the next few years is create one set of requirements that may be operated by two different regulatory systems and sets of legal powers, but that as far as possible are asking the same things."[384]

211. ***We recommend that the "likely to be accessed by children" test in the draft Online Safety Bill should be the same as the test underpinning the Age Appropriate Design Code. This regulatory alignment would simplify compliance for businesses, whilst giving greater clarity to people who use the service, and greater protection to children. We agree that the Information Commissioner's Office and Ofcom should issue a Joint Statement on how the two regulatory systems will interact once the Online Safety Bill has been introduced. They should be given powers to cooperate on shared investigations, with appropriate oversight.***

Beyond user-to-user and search

212. The draft Bill covers user-to-user services and search services.[385] It includes messaging apps (such as Facebook Messenger, WhatsApp, Telegram, etc) but excludes "one to one aural communication", SMS messages and email, as well as other types of content, discussed later.[386]

213. One of the major themes in the evidence we received was the potentially limiting nature of the focus on user-to-user and search. For example, Schillings LLP raised the role

379 Draft Online Safety Bill, CP 405, May 2021, Clause 26
380 Written evidence from Microsoft (OSB0076)
381 Written evidence from Common Sense (OSB0018)
382 Written evidence from techUK (OSB0098)
383 Written evidence from: Google (OSB0175); Microsoft (OSB0076)
384 Q 259 (Dame Melanie Dawes)
385 Draft Online Safety Bill, CP 405, May 2021, Clause 2
386 Draft Online Safety Bill, CP 405, May 2021, Clause 39(2)

of file transfer services, webhosts and "cyber lockers".[387] On the other hand, Ofcom and the Secretary of State, as well as some industry bodies whose members fall on both sides of the draft Bill's scope, all made the point that the Bill's scope could not be unlimited and still be effective.[388]

214. At the same time, concerns were raised by children's rights groups about services likely to be accessed by children that fall outwith the scope of the draft Bill, but nonetheless contain content and activity that creates a risk of harm to children. As John Carr OBE, Secretary at Children's Charities' Coalition on Internet Safety, put it: "What counts is not the nature of the platform or the environment but the nature of the likely harm irrespective of how or where it appears on a child's screen or how or by whom it was put there."[389] App stores who mis-advertise age restrictions and allow easy access to age-restricted apps were one case raised by 5Rights.[390] Many organisations raised repeatedly the issue of commercial pornography.[391]

Pornography

215. Professor Sonia Livingstone, Department of Media and Communications at LSE, described including pornography on sites that do not host user-to-user content, and are therefore not covered by the draft bill, as the "number one concern of children, and indeed many adults".[392] The evidence we received would appear to support this. Submissions from media safety groups, children's rights groups, campaigners on violence against women and girls, and others strongly supported either bringing such material within scope of the Bill or implementing the never-commenced Part 3 of the Digital Services Act, which the draft Bill would repeal.[393]

216. The day after launching the draft Bill, the then-Secretary of State told the DCMS Select Committee: "On the issue of commercial pornography, the biggest risk is kids stumbling across it but there is a greater risk from social media and user-generated content … I believe that the preponderance of commercial pornography sites have user-generated content on them, so most of them will be in scope."[394]

217. Most of those who submitted evidence accepted that the draft Bill would capture the largest sites hosting free pornography, as most host user-to-user content as defined in the draft Bill. However, there was widespread concern in the evidence cited above that

387 Written evidence from Schillings International LLP (OSB0183)

388 Q 284 (Rt Hon Nadine Dorries MP), Q 259 (Dame Melanie Dawes), for example, Written evidence from UK Interactive Entertainment (OSB0080)

389 Written evidence from Mr John Carr (Secretary at Children's Charities' Coalition on Internet Safety) (OSB0167)

390 5Rights Foundation, *Systemic breaches to the Age Appropriate Design Code*: https://5rightsfoundation.com/uploads/Letter_5RightsFoundation-BreachesoftheAgeAppropriateDesignCode.pdf [accessed 9 December 2021]

391 Q 62 (Izzy Wick)

392 Q 69

393 For example, Written evidence from: BBFC (OSB0006); Professor Clare McGlynn (Professor of Law at Durham University) (OSB0014); Barnardo's (OSB0017); Office of the Children's Commissioner (OSB0019); The Naked Truth Project (NTP) (OSB0023); All-Party Parliamentary Group on Commercial Sexual Exploitation (OSB0037); COST Action CA16207 - European Network for Problematic Usage of the Internet (OSB0038); Advisory Committee For Scotland (OSB0067); Dr Elly Hanson (Clinical Psychologist & researcher at I am independent) (OSB0078); Care (OSB0085); CEASE (Centre to End All Sexual Exploitation) (OSB0104); The Age Verification Providers Association (OSB0122); Parent Zone (OSB0124); Baroness Floella Benjamin, DBE (OSB0161); BT Group (OSB0163); Independent Schools Council (OSB0187); Dame Margaret Hodge (Member of Parliament for Barking and Dagenham at House of Commons) (OSB0201); NSPCC (OSB0109)

394 Oral evidence taken before the Digital, Culture, Media and Sport Committee, 13 May 2021 (Session 2021–22), QQ 26 and 27 (Rt Hon Oliver Dowden MP)

they could evade the draft Bill by removing that functionality—as Pornhub largely did in December 2020 following criticism of its hosting of illegal content—or that children would simply move to sites not covered by the draft Bill.[395] As the Centre to End All Sexual Exploitation told us, the pornography site XVideos received a boost in traffic after Pornhub introduced safeguarding reforms. They and the British Board of Film Categorisation stressed that standards had to be universal and strictly enforced to be effective. Otherwise, they would simply divert traffic away from the sites who do introduce age assurance/verification.[396]

218. We set out briefly in Chapter 2 some of the compelling and disturbing evidence we received concerning children's access to pornography and the impact that it has on them. It is worth recounting here the evidence of Prof McGlynn and Dr Elly Hanson, Clinical Psychologist, that the largest pornographic sites are immediately accessible, have no age verification at all and that their "landing pages" auto-play videos with themes around violence against women, lack of consent (including rape of sleeping women) and sex between step-relatives. Even when bondage, domination, submission, and masochism (BDSM) content was excluded, one in eight videos on the homepages of XVideos, Pornhub and XHamster depicted sexual violence or non-consensual conduct, including unconscious women and girls being raped and footage from "spy cams". All three are among the UK's top 25 most visited websites. [397]

219. Many of the submissions we received called for specific provisions in the Bill for pornographic sites or the revival of the provisions in the Digital Services Act. 5Rights in their evidence called for a broader approach, aligning the Bill with the AADC by including "services likely to be accessed by children" as a third category of regulated service under Clause 3(2). They identified this as their priority change to the draft Bill, seeing it as placing the onus on pornography, app stores and other sites without user-to-user content to either ensure they are not "likely to be accessed by children" or to comply with the child safety duties of the draft Bill.[398] Either way, pornography sites would have a legal obligation to prevent children from accessing their content.

220. We asked Ofcom whether extending the scope of the Bill to include the scope of the AADC would avoid the risk that the Bill might be brought into disrepute by pornographic sites escaping regulation by removing user-to-user content. Dame Melanie agreed that it would achieve the aim and acknowledged the risk, whilst stressing the concerns about scope outlined above.[399] The Secretary of State disagreed with such a change, seeing pornography as a separate issue, and telling us: "We need to keep the scope of the Bill very tight in order to keep it watertight and effective, so that it works. This is not the Bill to fix all online problems and harms. It is important to say that. This Bill is not to fix the internet. This Bill is solely aimed at platforms that we know do harm to children."[400]

395 For example, Written evidence from: Care (OSB0085); CEASE (Centre to End All Sexual Exploitation) (OSB0104); BBC News, 'Pornhub removes all user-uploaded videos amid legality row': https://www.bbc.co.uk/news/technology-55304115 [accessed 30 November 2021]; Written evidence from BBFC (OSB0006)

396 Written evidence from: CEASE (Centre to End All Sexual Exploitation) (OSB0104); BBFC (OSB0006)

397 Written evidence from: Professor Clare McGlynn (Professor of Law at Durham University) (OSB0014); Dr Elly Hanson (Clinical Psychologist & researcher) (OSB0078); CEASE (Centre to End All Sexual Exploitation) (OSB0104)

398 Q 62 (Izzy Wick)

399 Q 259 (Dame Melanie Dawes)

400 Q 284 (Rt Hon Nadine Dorries MP)

221. **Easy, often unwitting or unintended, access by children to pornography was one of the largest online concerns raised with us during our scrutiny of the draft Bill. It is evident to us that the credibility of the Bill will be undermined if the largest online pornography providers simply remove user-to-user elements from their sites and continue showing extreme content and content that creates a risk of harm to children.**

222. **Whilst there is a case for specific provisions in the Bill relating to pornography, we feel there is more to be gained by further aligning the Bill with the Age Appropriate Design Code. Whilst we understand the concerns over scope and regulatory burden, this provision would only bring within the scope of the Bill services already covered by the scope of the Age Appropriate Design Code. Both regulatory systems are risk-based and require the regulator to act proportionately. This step would address the specific concern around pornography, requiring all such sites to demonstrate that they have taken appropriate steps to prevent children from accessing their content. It would also bring other sites or services that create a risk of harm into scope whilst bringing us closer to the goal of aligned online regulation across data protection and online safety. We believe that our proposal on expanding the role of risk profiles, discussed later in this report, will be key to ensure that the Bill's provisions impact the riskiest services and are not disproportionate on those at lower risk.**

223. ***All statutory requirements on user-to-user services, for both adults and children, should also apply to Internet Society Services likely to be accessed by children, as defined by the Age Appropriate Design Code. This would have many advantages. In particular, it would ensure all pornographic websites would have to prevent children from accessing their content. Many such online services present a threat to children both by allowing them access and by hosting illegal videos of extreme content.***

Age Assurance and verification

224. We discuss age assurance above as one of the possible ways that companies can mitigate risks to children resulting from them accessing unsuitable services. This is a fast-growing area with new technological methods being identified including some that use AI, for example, in facial analysis.

225. The impact of some of the information, behaviours and pressures that exist for children online is well-documented.[401] Jim Steyer, CEO of Common Sense Media, told us: "the psychological impacts on young people—on children and teens—are discernibly more important and significant because their brains are still developing." He went on to describe research conducted by Common Sense Media about the negative impacts of some social media services on their "self-esteem, their sense of anxiety and depression, their body image, and their sense that their body images and their overall existence, if you will, does not measure up to the idealised images they can see from influencers on platforms like Instagram."[402] Professor Jonathan Haidt, Ethical Leadership at New York University Stern School of Business, linked this to a "sudden trend" in the US between 2015 and 2021 when the "pre-teen suicide rate in the USA is more than double in those couple of years … Something terrible and huge is happening."[403]

401 See para 16 of this report.
402 QQ 148–153
403 QQ 148–153

226. Children's online experiences are not limited to products and services that are intended for them. Ms Wick said: "The minimum age of use for most social media platforms is 13, but we know that a huge number of under-13s use these platforms. If the companies recognised this, their use base would drop quite significantly."[404] The Office of the Children's Commissioner noted: "Overall, we received little clarity from the companies on how many children they estimated to be using their services, and how many underage users were being identified, although there were notable exceptions to this." [405]

Role of the draft Bill–minimum standards

227. At the moment there is no single regulatory code in the UK that sets out rules for age assurance. There is the Age Appropriate Design Code, which sets out rules for data protection. There is the Video Sharing Platform Regulations, which requires service providers to protect "under-18s from videos containing pornography, extreme content and other material which might impair their physical, mental or moral development."[406] Finally there are the provisions of Part 3 of the Digital Economy Act 2017 that would have required mandatory age assurance for commercial pornography sites, but that have never been brought into force. The draft Bill repeals both these and the Video Sharing Platform Regulations.

228. Ms Wick told us that these measures had pushed forward the development of age assurance. In her view the draft Bill was " … to set the expectations for companies and establish rules of the road. We need Ofcom to produce minimum standards on very basic things such as … age assurance. There should be a requirement in the Bill for Ofcom to do that, which will then establish the floor of protection."[407] Summarising the outcome of our roundtable, Professor Lee Edwards, Strategic Communications and Public Engagement at LSE, noted:

> "There was broad consensus among participants about the risks of platforms being granted too much leeway regarding risk assessments and age verification. Without clear guidelines on the powers of Ofcom, there is a danger of companies setting their own standards. Likewise, a strict separation is necessary between verification providers, their clients, and advertisers." [408]

Privacy

229. On the other hand, numerous witnesses expressed concern about the impact of age assurance on privacy. Big Brother Watch feared that " … mandating age verification … would be hugely damaging to privacy rights online." [409] Demos wrote specifically of children's right to privacy and their concern that:

> "Although there are many third-party identity providers, it is likely that this market would be instantly captured by the large tech companies who already

404 Q 62
405 Written evidence from Office of the Children's Commissioner (OSB0019)
406 Ofcom, 'Video-sharing platform regulation': https://www.ofcom.org.uk/online-safety/information-for-industry/vsp-regulation [accessed 15 November 2021]
407 QQ 52–68
408 Written evidence LSE Department of Media and Communications—Anonymity & Age Verification Roundtable (OSB0236)
409 Written evidence from Big Brother Watch (OSB0136)

> facilitate identity provision across platforms, such as Facebook and Google. This would further consolidate their market power and their control of and ability to use and monetise people's personal data." [410]

230. Internet Matters highlighted concerns of inadvertent exclusion by age assurance processes: "Some vulnerable groups are at risk of being unable to access content they are entitled to, for example, some care experienced young people will not have easy access to an acceptable form of official ID required for age verification. Some young people will also be unable to comply with age assurance mechanisms for physical or cognitive reasons." [411]

231. Others recognised concerns relating to privacy but did not feel they were wholly justified. 5Rights explained: "people are rightly concerned about privacy implications, as are we when it comes to children's privacy, but age assurance is not the same thing as identification, and you can establish a user's age without knowing anything else about them. The technology exists. What is missing is the governance around it." [412]

Independent age assurance sector

232. In the report of our roundtable discussion held on 27 October 2021, Professor Lee Edwards noted:

> "A range of standards and technology options exist for age verification. These include hard verification via identity documents, facial recognition systems, or age estimation. While all technologies come with trade-offs, the significance of finding privacy-preserving solutions was highlighted." [413]

233. We heard that the developing age assurance sector was willing to follow minimum standards set by government.[414] For example, the Age Verification Providers Association recommended: "an independent privacy-protecting standards based, open competitive and interoperable age verification sector as a foundation for a safer internet for children." In their written evidence, Yoti recommended: "standards based and independently accredited approaches to age verification and identity verification for social media registration." [415] Yet some element of compulsion or regulation is likely to be required. The Office of the Children's Commissioner said:

> "Many of the tech companies recognised the need to implement or improve their age assurance systems in order to more effectively enforce their minimum ages. However, some indicated that they did not feel that the issue was a problem for their service, or that the design of the service meant implementing age assurance was difficult, perhaps impossible."[416]

410 Written evidence from Demos (OSB0159)
411 Written evidence from Internet Matters (OSB0103)
412 QQ 52–68
413 Written evidence from LSE Department of Media and Communications—Anonymity & Age Verification Roundtable (OSB0236)
414 Written evidence from: The Age Verification Providers Association (OSB0122); Match Group (OSB0053)
415 Written evidence from Yoti (OSB0130)
416 Written evidence from Office of the Children's Commissioner (OSB0019)

Age Assurance (Minimum Standards) Bill

234. The Age Assurance (Minimum Standards) Bill, a Private Member's Bill, was introduced in the House of Lords on 27 May 2021 and had its second reading on 19 November 2021.[417] If enacted, the Bill would require that age assurance systems for online or digital services or products used by consumers or operated in the UK must meet certain minimum standards. It requires Ofcom to publish those minimum standards within six months of the passing of the Bill. During the Bill's second reading, the Parliamentary Under-Secretary of State at DCMS explained that, while sharing its aims, the Government's view was that this draft Online Safety Bill was the better route by which they could be met, with the regulator including "steps on age assurance in its regulatory codes, as part of which Ofcom can include specific standards and name them."[418]

235. **There is currently no single regulatory or statutory code in the UK that sets out rules for age assurance. We believe that existing codes, and the duties outlined in the draft Bill, cannot be implemented properly without a statutory system of regulation of age assurance, that is trusted, effective and preserves privacy. We believe that an independent, privacy-protecting age assurance sector operating to a set of minimum standards appropriate for different methods of age assurance in different circumstances is key to any system that aims to protect children from harm online. Such a system:**

a) **should be for independent commercial providers as well those built by the service providers themselves;**

b) **should impose standards appropriate to the content and age of the user and be compatible with existing law, including international treaties such as the UN Convention of the Rights of the Child, to provide necessary protections for privacy and data protection; and**

c) **should provide a route of redress for users to challenge specific conclusions reached on age.**

A binding Code of Practice would provide a clear basis for service providers whose risk assessment identifies their content as likely to be accessed by children to put in place mitigations in the form of a rigorous system of age assurance.

236. ***We recommend that the Bill require Ofcom to establish minimum standards for age assurance technology and governance linked to risk profiles to ensure that third-party and provider-designed assurance technologies are privacy-enhancing, rights-protecting, and that in commissioning such services providers are restricted in the data for which they can ask. Ofcom should also require that service providers demonstrate to them how they monitor the effectiveness of these systems to ensure that they meet the minimum standards required.***

237. ***The Government should ask Ofcom to prioritise the development of a mandatory age assurance technology and governance code as a priority ahead of the Bill becoming law and, in doing so, set out risk profiles so that the use of such systems is clearly proportionate to the risk. The code must bear in mind that children have rights to freedom of association, participation, and information, as well as the right to protections. We expect this to be in place within three to six months of the Bill receiving Royal Assent.***

417 Age Assurance (Minimum Standards) Bill [HL] [Bill 19 (2021–22)]
418 HL Deb, 19 November 2021, col 538

6 Scope of the draft Bill

Meaning of "regulated service"

238. The draft Bill covers user-to-user services and search services.[419] It includes messaging apps (such as Facebook Messenger, WhatsApp, Telegram, etc) but excludes "one to one aural communication", SMS messages and email as well as particular types of content, discussed later.[420]

239. The draft Bill divides regulated services into three categories:—Category 1 (likely to include the largest user-to-user services), Category 2A (search services) and Category 2B (user-to-user services that do not meet the threshold to be Category 1).[421] It is ambiguous as to whether some user-to-user and search engines may not meet the threshold for Category 2 and may therefore be outside the regulator regime entirely. The draft Bill requires Ofcom to maintain a register of services in each category, determined by thresholds set by the Secretary of State in accordance with Schedule 4. Certain duties of the draft Bill only apply to Category 1 services—for example, the safety duties for adults and the duties to protect content of democratic importance and journalistic content.[422]

Categorisation

240. As set out above, the draft Bill divides service providers into categories based on size and functionality. The categories in the draft Bill restrict some duties to only the largest and most high-risk services—primarily the safety duty in respect of content that is harmful to adults and the duties to protect journalistic content and content of democratic importance. The Impact Assessment for the draft Bill states:

> " … the current estimate based on the policy intention is that only up to 20 of the largest and highest risk services will meet the Category 1 thresholds, likely to be large social media platforms and potentially some gaming platforms and online adult services."[423]

241. The idea of categorisation was welcomed by some witnesses. Mr Ahmed told us: "There are rational reasons for splitting out the larger platforms from the smaller platforms. … Like The Disinformation Dozen, you should focus as much as possible on where the greatest quantum of harm is caused."[424]

242. Whilst the draft Bill sets out the principle of thresholds, the setting of them is left to secondary legislation and the Government has given little guidance on where they might end up, especially the thresholds for the 2A and 2B categories. Uncertainty over categorisation was a significant theme in the evidence we had from service providers. TechUK noted that the regulation was likely to cover around 24,000 in-scope services,

419 Draft Online Safety Bill, CP 405, May 2021, Clause 2
420 Draft Online Safety Bill, CP 405, May 2021, Clause 39(2)
421 Draft Online Safety Bill, CP 405, May 2021, Clause 59
422 Draft Online Safety Bill, CP 405, May 2021, Clauses 11, 13, 14
423 Department of Digital, Culture, Media and Sport, *The Online Safety Bill: Impact Assessment* (May 2021) RPC-DCMS-4347(2), para 116: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/985283/Draft_Online_Safety_Bill_-_Impact_Assessment_Web_Accessible.pdf [accessed 18 November 2021]
424 Q 8 (Imran Ahmed)

with purposes ranging from educational to professional to social.[425] They called for clarity on the thresholds for Category 1 and 2 services and how different services operating from the same company might be treated. They argued there was a real challenge for companies to prepare for the regulation without this information.[426]

243. The categorisation in the draft Bill has also been criticised as underestimating the impact of small, high-risk companies. Some providers, like Microsoft, argued for a more nuanced approach to categorisation, taking account of risk factors such as the provider taking specific actions to boost the virality of content, a history of illegal content on the service or the ability to use it anonymously.[427] Large services like Facebook argue that much of the activity that creates a risk of harm on their services comes originally from smaller services.[428] Organisations like Hope not Hate, the Antisemitism Policy Trust, and Stonewall stressed the role of "alternative" services such as 4Chan and BitChute in hosting and spreading extremist content or misinformation.[429] The Samaritans told us they are regularly contacted by members of the public, including the bereaved parents of children who have died by suicide, concerned about children as young as 12 accessing smaller services that promote and assist suicide.[430]

244. During our visit to Brussels, we had a useful discussion about the work of the Centre on Regulation in Europe and, particularly, Dr Sally Broughton Micova's (Lecturer in Communications Policy and Politics at the University of East Anglia) work on the relationship of size to risk for online services. That work highlights that the "public" nature of a service and the risk of it aggregating private harm into societal harm does not just depend on size. It also depends on the service's interconnectedness to other services, its impact on media plurality and its impact within a relatively constrained geographic area such as a Member State in an EU context.[431]

245. We put the idea of a more nuanced approach to categorisation to Ofcom. They agreed that the Bill should not lose sight of "smaller but extremely risky" services, though they reminded us that most of the draft Bill (including the illegal content and content harmful to children duties) apply to Category 2B providers.[432]

246. ***We recommend that the categorisation of services in the draft Bill be overhauled. It should adopt a more nuanced approach, based not just on size and high-level functionality, but factors such as risk, reach, user base, safety performance, and business model. The draft Bill already has a mechanism to do this: the risk profiles that Ofcom is required to draw up. We make recommendations in Chapter 8 about how the role of the risk profiles could be enhanced. We recommend that the risk profiles replace the "categories" in the Bill as the main way to determine the statutory requirements that will fall on different online services. This will ensure that small, but high risk, services***

425 Written evidence from techUK (OSB0098)

426 Written evidence from: techUK (OSB0098); others who raised similar concerns included Mumsnet (OSB0031); and Glassdoor (OSB0033)

427 Written evidence from Microsoft (OSB0076)

428 Written evidence from Facebook (OSB0147)

429 For example, written evidence from: HOPE not hate (OSB0048); Q 38 (Danny Stone); Q 38 (Nancy Kelley); Q 57 (Nina Jankowicz)

430 Written evidence from the Samaritans (OSB0182)

431 The Centre for Regulation in Europe (CERRE), *Issue Paper: Sally Broughton, Micova; What is the Harm in Size? Very Large Online Platforms in the Digital Services Act* (October 2021): https://cerre.eu/wp-content/uploads/2021/10/211019_CERRE_IP_What-is-the-harm-in-size_FINAL.pdf [accessed 18 November 2021]

432 Q 258 (Dame Melanie Dawes)

are appropriately regulated; whilst guaranteeing that low risk services, large or small, are not subject to unnecessary regulatory requirements.

Search engines

247. As noted above, the draft Bill includes search engines in a separate category—Category 2A—from user-to-user services. This means they cannot come into Category 1 and cannot be subject to the safety duties relating to adults or the protections about journalistic content or content of democratic importance. In many other respects, however, the duties and responsibilities that the draft Bill places on them are broadly similar to those for user-to-user services.

248. Search providers felt that their duties under the draft Bill were not sufficiently differentiated from the requirements on user-to-user services. Google told us:

> "search services do not host user-generated content. They are an index of trillions of web pages. We provide users with the ability to find relevant information relative to those index pages. The broad definition of online harms would require us potentially to have to make contextual decisions about the nature of content that we do not host, which is on another service, and require us to monitor those billions of web pages in doing so."[433]

249. Groups representing the targets of online hate rejected the idea of search engines as passive "indexes". Mr Stone told us:

> "The search companies are having a laugh at the Bill's expense if they are not included in Category 1. Google was directing people to the search "Are Jews evil?" Microsoft Bing was directing people to "Jews are", and then a rude word about Jews. Currently, if you were to search the word "goyim"—originally a Yiddish word for non-Jews, which is being used in a pejorative sense now—on Microsoft Bing, you will get an antisemitic website and a suspended Twitter account as the top search results. Alexa and Siri are completely outside the bounds of responsibility of this Bill. It would be a travesty if they are left out of Category 1. They should absolutely be in there."[434]

250. For groups like these, the exclusion of search engines and smaller services from the duties in respect of content harmful to adults (in particular) leaves a significant gap in the draft Bill.

251. **We recognise that search engines operate differently from social media and that the systems and processes required to meet the separate duties that the draft Bill places on them are different. The codes of practice drawn up by Ofcom will need to recognise the specific circumstances of search engines to meet Ofcom's duties on proportionality. Search engines are more than passive indexes. They rely on algorithmic ranking and often include automatic design features like autocomplete and voice activated searches that can steer people in the direction of content that puts them or others at risk of harm. Most search engines already have systems and processes in place to address these and comply with other legislation. It is reasonable to expect them to come under**

433 Q 224 (Markham C. Erickson)
434 Q 46 (Danny Stone MBE)

the Bill's requirements and, in particular, for them to conduct risk assessments of their system design to ensure it mitigates rather exacerbates risks of harm. We anticipate that they will have their own risk profiles.

End-to-end encryption

252. End-to-end encryption (E2EE) allows messages to be transferred securely between devices by preventing service providers, and unassociated third parties, from viewing communications' content. A range of messenger services, including the Facebook-owned WhatsApp as well as Telegram and Signal, use encryption to ensure people's privacy in their private messages. However, this can present challenges for ensuring online safety because services are technologically unable to access people's content and cannot apply their usual moderation processes.

253. Currently, most E2EE services identify activity that creates a risk of harm by either using metadata signals (such as the number of messages being sent) or relying on people to submit reports. Both approaches have substantial limitations. For instance, metadata signals may help to identify fraud and inauthentic behaviour but are unlikely to identify hateful content or some forms of content that creates a risk of harm, for example, content related to eating disorders. User reports are unlikely to work for interactions between likeminded individuals, such as terrorists or their sympathisers, child abusers or extremists. The risk presented by E2EE services which lack appropriate safeguards is particularly acute when the communications involve more than just two individuals (as with 1-to-1 services). Some encrypted services allow the creation of groups involving hundreds or thousands of members. We heard concerns from those who believed that the draft Bill would undermine the privacy offered by E2EE.[435] We also heard from other witnesses who drew attention to the use of encrypted services in illegal or harmful activity, including CSEA and fraud.[436]

254. In their evidence submission, the ICO stated that E2EE and online safety should not be seen as "a false dichotomy", emphasising that proportionate responses need to be developed which do not "unduly interfer[e]" with the benefits presented by E2EE.[437] Balancing the benefits of E2EE in terms of protecting individuals' freedom and privacy, with online safety requirements is essential for the Bill to be effective. However, it is unclear how this will be achieved, and mature technical solutions which enable content on E2EE services to be moderated are not widely available. If the challenges presented by E2EE are not resolved then, in the extreme, there are two potential negative outcomes: (1) E2EE becomes infeasible because services cannot meet online safety requirements; or (2) E2EE is overapplied because it lets services avoid their regulatory obligations.

255. We heard during our visit to Brussels that there are ways that risks on E2EE services can be mitigated without breaking encryption or compromising the privacy of genuinely private communications. These included design features being built into the services themselves such as limiting the frictionless mass sharing of material, age assurance designed to prevent abusers from luring children into conversations on encrypted services,

435 Written evidence from: Tech Against Terrorism (OSB0052); British & Irish Law, Education & Technology Association (OSB0073); Sophie Zhang (Former Facebook Employee) (OSB0214)

436 Written evidence from: UK Finance (OSB0088); NSPCC (OSB0109); Mrs Gina Miller (OSB0112)

437 Written evidence from Information Commissioner's Office (OSB0211)

and media literacy campaigns on the message "report don't share" in respect of illegal content.

256. Concerns have been raised about the use of privacy-protecting and encrypted services to access the Internet, including proxies and VPNs. We heard from Mobile UK, the trade association for UK's mobile network operators, about Apple Private Relay.[438] Introduced in "beta" mode in 2021, Private Relay uses separate internet relays to process who is using the Internet and what they are searching for. Apple claims that this means it is technologically unable to break privacy protection, similar to a VPN. In contrast to a VPN, Private Relay is intended to be easier to use, cheaper and add less of a time delay. It only works for Apple's Safari browser but because it is included with an iCloud+ subscription, Private Relay could become widely used. The use of privacy-protecting services to access the Internet could reduce the number of points at which users are protected from unsafe content. We heard from BT Group that they "render ineffective the software that ISPs and mobile companies currently use to block images of child abuse and other extreme or illegal content."[439]

257. **The Government needs to provide more clarity on how providers with encrypted services should comply with the safety duties ahead of the Bill being introduced into Parliament.**

258. ***We recommend that end-to-end encryption should be identified as a specific risk factor in risk profiles and risk assessments. Providers should be required to identify and address risks arising from the encrypted nature of their services under the Safety by Design requirements.***

Exclusion of paid-for advertising from scope

259. Paid-for adverts are not included in the scope of the draft Bill. We heard that this exclusion creates a gateway for various harms to be spread online. The FCA told us that: "the problem [of online fraud] is most manifest in the paid-for space, so it does not make sense for the Bill not to deal with the very heart of the problem, which is the paid-for advertising space."[440] Similarly, Which? noted: "Paid-for advertising on online platforms is a primary method used by criminals to target consumers and engage them in a [financial] scam, as it gives them instant access to large numbers of target audiences."[441]

260. Several witnesses suggested that if paid-for advertising remained excluded from scope, criminals might switch to paying for fraudulent content to be disseminated.[442] Which? also told us they had investigated how easy it was to advertise a scam by creating a fake drinking water and hydration service. They were able to have it advertised on Facebook and Google with minimum checks, as well as pay for it to appear above the NHS website in Google searches about hydration.[443] The Advertising Standards Authority confirmed

438 Mobile UK, *Digital, Culture, Media and Sport Sub-Committee: Call for Evidence on Online Safety and Online Harms* (September 2021): https://uploads-ssl.webflow.com/5b7ab54b285deca6a63ee27b/61680f8d16cc3e792c4c2c1e_MobileUK_Online_safety_draft_bill_020921.pdf [accessed 9 December 2021]

439 Written evidence from BT Group (OSB0163)

440 Q 121

441 Written evidence from Which? (OSB0115)

442 This concern is raised in written evidence from: Reset (OSB0138) and Dame Margaret Hodge (Member of Parliament for Barking and Dagenham at House of Commons) (OSB0201), and oral evidence by the FCA (Q 120), Which? (Q 112 (Rocio Concha)), Martin Lewis (Q 112 (Martin Lewis)), and Ofcom (Q 263).

443 Q 111

that it knew "from recent research that increasing concerns about scams are influencing the public's trust in online ads."[444]

261. The exclusion of paid-for adverts from the scope of the draft Bill leaves little incentive for operators to remove scam adverts. Regardless of their legitimacy, they generate revenue for platforms—and as we heard from Dame Margaret Hodge MP: "This will continue to benefit the fraudsters and the chief executives." Dame Margaret added that some platforms currently benefit from the FCA paying them to place legitimate adverts above scam ones, making it "a 'win-win' for them."[445] For example, the FCA invested £600,000 in preventing scam adverts on Google in 2020 through public information advertising.[446]

262. We also received evidence of the risk of non-financial harms if paid-for advertisements were excluded from the scope of the Bill. The Center for Countering Digital Hate told us that: "Facebook routinely accepted money to advertise anti-vaccine messages to its users until announcing it would end the practice in October 2020. Even then, our research showed that Facebook continued to broadcast anti-vaccine adverts worth at least $10,000 to its users."[447] Who Targets Me called for paid-for advertising to be brought into scope so that political advertising was subject to regulation. They hope that this would mitigate problems such as foreign powers interfering in UK elections via advertising and the spread of harmful political disinformation.[448]

263. Ms Edelson told us that: "ad tech [advertising technology] can be really powerful in helping scammers to identify … vulnerable populations."[449] For example, she told us: "We saw anti-vax [anti-vaccine] content being promoted to pregnant women in the United States … as a means to sell their vaccine harm reduction supplements."[450] Global Action Plan highlighted to us that studies have "found that it was possible, using Facebook's admanager, to target children on Facebook aged between 13 and 17 based on such interests as alcohol, smoking and vaping, gambling, extreme weight loss, fast foods and online dating services."[451]

264. Most of the evidence we received called for paid-for advertising to be brought into scope of the Bill.[452] We were also told that the public was supportive of such a move in relation to financial scams. Research by Aviva found that 87 per cent of people felt that the

444 Q 118

445 Written evidence from Dame Margaret Hodge (Member of Parliament for Barking and Dagenham at House of Commons) (OSB0201)

446 Written evidence from Dame Margaret Hodge (Member of Parliament for Barking and Dagenham at House of Commons) (OSB0201)

447 Written evidence from Center for Countering Digital Hate (OSB0009)

448 Written evidence from Who Targets Me (OSB0086)

449 Q 108

450 Q 108

451 Written evidence from Global Action Plan (OSB0027)

452 See written evidence from: Keoghs LLP (OSB0003); Somerset Bridge Group Ltd (OSB0004); Work and Pensions Committee (OSB0020); Quilter (OSB0024); Money and Mental Health Policy Institute (OSB0036); Aviva Plc (OSB0042); Financial Conduct Authority (OSB0044); CIFAS (OSB0051); Mr Mark Taber (Consumer Finance Expert, Campaigner & Media Contributor at Mark Taber) (OSB0077); Association of British Insurers (OSB0079); Direct Line Group (OSB0082); UK Finance (OSB0088); Barclays Bank (OSB0106); MoneySavingExpert (OSB0113); Which? (OSB0115); Innovate Finance (OSB0116); Revolut (OSB0117); Lloyds Banking Group plc (OSB0135); Office of the City Remembrancer, City of London Corporation (OSB0148); The Investment Association (OSB0162); BT Group (OSB0163); Paul Davis (Director of Fraud at TSB Bank Plc) (OSB0164); Sky, BT, Channel 4, COBA, ITV, NBC Universal, TalkTalk, Virgin Media O2, Warner Media (OSB0177); Rt Hon. Mel Stride MP (Chair at House of Commons Treasury Select Committee) (OSB0209); Dame Margaret Hodge (Member of Parliament for Barking and Dagenham at House of Commons) (OSB0201); Hargreaves Lansdown (OSB0197)

Government should introduce legislation to ensure that search engines and social media platforms do not promote financial scams through advertising.[453]

265. Google disagreed, telling us that:

> "Advertising and financial fraud involve a complex and highly specialised set of issues and existing rules, and requires consideration of the implications of regulation for legitimate competition and innovation in the UK's dynamic fintech sector. The Advertising Standards Authority and the Financial Conduct Authority are best placed to address these issues."[454]

They also wrote that such an extension in scope: "would significantly add to the 24,000 companies the Government estimates will be affected by the Bill."[455] At the same time, Google are a good example of what can be achieved when a platform decides to co-operate with a regulator. By changing their terms and conditions to only allow adverts for financial services from FCA regulated firms they reduced the number of scam adverts on their platform considerably. Other service providers have not yet implemented this measure.[456]

266. We heard from Ofcom that if paid-for advertisements are brought into scope, "retaining the focus on systems and processes rather than individual fraudulent or other content" should remain their priority.[457] This is consistent with the approach that Ofcom has already taken to the regulation of advertising on video sharing platforms, where they have a duty to ensure that standards around advertising on those platforms are met. Here Ofcom has designated the ASA as co-regulator with day-to-day responsibility for advertising with Ofcom acting as the statutory backstop regulator.[458] They also underlined that "the onus and strategy [should] be clearly owned by the criminal enforcement agencies" when it comes to dealing with individual cases of fraud.[459]

267. Rt Hon Nadine Dorries MP, the Secretary of State for DCMS, told us that paid-for advertising and scams were being considered as part of DCMS's online advertising programme rather than being included in the Online Safety Bill. She also told us that her legal advice is that extending the scope to include paid-for adverts: "would not work and it would extend the scope of the Bill in a way that would not be appropriate."[460] Nonetheless, she did invite us to examine "highly targeted" amendments.[461] Mr Philp wrote to us that bringing paid-for advertising into scope would difficult. Firstly, because "doing so would require a reconsideration of the services in scope of the Bill", since online advertising can involve companies different to the ones the Government has aimed to regulate in the draft Bill.[462] Secondly, he told us that:

> "Safety duties on user-to-user and search services may not be appropriate, and in some cases not feasible, for application to advertisers given the very

453 Written evidence from Aviva (OSB0042)
454 Written evidence from Google (OSB0175)
455 Written evidence from Google (OSB0175)
456 'Instagram favourite site for scammers', *The Times* (26 November 2021): https://www.thetimes.co.uk/article/instagram-favourite-site-for-scammers-8qdmq0ffh [accessed 9 December 2021]; Q 223
457 Q 263
458 Ofcom, *The regulation of advertising on video-sharing platforms* (7 December 2021): https://www.ofcom.org.uk/__data/assets/pdf_file/0022/229009/vsp-advertising-statement.pdf [accessed 9 December 2021]
459 Q 263
460 Q 290 (Nadine Dorries)
461 Q 292 (Nadine Dorries)
462 Written evidence from Department of Digital, Culture, Media & Sport (OSB0243)

> different way in which they operate. Advertising actors rely on very different contractual agreements to publish and disseminate content, in comparison to the user-to-user and search services in scope of the Online Safety Bill. A whole new set of duties would be required to comprehensively address the range of actors involved in the advertising market."[463]

268. **The exclusion of paid-for advertising from the scope of the Online Safety Bill would obstruct the Government's stated aim of tackling online fraud and activity that creates a risk of harm more generally. Excluding paid-for advertising will leave service providers with little incentive to remove harmful adverts, and risks encouraging further proliferation of such content.**

269. ***We therefore recommend that clause 39(2) is amended to remove "(d) paid-for advertisements" to bring such adverts into scope. Clause 39(7) and clause 134(5) would therefore also have to be removed.***

270. **Ofcom should be responsible for acting against service providers who consistently allow paid-for advertisements that create a risk of harm to be placed on their platform. However, we agree that regulating advertisers themselves (except insofar as they come under other provisions of the Bill), individual cases of advertising that are illegal, and pursuing the criminals behind illegal adverts should remain matters for the existing regulatory bodies and the police.**

271. ***We recommend that the Bill make clear Ofcom's role will be to enforce the safety duties on providers covered by the online safety regulation, not regulate the day-to-day content of adverts or the actions of advertisers. That is the role of the Advertising Standards Authority. The Bill should set out this division of regulatory responsibility.***

Economic harms

272. Clause 41(6) excludes online consumer harms from the scope of the draft Bill, such as the sale of goods that are of an unsafe standard or the services of someone not qualified to perform the service. Clause 39(2)(d) sees reviews of products appearing on retailers' websites placed out of scope. We heard opposition to this from Sky and other media and creative businesses, who argued that: "the exclusion of these harms on the face of the legislation stands in contrast to the Government's stated ambition of a 'coherent, single regulatory framework' for online platforms."[464] These businesses were also concerned about the exclusion of intellectual property infringements from the draft Bill.[465]

273. In contrast, the British Retail Consortium supported the exclusion, claiming that its removal "could have an adverse impact on retail, not least the smaller retailers"[466] because of the regulatory burden it would place on them. They believe that "as it stands customer reviews on products that are sold by a third party to a customer via a marketplace are in scope",[467] unlike reviews on products retailers sell themselves. They urged the Government

463 Written evidence from Department of Digital, Culture, Media & Sport (OSB0243)
464 Written evidence from Sky (OSB0165)
465 For example, written evidence from Sky, BT, Channel 4, COBA, ITV, NBC Universal, TalkTalk, Virgin Media O2, Warner Media (OSB0177); Alliance for Intellectual Property (OSB0016)
466 Written evidence from British Retail Consortium (BRC) (OSB0087)
467 *Ibid.*

to exclude all reviews from scope to avoid bringing more retail companies into scope, and because reviews help consumers make informed choices.

274. The CMA told us they "consider that existing consumer law requires platform operators to take reasonable and proportionate steps to effectively protect consumers from economically harmful illegal content."[468] As such, they have already acted to, for example, "tackle the trading of fake and misleading online reviews on Facebook and eBay".[469] They therefore argue that online consumer harms should remain excluded from scope and that the Government should "use an alternative or existing legislative initiative to ensure the necessary protections for consumers".[470] The CMA told us that bringing consumer harms into scope could mean:

> "Many platform operators are likely to remain unclear about the full extent of their legal responsibilities in connection with economically harmful content … those operators may fail to take steps to implement appropriate systems and processes to effectively tackle such content until regulatory action is taken."[471]

275. **We recognise that economic harms other than fraud, such as those impacting consumers, and infringement of intellectual property rights, are an online problem that must be tackled. However, the Online Safety Bill is not the best piece of legislation to achieve this. Economic harms should be addressed in the upcoming Digital Competition Bill. We urge the Government to ensure this legislation is brought forward as soon as possible.**

468 Written evidence from Competition and Markets Authority (OSB0160)
469 *Ibid.*
470 *Ibid.*
471 *Ibid.*

7 Freedom of speech requirements, journalism, and content of democratic importance

Freedom of expression: Clause 12

276. Balancing people's right to freedom of expression with online safety was one of the most controversial subjects in our inquiry. The draft Bill attempts to tackle this in part by the inclusion of Clause 12, a duty on service providers to "have regard to the importance of protecting users' right to freedom of expression within the law, and protecting users from unwarranted infringements of privacy, when deciding on, and implementing, safety policies and procedures."[472] This duty applies to all service providers, with additional responsibilities placed on Category 1 providers, who must carry out impact assessments and specify how they fulfil this duty in their terms of service.[473] Clause 23 places a similar duty on search services.[474]

Freedom of expression and online safety

277. Article 10 of the ECHR, incorporated into UK law by the Human Rights Act 1998, states:

> "1. Everyone has the right to freedom of expression. This right shall include freedom to hold opinions and to receive and impart information and ideas without interference by public authority and regardless of frontiers. This Article shall not prevent States from requiring the licensing of broadcasting, television or cinema enterprises."[475]

Article 10 is not an absolute right but may be restricted by the state "in the interests of national security, territorial integrity or public safety, for the prevention of disorder or crime, for the protection of health or morals, for the protection of the reputation or rights of others, for preventing the disclosure of information received in confidence, or for maintaining the authority and impartiality of the judiciary."[476] In Handyside v United Kingdom, the European Court of Human Rights confirmed that Article 10 includes the right to say things that "offend, shock or disturb the State or any sector of the population".[477]

472 Draft Online Safety Bill, CP 405, May 2021, Clause 12(2)
473 Draft Online Safety Bill, CP 405, May 2021, Clause 12(3), (5)
474 Draft Online Safety Bill, CP 405, May 2021, Clause 23
475 Council of Europe: European Court of Human Rights, *European Convention on Human Rights* (August 2021) p 12: https://www.echr.coe.int/documents/convention_eng.pdf [accessed 22 November 2021]
476 Council of Europe: European Court of Human Rights, *European Convention on Human Rights* (August 2021) p 12: https://www.echr.coe.int/documents/convention_eng.pdf [accessed 22 November 2021]
477 European Court of Human Rights, *Handyside v United Kingdom* (December 1976): https://www.bailii.org/eu/cases/ECHR/1976/5.html [accessed 7 December 2021]

278. Any restriction must be "prescribed by law",[478] "necessary in a democratic society",[479] and proportionate. We have heard concerns that the draft Bill does not fulfil these criteria, particularly in its definition of harm to adults. Mr d'Ancona told us:

> "I think that with words like "harm" and "safety" there is a slippage or a kind of semantic mission creep going on in their use. We used to talk about safety, and what we really meant was physical safety. Now, when people talk about safety, they often mean convenience or comfort. It is not the task of democratic legislators to make people feel comfortable. I think that is stretching the job description."[480]

279. Since the publication of the White Paper several rights organisations have talked of the "chilling effect" the legislation will have on freedom of expression online.[481] This encompasses both censorship by moderation and removal of content, as discussed in Chapter 2, and the self-censorship that may follow when people are unsure what they can and cannot say online. We have also heard that, due to the volume of content online, the only way service providers can approach their duties is by using AI, and that this risked "overzealous removal of legitimate speech and the limiting of freedom of expression."[482] The size of the penalties, and the potential for criminal liability for managers, have also been highlighted as potential causes of excessive censorship. Open Rights Group said:

> "It will create a culture of fear which results in a phenomenon known as "collateral censorship", where service providers and companies feel they have no choice but to take down vast swathes of content which may be perfectly legal, perfectly subjective, and perfectly harmless, less they face sanctions, penalties, and even personal arrests for getting it wrong."[483]

280. At the same time, we heard that inaction and a lack of regulation on online safety is impacting the freedom of expression of people now, particularly marginalised groups. Ms Jankowicz told us:

> "The idea [of online abuse] is to quash women's right to freedom of expression. It is to take them out of the public eye, because women, according to the purveyors of this abuse, do not deserve to be there. We need to stand up to that. We need to protect that right to work, that right to freedom of expression. That is at the core of this misogynistic abuse."[484]

Hope Not Hate said:

> "If done properly, the inclusion of legal but harmful content within the scope of this legislation could dramatically increase the ability for a wider range of

478 Foreseeability is inherent to fulfilling the 'prescribed by law' requirement in Article 10 cases: *Sunday Times v United Kingdom* (1979) 2 EHRR 245

479 "Necessary" has been strongly interpreted: it is not synonymous with "indispensable", neither has it the flexibility of such expressions as "admissible", "ordinary", "useful", "reasonable" or "desirable": *Handyside v United Kingdom* (1976) 1 EHRR 737, 754, para 48, Shayler v (2002) UKHL 11, (2003) AC 247, per Lord Bingham, para 23

480 Q 135

481 For example: Big Brother Watch, Big Brother Watch's *Response to the Online Harms White Paper Consultation* (July 2019): https://bigbrotherwatch.org.uk/wp-content/uploads/2020/02/Big-Brother-Watch-consultation-response-on-The-Online-Harms-White-Paper-July-2019.pdf [accessed 22 November 2021]

482 Q 135

483 Written evidence from Open Rights Group (OSB0118)

484 Q 55

> people to exercise their free speech online by increasing the plurality of voices on platforms, especially from minority and persecuted communities."[485]

281. Prof Haidt told us that "freedom of speech is not freedom of reach", and that instead of looking at taking down individual pieces of content, service providers should focus on reducing amplification.[486] Ms Carlo said she remained uneasy about this, as it would still involve suppression of legal content.[487] Ms Zhang also urged caution, as it could set an "unfortunate precedent" that could be copied by authoritarian countries and used to shut down protest.[488] However, Maria Ressa told us "we definitely need legislation", adding that "doing nothing pushes the world closer to fascism."[489]

Clause 12

282. While the intention behind Clause 12 (and Clause 23 for search services) has been welcomed, many have noted that the phrasing seems weak by comparison to the other duties placed on service providers, particularly that they must "have regard to" the importance of protecting people's rights. Glassdoor said: "There is a strong risk that process-based safety duties will win out when companies are faced with the task of determining where the greater regulatory risk lies."[490] The Adam Smith Institute also felt that the duty in Clause 12 is "overwritten" by the safety duties, which they describe as "extraordinarily broad and threatening".[491] Mr Millar noted that the draft Bill put competing duties on service providers without providing guidance on how they could be balanced.[492]

283. We asked why the Government had chosen this phrasing, over something stronger such as "ensuring actions are consistent with". DCMS Minister Chris Philp MP told us:

> "'consistent with', and other similar formulations, might suggest that service providers owe a duty to their users under the ECHR or Human Rights Act 1998. The ECHR only imposes obligations in relation to freedom of expression on public bodies, and private actors are not required to uphold freedom of expression."[493]

The ECHR does apply to the UK Government, and to Ofcom, and so their directions to service providers must still comply with Article 10.

284. **We propose a series of recommendations throughout this report to strengthen protection for freedom of expression. These include greater independence for Ofcom, routes for individual redress beyond service providers, tighter definitions around content that creates a risk of harm, a greater emphasis on safety by design, a broader requirement to be consistent in the applications of terms of service, stronger minimum standards and mandatory codes of practice set by Ofcom (who are required to be compliant with human rights law), and stronger protections for news publisher**

485 Written evidence from HOPE not hate (OSB0048)
486 Q 149
487 Q 149
488 Q 134
489 Q 193
490 Written evidence from Glassdoor (OSB0033)
491 Written evidence from Adam Smith Institute (OSB0129)
492 Written evidence from Gavin Millar QC (OSB0221)
493 Written evidence from Department of Digital, Culture, Media and Sport (OSB0248)

content. We believe these will be more effective than adjustments to the wording of Clause 12.

Journalism and content of democratic importance

285. Journalism and political debate are fundamental aspects of freedom of expression in a democratic society and receive a higher level of protection than everyday speech under the ECHR.[494] Following concerns raised about the impact of regulation on freedom of expression and particularly media freedom in the White Paper consultation, the draft Bill places specific duties on Category 1 providers aimed at preventing excessive moderation of journalism or content of democratic importance. While these exemptions have been broadly welcomed, we have heard concerns about their definitions and the ability this will give the service providers to apply them consistently.

News publisher content

286. The draft Bill attempts to protect journalism and media freedom by two means. The first is by placing "news publisher content" outside of scope entirely, including on publishers' own websites, in search results and if it is shared on user-to-user services, providing it is shared in full and without editing.[495] "Recognised news publishers" are defined by Clause 40, which includes among others the requirements for them to produce "news-related material" with editorial control, a complaints procedure, a registered business address in the UK, and to be subject to a standards code.[496] Comments on a news publisher's site are also exempt, by means of the "limited functionality" exemption.[497]

287. Witnesses such as Hacked Off and the Independent Media Association disagreed with the definition used for "recognised news publishers", feeling it could leave content that creates a risk of harm outside of scope while failing to protect independent publishers, who may not have a registered office.[498] Hacked Off gave examples of websites that may qualify for the exemption, despite publishing racist stories and conspiracy theories.[499] IMPRESS suggested that the inclusion of requirements for a standards code and complaints procedure, which make the draft Bill's definition differ from that of a "relevant publisher" used in the Crime and Courts Act 2013, offer no extra protection, as they can be set by the publishers themselves with no minimum criteria.[500] The Professional Publishers Association called for the "news-related material" requirement to be changed in favour of including consumer magazines and business media, which may currently fall outside of the definition if not focussing on current affairs.[501]

288. A concern we heard from news organisations was that it is "news publisher content" that is exempt under Clause 39, rather than news publishers' websites in their entirety being explicitly placed outside of scope. While the limited functionality exemption covers below-the-line comments, News Media Association told us they were concerned because this can be repealed by the Secretary of State, and because the inclusion of other features

494 On 'political speech' see *R v BBC, ex p ProLife Alliance* (2003) UKHL 23
495 Draft Online Safety Bill, CP 405, May 2021, Clause 39(8)
496 Draft Online Safety Bill, CP 405, May 2021, Clause 40(2)
497 Draft Online Safety Bill, CP 405, May 2021, Schedule 1, part 5
498 Written evidence from: Hacked Off (OSB0041); The Independent Media Association (OSB0064)
499 Written evidence from Hacked Off, Annex A (OSB0041)
500 Written evidence from IMPRESS (OSB0092)
501 Written evidence from Professional Publishers Association (PPA) (OSB0154)

such as games or online workshops may bring a news publisher's site into scope of the regulations.[502] The National Union of Journalists took a different view on the exclusion of comments sections:

> "Material that doesn't pass the editorial or legal threshold for other published material—as abuse, threats and defamatory content clearly does not—should not be publishable on the sites of media outlets in 'below the line' commentary dressed up as reader engagement."[503]

We heard in evidence that newspapers can and have been held liable for comments on their own sites, that they already risk assess and apply moderation techniques, and that the Independent Press Standards Organisation has regulatory oversight.[504]

Competition and media plurality

289. Ofcom noted in its recent report on the future of media plurality that "online intermediaries and their algorithms control the prominence they give to different news sources and stories" and the basis on which they "serve news via their algorithms is not sufficiently transparent."[505] These were identified as risks to media plurality in the UK that are not captured under the existing regulatory framework. We have heard throughout this inquiry about the power the biggest service providers hold in controlling the news and information people see, and DMG Media detail in their written evidence the impact search engine control has on the visibility of different news sites.[506] We heard powerful evidence from Ms Ressa, Ms Zhang and Ms Haugen about the influence social media companies have both in the UK and abroad, particularly the global South, as they monopolise the news.[507] On the other hand, Dr Martin Moore, Senior Lecturer in Political Communication Education and Director of the Centre for the Study of Media, Communication and Power at King's College London, discussed how identifying recognised news publishers may have unwanted consequences by creating a list of statutory-recognised news publishers. [508]

290. Work is ongoing in this area with the creation of the Digital Markets Unit (DMU) within the Competition and Markets Authority (CMA) to promote competition[509], with powers to be provided by a Digital Competition Bill. DMG Media told us that, ahead of the establishment of the DMU's regulatory powers, it was important that the online safety legislation included a full exemption for news publishers, to prevent services such as Google using it to discriminate against "those it does not favour".[510]

291. ***We recommend that Ofcom be required to produce an annual report on the impact of regulated services on media plurality.***

502 Written evidence from News Media Association (OSB0107)
503 Written evidence from The National Union of Journalists (NUJ) (OSB0166)
504 Q 147
505 Ofcom, *The future of media plurality in the UK* (November 2021) p 1: https://www.ofcom.org.uk/__data/assets/pdf_file/0019/228124/statement-future-of-media-plurality.pdf [accessed 30 November 2021]
506 Written evidence from DMG Media (OSB0133)
507 Q 128; Q 144; Q 196
508 Written evidence from Dr Martin Moore (Senior Lecturer at King's College London) (OSB0063)
509 Competition and Markets Authority, 'Digital Markets Unit': https://www.gov.uk/government/collections/digital-markets-unit [accessed 23 November 2021]
510 Written evidence from DMG Media (OSB0133)

Journalistic content

292. In addition to the exemption for recognised news publishers, Clause 14 places a duty on Category 1 providers to protect journalistic content by using "systems and processes designed to ensure that the importance of the free expression of journalistic content is taken into account" when making moderation decisions.[511] This will apply beyond recognised news publishers and cover citizen journalism. Service providers must also establish a dedicated and expedited complaints procedure so that journalists may appeal when content is moderated or removed, and content must be swiftly reinstated if complaints are upheld.[512] Journalistic content is defined as either news publisher content, or regulated content that is "generated for the purposes of journalism", and UK linked.[513] During our inquiry, we have heard concerns both about the definitions used and the duty itself.

293. While there are evidently some concerns about the "news publisher" definition, we heard that "journalistic content" was much harder to apply, with providers having to determine whether content had been generated for the "purposes of journalism". The House of Lords Communications and Digital Committee thought that the inclusion of citizen journalism could overwhelm the appeals system and recommended that citizen journalism should be clearly defined. The Government said in its response that further clarification was not necessary and that it was intended to be interpreted broadly to include "content produced by individuals, freelancers and others."[514] Facebook told us the broad definition could be abused by those claiming to be citizen journalists "to ensure their content is given protections."[515] Mr Millar offered a definition: "content by which the user who generates it disseminates information and ideas to the public (or a section of the public) which they reasonably perceive to be of public interest."[516]

294. The duty to protect journalistic content does not mean such content cannot be removed in any circumstance. Clause 14 states that Category 1 services must ensure "the importance of the free expression of journalistic content is taken into account" when making moderation decisions and provide an expedited complaints procedure with swift reinstatement of content when appeals are upheld. It does not mandate that all such appeals be upheld simply because the content is journalistic—providers may still remove content they judge to create a risk of harm that breaks their terms of service. In the draft Bill they will be required to make their policies clear and enforce them consistently, and it will be for Ofcom to determine if they are balancing the application of their duties correctly.

295. We have heard concern from news publishers that, while their content is placed outside of scope, and Clause 14 seemingly offers further protection, there is not a strong enough disincentive to stop providers from removing it at all. Peter Wright, Editor Emeritus at DMG Media, told us that he feared that the use of algorithms may continue to mean news content is caught by automatic moderation as they are a "blunt instrument",

511 Draft Online Safety Bill, CP 405, May 2021, Clause 14(2)

512 Draft Online Safety Bill, CP 405, May 2021, Clause 14(3),(4),(5)

513 Draft Online Safety Bill, CP 405, May 2021, Clause 14(8). 'UK-linked' means the UK is a target market, or the content is or is likely to be of interest to a significant number of United Kingdom users.

514 Department of Digital, Culture , Media and Sport, 'Government response to the House of Lords Communications Committee's report on Freedom of Expression in the Digital Age': https://committees.parliament.uk/publications/7704/documents/80449/default/ [accessed 30 November 2021]

515 Written evidence from Facebook (OSB0147)

516 Written evidence from Gavin Millar QC (OSB0221)

and Alison Gow, President of the Society of Editors, added that by the time the issue is caught by human moderators, it is often too late.[517] Ms Haugen gave us the example that "76 per cent of counterterrorism speech in an at-risk country was getting flagged as terrorism and taken down" and that "any system where the solution is AI is a system that is going to fail."[518] Mr Perrin told us that since traditional print media is already self-regulated, "introducing a new layer when one is trying to regulate in a complex new sector was always going to be counterproductive."[519] Ms Ressa warned that algorithmic design and the "incentive scheme" of the internet was already pushing people away from quality journalism "towards clickbait", and told us "the news agenda needs to be protected."[520] We heard that the "perishable" nature of news meant that even an expedited complaints system may be too slow, and instead providers should be required not to restrict access to news publisher content.[521] We asked the Government for their views on this:

> "'Positive requirements', which actively prevent social media companies from removing any news publisher content, regardless of whether they consider it to comply with their terms of service or to be suitable for their audience, carry significant risk. This approach would constitute a significant interference with private companies' ability to set their own terms and conditions regarding legal content. Moreover, it could create perverse outcomes if companies were prevented from removing this type of content in all circumstances."[522]

Content of democratic importance

296. Alongside the protections for journalism, the draft Bill contains a duty to protect "content of democratic importance". Clause 13 is similar to Clause 14's protections for journalism, in that providers must ensure their systems and processes are "designed to ensure that the importance of free expression of content of democratic importance is taken into account" when making moderation decisions or taking action against users.[523] It requires clear and transparent policies to be published and providers must ensure they are applied equally across a "diversity of political opinion", but unlike the protection for journalists there is no dedicated complaints route to appeal decisions, so people would need to use the standard complaints process. The definition encompasses news publisher content, and regulated content that "is or appears to be specifically intended to contribute to democratic political debate in the United Kingdom or a part or area of the United Kingdom."[524]

297. Legal to Say, Legal to Type said the exemption, alongside that for journalism, creates a two-tier system, with "free speech for journalists and politicians, and censorship for ordinary citizens."[525] They queried whether individuals could claim protection for a broad spectrum of content that creates a risk of harm if they have stood for office, a concern also reflected by Mr Stone, who gave the example of misogynistic abuse of a Women's Equality

517 Q 143
518 Q 173
519 Q 76
520 Q 194
521 Written evidence from News Media Association (OSB0107)
522 Written evidence from Department of Digital, Culture, Media & Sport (OSB0243)
523 Draft Online Safety Bill, CP 405, May 2021, Clause 13(2)
524 Draft Online Safety Bill, CP 405, May 2021, Clause 13(5)
525 Written evidence from Legal to Say, Legal to Type (OSB0049)

Party candidate by another candidate.[526] It should be noted that as with the journalism exemption, Clause 13 does not mean content cannot be removed at all when judged to be create a risk of harm, and illegal content would still be required to be removed. It may make service providers more cautious about removal, as it is intended to, but some worry that the broad definitions of this and the journalism exemption create loopholes that may undermine confidence in the legislation.[527] Facebook said that "private companies should not be the arbiters of what constitutes journalism or what is democratically important."[528]

298. The Explanatory Notes give as examples "content promoting or opposing government policy and content promoting or opposing a political party."[529] The House of Lords Communications and Digital Committee felt that:

> "The definition of 'content of democratic importance' in the draft Bill is too narrow. It should be expanded to ensure that contributions to all political debates—not only those debates which are about, or initiated by, politicians and political parties, and about policy, rather than social change—would be covered."[530]

In its response, the Government said that the definition "covers all political debates, including where these are advanced by grassroots campaigns and smaller parties" and that the measures were designed "to protect content of democratic importance, rather than to protect specific actors."[531] It was also noted in our session with ministers that currently there are no protections in place for such speech and journalistic content, with Mr Philp stating on the subject of the definitions that "while I am sure they are not perfect and we can try to improve them, currently there is nothing at all."[532]

Public interest

299. In their response to the House of Lords Communications and Digital Committee, the Government said that when deciding how to balance the safety duties with protecting democratic content "platforms will need to consider whether the public interest in seeing some types of content outweighs the risk of harm it creates, or vice versa."[533] We heard providers sometimes censor what can be discussed elsewhere, including content that could be "of great public importance". Ms Carlo gave the example of discussions around the origins of COVID-19 being removed from social media, which would have been allowed in mainstream newspapers.[534]

300. The Government has said these protections are designed to protect content rather than specific actors. However, a great deal of evidence we have heard centred on people's

526 Q 41
527 Q 57
528 Written evidence from Facebook (OSB0147)
529 Explanatory Notes to the draft Online Safety Bill [Bill CP 405-EN], para 95
530 Communications and Digital Committee, *Free for All? Freedom of Expression in the Digital Age* (1st Report, Session 2021–22, HL Paper 54), para 80
531 Department for Digital, Culture, Media and Sport, *Government response to the House of Lords Communications Committee's report Freedom of Expression in the Digital Age* (October 2021), para 7–8: https://committees.parliament.uk/publications/7704/documents/80449/default/ [accessed 9 December 2021]
532 Q 281
533 Department for Digital, Culture, Media and Sport, *Government response to the House of Lords Communications Committee's report Freedom of Expression in the Digital Age* (October 2021), para 9: https://committees.parliament.uk/publications/7704/documents/80449/default/ [accessed 9 December 2021]
534 Q 138

understanding of what a journalist may be, and who is entitled to speak about matters of democratic importance. On this subject, responding to a question on whether someone's standing as a political candidate would give their speech unwarranted protection, Mr Millar suggested that the person's identity was of less importance than whether what they were saying was in the public interest.[535] He elaborated in writing:

> "In the domestic and international law of free speech it is well established that speech on matters of public interest in a democratic society is deserving of the strongest protection. Political speech is the paradigm example of this. Journalism on such matters is also particularly strongly protected.
>
> But speech on matters of public interest in a democratic society is a flexible category of speech. It is not closed. It covers more than just political speech/speech about the activities of government and/or journalism on such matters."[536]

We also heard in our roundtable event on freedom of expression on 3 November that "public interest" may be a term that is better understood than the novel definitions of "journalistic content" and "content of democratic importance" and would more easily catch discussions on topics that may not be subject to high level political discussion.[537]

301. This term has been used by the Law Commission in developing their proposals for a harm-based offence, which the Government have indicated they are minded to implement.[538] One test that would need to be met is the defendant lacking a "reasonable excuse", and in deciding if this has been met "the court must have regard to whether the communication was or was meant as a contribution to a matter of public interest."[539] It is already used in relation to balancing the right of an individual to privacy with the freedom of the press, in whistle-blowing protections, and in the Freedom of Information Act, as well as the ECHR and Human Rights Act 1998, to which Ofcom is subject.[540] Given the precedent for use of this term more generally, and its likely role in determining whether a new harm-based offence has been committed, we wrote to the Government to seek their views on using this instead of the novel definitions in Clauses 13 and 14. They said in response:

> " … given the complexities of defining what the 'public interest' is, there may be concerns about requiring private companies to define what types of content are in the public interest. Our existing approach sets out more precisely the types of content that the government believes it is particularly important to protect."[541]

302. While determining public interest carries some of the inherent difficulties of asking providers to identify journalistic or democratic content, there is guidance available already

535 Q 144
536 Written evidence from Gavin Millar QC (OSB0221)
537 Written evidence from LSE, Department of Media & Communications—Freedom of Expression Roundtable (OSB0247)
538 Q 278
539 The Law Commission, *Modernising Communication Offences: a summary of the final report,* (July 2021) p 8: https://s3-eu-west-2.amazonaws.com/lawcom-prod-storage-11jsxou24uy7q/uploads/2021/07/Summary-of-Modernising-Communications-Offences-2021.pdf [accessed 1 November 2021]
540 European Court of Human Rights, *Guide on Article 10 of the European Convention on Human Rights: Freedom of Expression* (April 2021): https://www.echr.coe.int/documents/guide_art_10_eng.pdf [accessed 23 November 2021]; Information Commissioner's Office, *The Public Interest Test: Freedom of Information Act*: https://ico.org.uk/media/for-organisations/documents/1183/the_public_interest_test.pdf [accessed 1 December 2021]
541 Written evidence from Department of Digital, Culture, Media & Sport (OSB0243)

on applying public interest tests due to their use elsewhere. The ICO, for example, provides extensive guidance on applying public interest tests to Freedom of Information requests for public bodies[542]

Protecting high value speech

303. Clause 13 and 14 apply only to Category 1 providers who are subject to the extra duty to address content that is harmful to adults, which could carry a greater risk of censorship than the duties around illegal content.

304. ***We recommend that the news publisher content exemption is strengthened to include a requirement that news publisher content should not be moderated, restricted or removed unless it is content the publication of which clearly constitutes a criminal offence, or which has been found to be unlawful by order of a court within the appropriate jurisdiction. We recommend that the Government look at how bad actors can be excluded from the concept of news publisher. We suggest that they may wish to exclude those that have been repeatedly found to be in breach of The Ofcom Broadcasting Code, or are publications owned by foreign Governments. Ofcom should also examine the use of new or existing registers of publishers. We are concerned that some consumer and business magazines, and academic journals, may not be covered by the Clause 40 exemptions. We recommend that the Department consult with the relevant industry bodies to see how the exemption might be amended to cover this off, without creating loopholes in the legislation.***

305. **The draft Bill already makes a distinction between "news publisher content" and citizen journalism, in recognition that the former is subject to editorial control and there are existing mechanisms for accountability. There is also a clear difference between the categories, as one is based on "who" is sharing the content, and the other focuses on the purpose of the content, rather than the identity of those behind it. For both citizen journalism and content of democratic importance, the justification for special consideration appears to be that they are in the public interest to be shared. This should therefore be key to any final definition and providers will require guidance as to how to balance the risk of harm with the public interest. It is not, nor is it intended to be, a blanket exemption in the same way as that for news publisher content, but a counterbalance to prevent overzealous moderation, particularly in borderline cases.**

306. **Our recommendations to narrowly define content that is harmful to adults by way of reference to existing law should provide some of the extra clarity service providers need to help protect freedom of expression. At the same time, journalism and content of democratic importance have long been recognised as vital in a democratic society and should be given specific consideration and protection by providers, who have significant influence over the information we see. We have heard concerns around the definitions used however, and about the ability of the providers to interpret and apply them consistently. We feel that "democratic importance" may be both too broad—creating a loophole to be exploited by bad actors—and too narrow—excluding large parts of civil society. Similarly, we are concerned that any definition of journalistic content that is designed to capture citizen journalism would be so broad it would render the consistent application of the requirement almost impossible, and see the expedited**

542 Information Commissioner's Office, *The Public Interest Test: Freedom of Information Act*: https://ico.org.uk/media/for-organisations/documents/1183/the_public_interest_test.pdf [accessed 1 December 2021]

complaints route overwhelmed by people claiming without merit to be journalists in order to have their content reinstated. "Public interest" might be more useful in ensuring that content and activity is judged on its merit, rather than its author.

307. ***We recommend that the existing protections around journalistic content and content of democratic importance should be replaced by a single statutory requirement to have proportionate systems and process to protect 'content where there are reasonable grounds to believe it will be in the public interest'. Examples of content that would be likely to be in the public interest would be journalistic content, contributions to political or societal debate and whistleblowing. Ofcom should produce a binding Code of Practice on steps to be taken to protect such content and guidance on what is likely to be in the public interest, based on their existing experience and case law. This should include guidance on how appeals can be swiftly and fairly considered. Ofcom should provide guidance to companies in cases of systemic, unjustified take down of content that is likely to be in the public interest. This would amount to a failure to safeguard freedom of expression as required by the objectives of the legislation.***

8 Role of the regulator

The suitability of Ofcom as regulator

308. The draft Bill contains provision about the regulation of certain internet services by Ofcom (Clause 1(1)). In the White Paper response, the Government says it chose Ofcom to enforce the draft Bill because it is: "a well-established independent regulator with a strong reputation internationally and deep experience of balancing prevention of harm with freedom of speech considerations", as well as due to its role regulating video-sharing platforms.[543] Ofcom also carries out research on "market trends, online habits and attitudes" and will be able to "draw on strong relationships with industry, policymakers, academic experts, charities and other regulators".[544]

309. Witnesses generally agreed that Ofcom was the right choice of regulator.[545] Those who did oppose Ofcom's designation generally did so due to concerns about their expertise or that they might simply replicate the broadcasting model of regulation.[546] A third option arose during discussions, in which Ofcom's role could be understood as one of a series of co-regulating bodies with shared duties. For example, the Information Commissioner made the case that her Office (the ICO) should determine issues of data protection and privacy, whilst the IWF argued that they should be co-designated for content relating to the sexual abuse and exploitation of children.[547]

The powers of the regulator

310. Some of our witnesses expressed concerns that large tech companies might treat Ofcom with contempt, citing both the historic example of how bankers treated the Financial Standards Authority in the 1990s and more recent examples of how tech executives have "repeatedly shown contempt for elected officials and regulators".[548] In 2019, Facebook agreed to pay a £500,000 fine imposed by the ICO in relation to the processing and sharing of its users' personal data by Cambridge Analytica, but only as part of a settlement deal after an appeal to a First-tier Tribunal; the settlement included no admission of liability on Facebook's part.[549] In October 2021, the CMA fined Facebook £50.5 million for breaching an initial enforcement order relating to their merger with Giphy.[550] Joel Bamford, Senior Director of Mergers at the CMA, said: "We warned Facebook that its refusal to provide us with important information was a breach of the order but, even after losing its appeal

543 Department for Digital, Culture, Media and Sport and The Home Office, *Online Harms Consultation: Full Government Response to the consultation*, CP 354, December 2020, p 60: https://www.gov.uk/government/consultations/online-harms-white-paper/outcome/online-harms-white-paper-full-government-response [accessed 17 November 2021]

544 Ofcom, 'Ofcom to regulate harmful content online': https://www.ofcom.org.uk/about-ofcom/latest/features-and-news/ofcom-to-regulate-harmful-content-online [accessed 9 December 2021]

545 See, for example, written evidence from: Crown Prosecution Service (OSB0179), point 20, techUK (OSB0098), 5.1

546 Written evidence from: British & Irish Law, Education & Technology Association (OSB0073); 7.1–7.2.; Dr Kim Barker (Senior Lecturer in Law at Open University); Dr Olga Jurasz (Senior Lecturer in Law at Open University) (OSB0071); 11.1–11.4; Dr Edina Harbinja (Senior lecturer in law at Aston University, Aston Law School) (OSB0145); para 28–31; Dr Dimitris Xenos (Lecturer in Law at Cardiff Metropolitan University) (OSB0157)

547 Q 86; Written evidence from Internet Watch Foundation (IWF) (OSB0110), 7.1

548 Q 14, Written evidence from Center for Countering Digital Hate (OSB0009), p 9.

549 Hunton Andrews Kurth, 'Facebook reaches settlement with ICO over £500,000 data protection fine': https://www.huntonprivacyblog.com/2019/11/05/uk-ico-imposes-maximum-fine-on-facebook-for-compromising-user-data/ [accessed 1 December 2021]

550 Competition and Markets Authority, 'CMA fines Facebook over enforcement order breach': https://www.gov.uk/government/news/cma-fines-facebook-over-enforcement-order-breach [accessed 1 December 2021]

in two separate courts, Facebook continued to disregard its legal obligations."[551] On 30th November 2021, the CMA ordered Facebook to sell Giphy. Facebook intends to appeal the ruling.[552]

311. When Dame Melanie appeared before the Committee, she acknowledged that "this is a really challenging task", but asked: "more generally across the Bill, do we feel that we have what we need to act, and act quickly when we need to? The answer is broadly yes … The Bill gives us, broadly, the right overall things that we need."[553] Ofcom has suggested some small improvements to the Bill, including in safety duties and in the use of technology reports, but is largely content with the powers that have been extended to it in the Bill as it is currently drafted.

312. **Robust regulatory oversight is critical to ensuring the ambition of the Online Safety Bill is fully met. Tech companies must not be allowed to snub the regulator, to act with impunity, to continue to rely on self-regulation, or to abdicate responsibility for the harms which occur through the operation of their services or because of their governance structures. In turn, Ofcom must be able to move at pace to hold providers to account authoritatively to issue substantial fines, and assist the appropriate authorities with criminal prosecutions. The Bill extends substantial powers to the Regulator, but there are improvements to be made if the Government is to ensure the Bill is enforced effectively.**

International co-operation

313. Dame Melanie told us: "co-operation with international regulators is where we think the Bill could be slightly improved. We might find we need the ability to share with International Regulators in some circumstances."[554]

314. Ms Denham told us: "we need co-operation at a domestic level … but we also need the ability to collaborate and to cooperate at international level."[555]

315. ***Ofcom should have the power on the face of the Bill to share information and to co-operate with international regulators at its discretion.***

Risk Assessments

Naming the risk assessments

316. The draft Bill requires Ofcom to conduct an overall assessment of the types of risk across the range of online services and service providers to undertake their own risk assessments. During oral evidence, it was not always easy to differentiate between the Ofcom risk assessment and the service providers' own risk assessments. Dame Melanie

551 Competition and Markets Authority, 'CMA fines Facebook over enforcement order breach': https://www.gov.uk/government/news/cma-fines-facebook-over-enforcement-order-breach [accessed 1 December 2021]
552 Competition and Markets Authority, 'CMA directs Facebook to sell Giphy': https://www.gov.uk/government/news/cma-directs-facebook-to-sell-giphy [accessed 1 December 2021]
553 Q 250
554 Q 263
555 Q 86

said: "if we could have different names for our risk assessment and the platforms' risk assessment, it would be quite helpful."[556]

317. ***To help differentiate between the risk assessment undertaken by the regulator and that undertaken by the service providers, Ofcom's risk assessment should be renamed the "Ofcom register of risks of regulated services" (henceforth, register of risks). Ofcom should begin working on this immediately so that it is ready to be actioned when the Bill becomes law.***

Establishing risk profiles for companies of different kinds

318. Clause 61(3) states that: "Ofcom must develop risk profiles for different kinds of regulated services, categorising the services as Ofcom consider appropriate, taking into account (a) the characteristics of the services, and (b) the risk levels and other matters identified in the risk assessment". These characteristics include "the functionalities of the service, its user base, business model, governance and other systems and processes" (Clause 61(6)). In turn, service providers must consider the relevant risk profile when completing their risk assessments.

319. Although the Bill contains provisions for Ofcom to develop risk profiles based on the characteristics of services, some witnesses felt that this should be more central to the Bill, and that further clarity was needed on which characteristics Ofcom should consider. 5Rights argued that Ofcom should develop risk profiles for different kinds of regulated services, which should consider several factors when assessing the risk posed by a service including the characteristics of the service, platform design, risk level, and the service's business model and its overall corporate aim.[557] The UK Interactive Entertainment Association argued:

> " … a proportionate approach should be taken to the extent of requirements for transparency and risk assessment on different services. For instance, online services with minimal user-to-user interaction should not be expected to bear the same burdens as full social media platforms or other online services where user-to-user interaction is core to the service's offering."[558]

Similarly, Match Group argued that as its business model relies on user subscriptions, rather than "revenue streams like advertisements and data harvesting", it should be grouped with other similar businesses for the purposes of risk assessment.[559]

320. Allowing a more holistic approach to risk assessment will allow Ofcom to meet the challenges of regulating emerging technologies and their associated risks. We were challenged by Dr Edina Harbinja, Senior Lecturer in Law at Aston University, to consider, for example, how the Bill will manage "virtual reality, augmented reality, and the Metaverse that Facebook is now building", alongside "deep fakes, chatbots of us after we die, or before".[560] Both the Bill and its regulatory duties need to be 'future proof' and able to manage the emergence of future technologies and platforms, allowing Ofcom to

556 Q 261
557 Written evidence from 5Rights Foundation (OSB0096), point 5; see also Match Group (OSB0053), 56c.
558 Written evidence from UK Interactive Entertainment (OSB0080), para 24.
559 Written evidence from Match Group (OSB0053). See also written evidence from Microsoft (OSB0076)
560 Q 77. See also Q 271.

produce risk profiles for similar businesses will allow it more flexibility than the current system.

321. According to the Government's April 2021 Impact Assessment, 81 per cent of businesses in scope of the Bill are likely to be microbusinesses.[561] The Coalition for a Digital Economy argues that "smaller businesses should not be burdened with the same obligations as their larger counterparts", warning that "the proposed framework could have a significant and disproportionately negative financial impact on start-ups" with "a chilling impact on digital competition".[562] A study commissioned by DCMS which found that smaller services with video-sharing capabilities were spending over £45 per user to protect them from content that creates a risk of harm, compared to the biggest services, which spent £0.25–50 per user.[563]

322. We heard that a service's size or number of employees is not necessarily a vector for its risk, and thus a small service could potentially be a very harmful one. Characteristics of a service which were suggested might influence which risk profile companies will come under included:

a) Risks created by algorithms, including the promotion of divisive content and "out-group animosity" [and] rewarding hostility online with virality"[564]; and "pushing" people towards extremist content and groups.[565]

b) Risks created by a reliance on AI moderation, including risks to freedom of expression, for example, the over-zealous moderation of LGBTQ+ groups or women.[566] The Lords Communications and Digital Committee heard that content is twice as likely to be deleted if it is in Arabic or Urdu than if it is in English, whilst Legal to Say, Legal to Type reported that leading AI models for hate speech are 2.2 times more likely to flag tweets written in African American English.[567] We also heard evidence from the Board of Deputies of British Jews asking for in-country teams monitoring suspected breaches of community guidelines, as they "will be more likely to have political, cultural, and linguistic context for cases".[568]

c) Risks caused by unlimited, "one-click" sharing leading to e.g. the viral spread of false or illegal content, especially on end-to-end encrypted services. One 2018 study found the false stories were 70 per cent more likely to get retweeted than accurate stories.[569]

d) Risks caused by "designed addiction", including infinite scrolling pages and automatic, frictionless recommender tools which maximise "engagement" e.g. time spent on

561 Written evidence from Coadec (OSB0029), XXXII.

562 Written evidence from: Coadec (OSB0029), XXVIII; VI; see also: Snap Inc. (OSB0012), pp 6–7; British & Irish Law, Education & Technology Association (OSB0073), 6.3.1.

563 Written evidence from Coadec (OSB0029), XXXI

564 Written evidence from HOPE not hate (OSB0048), 6.17.

565 Q 155.

566 See written evidence from: LGBT Foundation (OSB0191); *ibid.*; Legal to Say, Legal to Type (OSB0049); p 3; Reddit, Inc. (OSB0058), p 7. We heard from Mumsnet that their pages are repeatedly blacklisted as 'obscene' by algorithms used by programmatic advertising agencies "because our users post about breasts (in the context of breastfeeding, or in discussions of clothing shapes) and vulvas and vaginas (in the context of discussions of their health and wellbeing). Trained on databases of largely male speech, algorithms are simply unable to interpret non-pornographic discussions of female anatomy". Mumsnet (OSB0031).

567 Q 44, Written evidence from Legal to Say, Legal to Type (OSB0049), p 3

568 Written evidence from Board of Deputies of British Jews (OSB0043), p 3. See also Q 154.

569 Written evidence from RSA (Royal Society for the Encouragement of Arts, Manufactures and Commerce) (OSB0070), 7.III.

the service, which then allows the service provider to generate saleable user data and advertising revenue. This type of design can, in some cases, take people down "rabbit holes that lead to a warren of conspiracy" and normalise content that creates a risk of harm or sensationalist content.[570] As 5Rights observes, "pro-suicide, self-harm, or eating disorder content is far more dangerous when served up automatically, proactively, and repeatedly by the recommender systems of platforms popular with young people".[571]

e) Risk of unsupervised contact between adults and children which may create circumstances where children can be "groomed" for abuse online or offline.[572] The NSPCC reports that "when children are contacted [online] by someone they don't know in person, in nearly three quarters (74 per cent) of cases, this contact initially takes place by private message."[573]

f) Risks caused by surveillance advertising (also called targeted or microtargeted advertising).[574] Surveillance advertising "requires the large-scale collection, profiling and sharing" of user data which "can be harvested for behavioural profiling and recommender algorithms which maximise "engagement" … to the detriment of all other considerations".[575] 5Rights states "there is not a single online harm or socio-digital problem that is not made worse by micro-targeting".[576] Surveillance advertising can be particularly harmful to children. The End Surveillance Advertising to Kids Coalition suggest that "based on average online time, a third of 14-year-olds could be exposed to 1,332 adverts a day—ten to twenty times as many adverts as children see on TV alone".[577] "Surveillance advertising frequently enables children to be targeted with harmful products".[578] A 2021 study found that it was possible to target 13–17 year olds on Facebook with adverts "based on interests such as alcohol, smoking and vaping, gambling, extreme weight loss, fast foods and online dating services".[579] Research published by New York University's Cyber Security for Democracy research team and imec-DistriNet at KU Leuven in Belgium has also highlighted the failings of Facebook's monitoring of political adverts. Between July 2020 and February 2021, globally, Facebook made the wrong decision for 83 percent of ads that had not been declared as political by their advertisers. Facebook both overcounted and undercounted political advertisements in this group. They also missed a higher proportion of political

570 Written evidence from Center for Countering Digital Hate (OSB0009), p 1.
571 Written evidence from Global Action Plan, on behalf of the End Surveillance Advertising to Kids coalition, The Mission and Public Affairs Council of the Church of England, Global Witness, New Economics Foundation, Foxglove Legal, Fairplay, 5Rights Foundation, Andrew Simms, New Weather Institute, Dr Elly Hanson, Avaaz (OSB0150), p 2.
572 Written evidence from: Parent Zone (OSB0124), p 2; Yoti (OSB0130), p 8; Reset (OSB0138), appendix 1; The Arise Foundation (OSB0198), p 5; Barnardo's (OSB0017), p 1.
573 Written evidence from NSPCC (OSB0109), p 4.
574 Written evidence from Global Action Plan (OSB0027)
575 Written evidence from Global Action Plan, on behalf of the End Surveillance Advertising to Kids coalition, The Mission and Public Affairs Council of the Church of England, Global Witness, New Economics Foundation, Foxglove Legal, Fairplay, 5Rights Foundation, Andrew Simms, New Weather Institute, Dr Elly Hanson, Avaaz (OSB0150)
576 Written evidence from Global Action Plan, on behalf of the End Surveillance Advertising to Kids coalition, The Mission and Public Affairs Council of the Church of England, Global Witness, New Economics Foundation, Foxglove Legal, Fairplay, 5Rights Foundation, Andrew Simms, New Weather Institute, Dr Elly Hanson, Avaaz (OSB0150)
577 Written evidence from Global Action Plan, on behalf of the End Surveillance Advertising to Kids coalition, The Mission and Public Affairs Council of the Church of England, Global Witness, New Economics Foundation, Foxglove Legal, Fairplay, 5Rights Foundation, Andrew Simms, New Weather Institute, Dr Elly Hanson, Avaaz (OSB0150)
578 Written evidence from Global Action Plan (OSB0027)
579 Written evidence from Global Action Plan, on behalf of the End Surveillance Advertising to Kids coalition, The Mission and Public Affairs Council of the Church of England, Global Witness, New Economics Foundation, Foxglove Legal, Fairplay, 5Rights Foundation, Andrew Simms, New Weather Institute, Dr Elly Hanson, Avaaz (OSB0150)

advertisements outside the United States. However, Facebook also allowed more than 70,000 political adverts to run during its moratorium on political ads around the U.S. 2020 elections.[580]

g) Risks caused by features designed to enhance reach or to maximise 'network effect' such as live-streaming, the ability to create 'groups' or to add multiple contacts/users at the same time.

h) Such other risks as Ofcom identifies in its overall register of risks.

323. ***The Bill's provision that Ofcom should develop risk profiles based on the characteristics of services should be strengthened. Ofcom should begin drawing up risk profiles immediately so that they are ready to be actioned when the Bill becomes law. Risk profiles should reflect differences in the characteristics of the service. These could include (but are not limited to) risks created by algorithms; risks created by a reliance on artificial intelligence moderation; risks created by unlimited 'one-click' sharing; risks caused by "engagement" maximising design features; risk of unsupervised contact between adults and children which may give rise to grooming; risks caused by surveillance advertising; and such other risks as Ofcom identifies in its overall risk assessment, as well as platform design, risk level, end-to-end encryption, algorithmic design, safety by design measures, and the service's business model and overall corporate aim. Ofcom should also be able to take into account whether a company has been the subject of a super complaint, other legal proceedings or publicly documented evidence of poor performance e.g. independent research, a poor monitoring report in the EU's Code of Conduct for Illegal Hate, or whistleblowers' evidence.***

Enforcement against the safety duties

324. Ofcom has expressed concerns that they might struggle to build up enough momentum for enforcement action if they are compelled to keep going back to the start of the process whenever they identify an issue. Dame Melanie told the Committee:

> "We think there is a slight risk that a service may not identify a risk, and then not be required under the safety duties to address that risk. Our concern is that, if we did then identify one of those problems, we would have to go all the way back to the risk assessments and get them to do it again before we were able to engage the safety duties for any kind of enforcement action."[581]

325. ***The Bill should be amended to clarify that Ofcom is able to take enforcement action if it identifies a breach of the safety duties, without requiring a provider to redo a risk assessment.***

580 Research published on 9 December 2021. Researchers with imec-DistriNet at KU Leuven in Belgium and New York University's Cybersecurity for Democracy conducted a comprehensive audit of Facebook's political advertisement detection and policy enforcement. The researchers examined 33.8 million Facebook ads that ran between July 2020 and February 2021—a timeframe that included elections in both the U.S. and Brazil. This is the first known study to quantify the performance of Facebook's political ad policy enforcement system at a large and representative scale.

581 Q 252

Establishing minimum quality standards for risk assessments

326. Under Clause 62, the Regulator will prepare guidance for providers of regulated services to assist them in complying with their duties to carry out risk assessments. However, the Bill as drafted does not specify minimum quality standards for the providers' risk assessments or require companies to take the risk profile produced by the Regulator into account when producing their own risk assessments. This was a source of widespread concern among witnesses, who argued that a lack of quality standards could give service providers an incentive to underplay or not go looking for risks that their services might cause.[582] Ms Haugen told us:

> "I believe that, if Facebook does not have standards for those risk assessments, it will give you a bad risk assessment, because Facebook has established over and over again that when asked for information it misleads the public. I do not have any expectation that it will give you a good risk assessment unless you articulate what a good one looks like."[583]

327. Ofcom agreed that the Bill would benefit from stronger provisions relating to minimum quality standards for risk assessments. Dame Melanie told us: "the way the Bill is drawn on risk assessments is good in large part … we are broadly there, but with the gap of adequacy in standards not being quite strong enough at the moment"[584] and "I certainly think it should be clearer in the Bill that risk assessments need to be of a certain standard".[585] Ofcom's written evidence elaborated, stating that whilst the duties on providers to complete a risk assessment were clear, it would be harder for them to take enforcement action against a provider for deliberately or negligently understating risk.[586]

328. Ofcom also suggested that the concept of "reasonable foreseeability" should be introduced into the risk assessment, meaning that references to "risk" or "level of risk" should mean risks or levels of risk that are reasonably foreseeable.[587] The idea that "companies should take reasonable steps to prevent reasonably foreseeable harms that occur through the operation of their services" was mooted by Carnegie UK Trust and supported by the NSPCC; a duty to address reasonably foreseeable harms was also proposed by the Antisemitism Policy Trust.[588] Facebook challenged the use of the term "reasonably foreseeable" in the Bill and asked for it to be further defined.[589] We note that reasonable foreseeability is both an objective standard and an established principle in law, and use it here to mean that a reasonable person could reasonably foresee that a given risk would occur on a service.

---

582 Written evidence from: Centenary Action Group, Glitch, Antisemitism Policy Trust, Stonewall, Women's Aid, Compassion in Politics, End Violence Against Women Coalition, Imkaan, Inclusion London, The Traveller Movement (OSB0047), p 9; Reset (OSB0138); See also Compassion in Politics (OSB0050), p 1; Carnegie UK (OSB0095), p 7: "Without regulation, internal risk assessments would then underplay the probability of harm, lack rigour or be quashed at a senior level."; and NSPCC (OSB0109), p 3: "the legislation introduces a risk of moral hazard for online services to overlook the more risk-inducing or complex aspects of their services."

583 Q 173

584 Q 252

585 Q 257

586 Written evidence from Ofcom (OSB0223)

587 Written evidence from Ofcom (OSB0223)

588 Written evidence from: Carnegie UK (OSB0095), p 1; NSPCC (OSB0109), p 2; Antisemitism Policy Trust (OSB0005), p 2. See also Mr John Carr (Secretary at Children's Charities' Coalition on Internet Safety) (OSB0167), para 29.

589 Written evidence from Facebook (OSB0147), p 14. See also written evidence from Dame Margaret Hodge (Member of Parliament for Barking and Dagenham at House of Commons) (OSB0201), pp 13–14.

329. During oral evidence, Mr Philp stated:

> " … we have constructed this so that there is no wiggle room for platforms that may try to fudge their risk assessment in relation to children. Ofcom will do its own sector risk assessment first, and the companies' own risk assessments will be measured against that … We will make sure that they cannot get themselves some sort of get out of jail free card by fudging or diluting their risk assessment. That will not be acceptable at all".[590]

330. We were told there is no shortage of models for minimum standards of risk assessment for regulators in the financial sector to implement minimum standards, with both clearly laid out threshold conditions—minimum standards that regulated entities must meet at all times in order to be permitted to carry on the regulated activities in which they are engaged—and high-level fundamental rules that express the general objective of promoting the safety of regulated entities.[591] The ICO sets minimum standards for the Data Protection regime, and the language of the "minimum standards" framework for a Data Protection Impact Assessment is well understood among regulated entities. Such a framework of minimum standards for risk assessments would be adaptable and allow for scalability, so even the smallest service providers could design and implement one.

331. The Government told us:

> "In line with the risk-based and proportionate approach to regulation, the Bill does not additionally seek to set a specific standard to determine what needs to be done to comply with [risk assessment] obligations. In this case companies will need to refer to the guidance about compliance with their assessment duties which Ofcom is required to publish under Clause 62, which should include risk profiles to establish the standards expected of them".[592]

332. **It should not be possible for a service provider to underestimate the level of risk on their service without fear of sanction. If Ofcom suspects such a breach, it should have the power to investigate, and, if necessary, to take swift action. We are not convinced that the draft Bill as it currently stands achieves this.**

333. ***Ofcom should be required to set binding minimum standards for the accuracy and completeness of risk assessments. Ofcom must be able to require a provider who returns a poor or incomplete risk assessment to redo that risk assessment. Risk assessments should be carried out by service providers as a response to the Online Safety Act before new products and services are rolled out, during the design process of new features, and kept up to date as they are implemented.***

334. ***The required content of service providers' risk assessments should follow the risk profiles developed by Ofcom, which in turn should be based on the differences in the characteristics of the service, platform design, risk level, and the service's business model and overall corporate aim. For example, a provider that does not have an engagement-based service would not need to address irrelevant risks associated with virality, whilst a site containing adult content would have to address the higher level of risks associated with children accessing the site.***

590 Q 284
591 Written evidence from: NSPCC (OSB0109); Antisemitism Policy Trust (OSB0005), point 4.
592 Written evidence from Department of Digital, Culture, Media & Sport (OSB0243), Q 20

335. ***The Bill should be amended to clarify that risk assessments should be directed to "reasonably foreseeable" risks, to allow Ofcom greater leeway to take enforcement action against a company that conducts an inadequate risk assessment.***

336. ***Ofcom should look to the Data Protection Impact Assessment as they come to form their own guidance for minimum standards for risk assessments for regulated services.***

Powers of audit

337. Algorithms can both increase and reduce the spread of content that creates a risk of harm. As Full Fact put it: "content moderation algorithms can do real good if they work well, and if they malfunction, they can cause real harm", yet "the safety consequences of deploying a certain content moderation algorithm are not always obvious".[593] Throughout the inquiry, we heard from witnesses who were concerned that Ofcom's powers of audit, particularly with regard to algorithms, did not go far enough, and who called, in the words of the Ada Lovelace Institute, for Ofcom to be given the power to "perform technical audits, assessments, and monitoring of platform behaviour, including algorithmic behaviour, whenever Ofcom deems appropriate."[594] As Mr Ahmed told the Committee: "[The Bill] needs independent auditing powers and the ability to go in and get other bodies, not just self-reporting. You cannot ask Facebook to mark their own homework. That is why we are where we are. Self-regulation is over. It has to be over."[595]

338. There is something of a discrepancy between the ICO's sense of what additional powers Ofcom needs "to be able to look under the bonnet" of the tech companies and what Ofcom feels the draft Bill already empowers it to do. When asked whether Ofcom's auditing powers were as strong as those held by the ICO, Ms Denham stated she "would like to see stronger powers of compulsory audit" given to Ofcom by the Bill.[596] However, Dame Melanie stated that she considers Ofcom's power to ask for a skilled person's report to be "the same sort of thing as, for example, the Information Commissioner's Office is able to use to get under the bonnet when it needs to".[597] Ofcom clarified in writing that they consider they "have broadly similar investigative and information gathering powers under the draft Online Safety Bill to those ICO has to carry out audits."[598]

593 Written evidence from Full Fact (OSB0056), p 5

594 Written evidence from Ada Lovelace Institute (OSB0101). See also written evidence from: Reset (OSB0203); Center for Countering Digital Hate (OSB0009); Common Sense (OSB0018); Full Fact (OSB0056), p 1; APPG Coalition (OSB0202), p 3; Common Sense (OSB0018), p 2; Glitch (OSB0097), p 9; Demos (OSB0159), p 4; Written evidence from LSE roundtable LSE Department of Media and Communications—Anonymity & Age Verification Roundtable (OSB0236) p 2

595 Q 14

596 Q 85

597 Q 250

598 Written evidence from Ofcom (OSB0223)

Box 1: The audit powers of the ICO

1) The Information Commissioner wrote to us to detailing the ICO's audit powers, which are "either consensual or compulsory, and may be deployed in an ex-post and ex-ante manner".

2) The Information Commissioner explained:

- "From an ex-post perspective … the ICO can seek its own assurances that an enforcement notice has been complied with by directly auditing the current practices in an organisation."
- "The majority of the ICO's audit activity however takes place where we have concerns about ongoing data processing … but the threshold for taking immediate enforcement action has not been reached. In such cases, we can undertake an audit … if our audit raises concerns then this may lead to a subsequent enforcement notice."
- "The deployment of audit powers as a check against compliance with a notice to improve data practices, is an important examination tool for the ICO; whilst proactive audits based on concerns also provide a level of consistent assurance for the public that improvements have been made by an organisation to the extent expected by the independent Regulator."

3) Ofcom stated that they "consider that Ofcom would have broadly similar investigative and information gathering powers under the draft Online Safety Bill to those ICO has to carry out audits".

Source: Information Commissioner's Office (OSB0210) 4.7–4.10

339. ***In bringing forward the final Bill, we recommend the Government publish an assessment of the audit powers given to Ofcom and a comparison to those held by the Information Commissioner's Office and the Financial Conduct Authority. Parliament should be reassured that the Bill will give Ofcom a suite of powers to match those of similar regulators. Within six months of the Act becoming law, Ofcom should report to Parliament on how it has used those powers.***

340. ***We recommend that the largest and highest-risk providers should be placed under a statutory responsibility to commission annual, independent third-party audits of the effects of their algorithms, and of their risk assessments and transparency reports. Ofcom should be given the explicit power to review these and undertake its own audit of these or any other regulated service when it feels it is required. Ofcom should develop a framework for the effective regulation of algorithms based on the requirement for, and auditing of, risk assessments.***

Coregulation

341. In their "Response to the Consultation on the Online Harms White Paper", the Government stated that it would "work with Ofcom to ensure that the regulator is able to work effectively with a range of organisations. This will be delivered through a range of

means **including co-designation powers**, memorandums of understanding, forums, and networks" (our emphasis).[599]

342. The current draft Bill does not explicitly mention co-regulation (with other regulators) or co-designation (with third parties) powers or give any detail on how the Government or Ofcom intends to achieve this. Ofcom has stated that it has delegation arrangements in place in other situations through the Deregulation and Contracting Out Act 1994 and the Communications Act 2003 and has suggested that it could delegate functions in this manner without adding additional provisions on the face of the Bill.[600] However, the IWF suggests, and we agree, that "it would have been beneficial to see more information published alongside the Bill about how such co-designation might be achieved or even a timeline on when such decisions will be taken". This would have helped such bodies prepare.[601]

343. During oral evidence with financial service regulators, the Committee heard that there was no objection to a cooperation duty, and a great appetite, in the words of Mark Steward, Executive Director of Enforcement and Market Oversight for the FCA, for "allowing information and intelligence to be shared between all the regulators on a mutual basis".[602] It was felt that this was important to prevent "things falling between the cracks". Michael Grenfell, Executive Director for Enforcement at the CMA, for example, suggested that Ofcom was unlikely to prioritise smaller consumer protection breaches, and so it might be prudent to "give parallel concurrent powers to other regulators too … to enforce those bits."

344. On 1st July 2020, Ofcom, the ICO, and the CMA came together to form the Digital Regulation Cooperation Forum (DRCF), which "aims to strengthen existing collaboration and coordination between the three regulators by harnessing their collective expertise when data, privacy, competition, communications, and content interact."[603] The Financial Conduct Authority (FCA) joined the DRCF as a full member in April 2021 (having previously been an observer member). The Committee welcomes the foundation of the DRCF, which is, in the words of the Information Commissioner, "a pathfinder in the areas of safety online and data protection and competition [which is] setting international norms now".[604] Ofcom, however, told us: "we and our fellow regulators could do with a little more by way of legislative support to be able to work together. I am thinking of things such as information powers and requirements to consult each other."[605] In their recent report *Digital regulation: joined-up and accountable*, the House of Lords Communications and Digital Committee identified that for the DRCF to operate effectively, cooperation between its members needs to be extended and formalised. Regulators within the DRCF need to be subject to statutory requirements to cooperate and consult with one another and to share information. This would allow them to share their powers and would facilitate joint

599 Department for Digital, Culture, Media and Sport and The Home Office, *Online Harms Consultation: Full Government Response to the consultation,* CP 354, December 2020, p 60: https://www.gov.uk/government/consultations/online-harms-white-paper/outcome/online-harms-white-paper-full-government-response [accessed 17 November 2021]

600 Written evidence from Ofcom (OSB0288)

601 Written evidence from Internet Watch Foundation (IWF) (OSB0110), 6.3, 6.4. See also written evidence from TalkTalk (OSB0200), pp 8–9.

602 Q 118–127

603 Information Commissioner's Office, 'UK regulators join forces to ensure online services work well for consumers and businesses': https://ico.org.uk/about-the-ico/news-and-events/news-and-blogs/2020/07/uk-regulators-join-forces-to-ensure-online-services-work-well-for-consumers-and-businesses/ [accessed 1 December 2021]

604 Q 90

605 Q 263

regulation. Placing the DRCF on a statutory footing with the power to resolve conflicts by directing its members would further support its functions.[606]

345. The ICO also holds the position that legislative support would aid cooperation between regulators. Ms Denham called for the regulators to be given "duties to respect the other [regulators'] regulatory objectives as well as information sharing between the regulators".[607] Ofcom similarly wrote to us to say: "we need to ensure that we are able to share information as needed, subject to appropriate safeguard, and that we are able to consult the ICO on privacy matters."[608] We note that Ofcom and the ICO already have a Memorandum of Understanding but agree that further clarity on the bounds of their respective remits and a greater emphasis on cooperation and sharing would provide clarity for both regulators and regulated services.

346. ***In taking on its responsibilities under the Bill, Ofcom will be working with a network of other regulators and third parties already working in the digital world. We recommend that the Bill provide a framework for how these bodies will work together including when and how they will share powers, take joint action, and conduct joint investigations.***

347. ***We reiterate the recommendations by the House of Lords Communications and Digital Committee in their Digital Regulation report: that regulators in the Digital Regulation Cooperation Forum should be under a statutory requirement to cooperate and consult with one another, such that they must respect one another's objectives, share information, share powers, take joint action, and conduct joint investigations; and that to further support coordination and cooperation between digital regulators including Ofcom, the Digital Regulation Cooperation Forum should be placed on a statutory footing with the power to resolve conflicts by directing its members.***

348. ***The draft Bill does not give Ofcom co-designatory powers. Ofcom is confident that it will be able to co-designate through other means. The Government must ensure that Ofcom has the power to co-designate efficiently and effectively, and if it does not, this power should be established on the face of the Bill.***

The regulation of child sexual exploitation and abuse material

349. Some concerns have been raised about whether Ofcom is the correct regulator to deal with CSEA material. Dr Dimitris Xenos, Lecturer in Law at Cardiff Metropolitan University, argued that Ofcom is a "soft moderator", and that "some types of harm, especially those relating to extreme and child pornography, torture and serious violence should be organised under a different regulatory framework with different legal obligations (criminal liability) and more robust monitoring bodies, such as the CPS."[609] In turn, the CPS "supports Ofcom as the chosen appointed regulator for the draft Online Safety Bill" but is concerned that the Bill "lacks detail about how Ofcom will interact with law enforcement and the CPS as a regulator of illegal content", and recommends that "Ofcom should establish internal mechanisms for reporting any indecent or illegal material they

606 Communications and Digital Committee, *Digital regulation: joined-up and accountable* (3rd Report, Session 2021–22, HL Paper 126)
607 Q 90
608 Written evidence from Ofcom (OSB0223)
609 Written evidence from Dr Dimitris Xenos (Lecturer in Law at Cardiff Metropolitan University) (OSB0157), p 4–5

receive directly to law enforcement and/or the IWF."[610] The IWF made a persuasive case that they should be co-designated by Ofcom to regulate CSEA content, an argument supported by the CPS and by TalkTalk.[611] Ofcom mentioned the need to have a "strong partnership" with "third-sector organisations like the IWF" but no formal arrangement has been made, reflecting the general lack of clarity on co-designation discussed above.[612]

350. The IWF specialises in tackling online CSEA material hosted anywhere in the world and non-photographic CSEA images hosted in the UK. In 2020, the UK's Independent Inquiry into Child Sexual Abuse described it as a "genuine success story" which "deserves to be publicly acknowledged as a vital part of how and why comparatively little child sexual abuse is hosted in the UK".[613] Many UK Internet Service Providers are members of the IWF or use its watch list to block CSEA content via third parties.[614]

351. The IWF has a Memorandum of Understanding between the CPS and the National Police Chiefs' Council, which "ensures immunity from prosecution for our analysts and recognises our role as "the appropriate authority" for the issuing of Takedown Notices in the UK".[615] The CPS was concerned that Ofcom might "receive unsolicited illegal material from the public through their public complaints procedures, and some of this material could relate to CSEA material. Receiving this material would constitute an offence under section 1 of the Protection of Children Act 1978".[616]

352. ***During the course of its duties, Ofcom will be required to investigate companies for a range of breaches, some of which will relate to suspected or known child sexual exploitation and abuse material. As child sexual exploitation and abuse investigations lie so far outside Ofcom's normal duties, we expect Ofcom to work closely with experts like the Internet Watch Foundation, to develop and update the child sexual exploitation and abuse Code of Practice; monitor providers to ensure compliance with the child sexual exploitation and abuse code; and during investigations relating to child sexual exploitation and abuse content.***

353. ***Ofcom may receive unsolicited child sexual exploitation and abuse material which would constitute an offence under Section 1 of the Protection of Children Act 1978. The Bill should be amended to provide Ofcom with a specific defence in law to allow it to perform its duties in this area without inadvertently committing an offence.***

Codes of Practice

354. During oral evidence, the Secretary of State was adamant that this Bill "has to be watertight. That includes the codes of practice and the terms and conditions."[617] The Bill requires Ofcom to prepare Codes of Practice for providers of regulated services describing recommended steps for the purposes of compliance with duties in relation to terrorism, CSEA, and other relevant duties. Currently, whilst safety duties under the Bill are binding, Codes of Practice are not. A service provider can demonstrate compliance with a safety

610 Written evidence from Crown Prosecution Service (OSB0179), para 20, 23, 24.
611 Written evidence from: Internet Watch Foundation (IWF) (OSB0110), para 1.7; TalkTalk (OSB0200), p 7; Crown Prosecution Service (OSB0179), para 23, 24.
612 Written evidence from Ofcom (OSB0021)
613 Written evidence from Internet Watch Foundation (IWF) (OSB0110), para 1.4
614 Written evidence from ISPA (The Internet Service Provider Association) (OSB0059), p 1.
615 Written evidence from Internet Watch Foundation (IWF) (OSB0110), para 3.2–3.3
616 Written evidence from Crown Prosecution Service (OSB0179), para 23.
617 Q 276

duty by taking steps set out in a Code of Practice, or in another way which would be assessed by Ofcom having regard to the online safety objectives and protections for freedom of speech and privacy.[618] As such, there are multiple routes that service providers can take to fulfil their safety duties.

355. Ofcom describe this approach as "leaning towards flexibility" but acknowledge that it "will make it harder for Ofcom to judge compliance with safety duties and ultimately to enforce against any breaches, particularly if the safety duties themselves are specified at a high level."[619] When asked whether she thought Ofcom's Codes of Practice should be binding, Dame Melanie replied:

> "They are statutory codes, but the way platforms are able to discharge their duties, particularly their safety duties, means that they can choose another route. … At some point it is right that there is flexibility for services to determine how they address the safety duties. At the same time, that makes it potentially harder for us to prove a breach against those duties, because it leaves open so many different options through which they could be addressed."[620]

356. Richard Wronka, Director for Online Harms at Ofcom, clarified that Ofcom can "make a requirement on services to take specific steps where we have identified that they have breached their safety duties", but they cannot set out "binding requirements before the event through codes of practice."[621] Reset advocated for "minimum standards for compliance with the safety duties, perhaps through binding codes of practice."[622]

357. During oral evidence, there was a lack of clarity about whether amendments to the Codes of Practice would be subject to affirmative or negative parliamentary procedure.[623] Amendments to the Codes of Practice require only negative procedure, which Carnegie UK Trust have argued "gives the executive too much power on matters of free expression" and advocated instead for giving Parliament "more influence at the outset and more flexibility for the Regulator downstream."[624]

618 Draft Online Safety Bill, CP 405, May 2021, Clause 36
619 Written evidence from Ofcom (OSB0021)
620 Q 252
621 Q 252
622 Written evidence from Reset (OSB0138)
623 Q 285
624 Written evidence from Carnegie UK (OSB0095)

Box 2: Indicative list of Codes of Practice

- Terrorism (interim code should be updated)
- CSEA (interim code should be updated)
- Regulated content and activity for adults
- Child online safety
- Safety by design
- Age assurance
- Freedom of speech (including content in the public interest)
- Moderation, reporting, complaints, and redress
- Accessibility and consistency of terms and conditions (including Online Safety Policies)
- Transparency reporting
- Digital literacy
- Risk assessment
- Any other Codes of Practice the Regulator deems necessary

358. ***The Bill should be amended to make clear that Codes of Practice should be binding on providers. Any flexibility should be entirely in the hands of and at the discretion of the Regulator, which should have the power to set minimum standards expected of providers. They should be subject to affirmative procedure in all cases.***

359. ***Ofcom should start working on Codes of Practice immediately, so they are ready for enforcement as soon as the Bill becomes law. A provisional list of Codes of Practice, including, but not necessarily limited to, those listed in Box 2 above should be included on the face of the Bill. Some of the Codes should be delegated to co-designated bodies with relevant expertise, which would allow work on multiple Codes to happen simultaneously and thus the entire endeavour to be completed more quickly. Once the Codes of Practice are completed, they should be published.***

Criminal liability

360. The draft Bill provides for criminal liability for senior managers who fail to comply with the information notice provisions. This provision can only come into force after the two-year review of the legislation required under Clause 115.[625] Throughout this inquiry, there has been debate about when criminal liability should come into force. The CCDH said that a two-year delay "would be a grave mistake" as "tech executives have repeatedly shown contempt for elected officials and regulators".[626] When she appeared before the

625 The offence is in Draft Online Safety Bill, CP 405, May 2021, Clause 73, the requirement for commencement is in Clause 140(4b). Written evidence from Department of Digital, Culture, Media & Sport (OSB0243), Q 19

626 Written evidence from Center for Countering Digital Hate (OSB0009)

Committee, the Secretary of State told us: "I say to the platforms, 'Take note now. It will not be two years. We are looking at truncating that to a very much shorter timeframe. … I am looking at three to six months for criminal liability."[627] We saw recently how Facebook criticised the CMA in respect of their £50.5 million fine for "consciously refusing" to supply all the required information under an Initial Enforcement Order.[628] Facebook described the fine as "grossly unreasonable and disproportionate" and questioned the CMA's authority to enforce it.[629]

361. We welcome the introduction of criminal sanctions as a demonstration of the seriousness with which the Government is taking the matter of holding tech executives to account. However, as it stands, a named senior manager can only be held liable for the following offences: failure to comply with an information notice; deliberately or recklessly providing or publishing false information; providing or publishing encrypted information with the intention of preventing the Regulator from understanding such information. As the NSPCC has pointed out, these criminal sanctions "would not apply in respect of actual product or safety decisions."[630] Ms Pelham told us she considered criminal responsibility for failure to ensure online safety was the single most impactful thing that the Bill could do.[631]

362. Governance structures were a key issue that came up in our evidence. Ms Haugen told us: "I think there is a real problem with the left hand not speaking to the right hand at Facebook", describing it as "a world that is too flat, where no one is really responsible". She put it plainly: "the organisational choices of Facebook are introducing systemic risk."[632]

363. Our sessions with the major service providers did little to allay our concerns about their governance structures. Antigone Davis, the Global Head of Safety at Facebook, does not report to Facebook's Audit and Risk Oversight Committee and could not tell us whether papers had been submitted to that Committee detailing the online harms discussed at our session, nor who would be submitting the risk assessment when the Bill becomes law.[633] In a subsequent letter, Facebook told us that the Committee reviews Community Standards and Safety Issues "at least annually" and are generally briefed "twice a year".[634] Leslie Miller, Vice President of Government Affairs and Public Policy at YouTube, assured us that the YouTube risk assessment "will certainly have a review by executives" but could not be more specific; Markham C. Erickson, who holds the same position at Google, could only say "it will be reviewed at the appropriate level" at Google.[635] On the other hand, Twitter assured the Committee that they produce a range of risk assessments, including a corporate governance risk document which their board sees "several times a year".[636]

627 Q 276
628 Competition and Markets Authority, *CMA fines Facebook over enforcement order breach* (October 2021): https://www.gov.uk/government/news/cma-fines-facebook-over-enforcement-order-breach [accessed 1 December 2021]
629 'Facebook criticises UK competition watchdog's concern over Giphy takeover', *Evening Standard (8 September 2021):* https://www.standard.co.uk/news/uk/facebook-giphy-b954378.html [accessed 17 November 2021]
630 Written evidence from NSPCC (OSB0109)
631 Q 62; see also: Antisemitism Policy Trust (OSB0005); APPG Coalition (OSB0202); Center for Countering Digital Hate (OSB0009); Refuge (OSB0084); NSPCC (OSB0109)
632 Q 181
633 Q 212
634 Written evidence from Meta (Facebook) (OSB0224)
635 Q 227
636 Q 241

364. The case against senior management liability was presented by the Open Rights Group, who argued that it would dissuade people from taking up jobs where they might be held criminally liable, that the offence targets a small handful of "specific and high-profile individuals, all of whom are American", that it will "create a culture of fear which results in a phenomenon known as ''collateral censorship'" where "vast swathes" of "perfectly harmless" content is taken down, and that it "sets a very poor global example".[637] The Open Rights Group argued that personal criminal liability for senior managers and company directors will "provide inspiration to authoritarian nations who look up to the UK as an example to follow: if the UK arrests company employees for the political speech carried on their platforms, why shouldn't they?"[638] Any enforcement process must be compatible with due process commitments under both the Human Rights Act and in common law.

365. On the other hand, Ms Pelham told us "the key to securing good regulation is the provisions on personal responsibility for senior managers in the social media companies. I have been in a position myself where I was personally responsible and, my goodness, it focuses your mind."[639]

366. The Government confirmed that Ofcom can only take action under the draft Bill against senior managers on failures to supply information. They told us that they have "targeted sanctions in this area as it is vital that Ofcom gets the information it needs to regulate the sector." They expect criminal sanctions will "instil strong engagement and cooperation with the regime among tech executives, and are satisfied that Ofcom's suite of enforcement powers will push strong compliance across the board."[640]

367. ***The Bill should require that companies' risk assessments be reported at Board level, to ensure that senior management know and can be held accountable for the risks present on the service, and the actions being taken to mitigate those risks.***

368. ***We recommend that a senior manager at board level or reporting to the board should be designated the "Safety Controller" and made liable for a new offence: the failure to comply with their obligations as regulated service providers when there is clear evidence of repeated and systemic failings that result in a significant risk of serious harm to users. We believe that this would be a proportionate last resort for the Regulator. Like any offence, it should only be initiated and provable at the end of an exhaustive legal process.***

369. ***The Committee welcomes the Secretary of State's commitment to introduce criminal liability within three to six months of Royal Assent and strongly recommends that criminal sanctions for failures to comply with information notices are introduced within three months of Royal Assent.***

Secretary of State powers

370. When she appeared before the Committee, the Secretary of State described her powers under the draft Bill as "novel".[641] Reset described them as "unprecedented, not only in the UK but also as compared to other online safety regulations" and said "they

637 Written evidence from Open Rights Group (OSB0118). See also written evidence from Internet Association (OSB0132) and Brother Watch (OSB0136)

638 Written evidence from Open Rights Group (OSB0118). See also written evidence from Brother Watch (OSB0136)

639 Q 59

640 Written evidence from Department of Digital, Culture, Media & Sport (OSB0243); Q 19

641 Q 279

undermine the independence of the UK's regime and cause unnecessary uncertainty for companies in scope".[642]

Box 3: Criticism of the powers of the Secretary of State

- Carnegie UK Trust states that the draft Bill "allows the Secretary of State to interfere with Ofcom's independence on content matters in four [principal] areas":
- The draft Bill "gives the Secretary of State relatively unconstrained powers to:
  - Set strategic priorities which OFCOM must take into account (109 and 57)
  - Set priority content in relation to each of the safety duties (41 and 47)
  - Direct OFCOM to make amendments to their codes to reflect Government policy (33)
- Give guidance to OFCOM on the exercise of their functions and powers (113)."

Source: https://www.carnegieuktrust.org.uk/blog-posts/secretary-of-states-powers-and-the-draft-online-safety-bill/

371. Clause 33.1 of the draft Bill empowers the Secretary of State to direct Ofcom to modify a code of practice submitted under section 32(1) where the Secretary of State believes that modifications are required (a) to ensure that the code of practice reflects government policy or (b) in respect of CSEA and/or terrorism content, for reasons of national security or public safety.

372. We heard from many witnesses who were concerned that the proposed powers of the Secretary of State to modify a code of practice so that it reflects government policy may undermine Ofcom's independence.[643] The IWF summarised the issue: "The possibility of too much central government constraint on Ofcom could undermine Ofcom's independence as a regulator and its ability to draft, implement and enforce mandatory Codes of Practice in a politically neutral way."[644] Prof Wilson told the Committee that it would be "better in the long run to grant Ofcom more independence and authority than the Bill does, because that will give the exercise of its regulatory powers more legitimacy."[645]

373. There is a case for retaining the Secretary of State's power to direct Ofcom in matters relating to CSEA and/or terrorism as far as they pertain to national security and public safety. As Ofcom put it: "there will clearly be some issues where the Government has access to expertise of information that the regulator does not, such as national security."[646] However, these powers should not be exercised without oversight or scrutiny. We note the Secretary of State's power to direct Ofcom to amend codes of practice so that they reflect

642 Written evidence from Reset (OSB0203). See also written evidence from Coadec (OSB0029).

643 See, for example, written evidence from: Carnegie UK (OSB0095); Professor Damian Tambini (Distinguished Policy Fellow and Associate Professor at London School of Economics and Political Science) (OSB0066); Barbora Bukovská (Senior Director, Law and Policy, Article 19) (Q138); TalkTalk (OSB0200); LSE Department of Media and Communications (OSB0001); Snap Inc. (OSB0012); Full Fact (OSB0056); ISPA (The Internet Service Provider Association) (OSB0059); Dr Martin Moore (Senior Lecturer at King's College London) (OSB0063); Written evidence from LSE, Department of Media & Communications—Freedom of Expression Roundtable (OSB0247)

644 Written evidence from Internet Watch Foundation (IWF) (OSB0110)

645 Q 138

646 Written evidence from Ofcom (OSB0021)

government policy is likely to be incompatible with best practice in other regulatory fields, for example the Council of Europe's Regulatory Best Practice Code.[647]

374. Clause 113 states that the Secretary of State may give guidance to Ofcom about the exercise of their functions under this Act; under section 1(3) of the Communications Act to carry out research in connection with online safety matters or to arrange for others to carry out research; and about the exercise of their media literacy functions under section 11 of the Communications Act. Dr Damian Tambini, Distinguished Policy Fellow and Associate Professor at LSE, described this power as "closer to authoritarian than to liberal democratic standards even with the safeguards".[648]

375. The Government told the Committee that: "the Secretary of State's powers under Clause 109 (to publish a statement of strategic priorities in relation to online safety matters) [cater] for long term changes in the digital and regulatory landscape" and that "a similar power already exists (under section 2A of the Communications Act 2003) for telecommunications, the management of the radio spectrum, and postal services".[649] Furthermore, "it is not the Government's intention that such a statement will be in place, or be needed, at the outset of the regime".[650] They said "these powers are part of the overall approach of balancing the need for regulatory independence with appropriate roles for parliament and government."[651]

376. ***The power for the Secretary of State to exempt services from regulation should be clarified to ensure that it does not apply to individual services.***

377. ***The powers for the Secretary of State to a) modify Codes of Practice to reflect Government policy and b) give guidance to Ofcom give too much power to interfere in Ofcom's independence and should be removed.***

378. ***Exercise of the Secretary of State's powers in respect of national security and public safety in respect of terrorism and child sexual exploitation and abuse content should be subject to review by the Joint Committee we propose later in this report.***

Media Literacy

Minimum standards for media literacy initiatives

379. The draft Bill places a duty on Ofcom to improve the media literacy of the public, building on the duty given to Ofcom in the Communications Act 2003 to promote media literacy. This duty largely involves improving awareness around how technology works and how to protect oneself online. We recognise that improved media literacy plays an important role in keeping people safe online, and as such welcome the publication of

647 See Council of Europe: Committee of Ministers, 'Recommendation Rec (2000)23 of the Committee of Ministers to member states on the independence and functions of regulatory authorities for the broadcasting sector': https://rm.coe.int/16804e0322.See [accessed 1 December 2021] which states that the rules governing regulatory authorities for the broadcasting sector are a key element of their independence and should be defined to protect them against any interference in particular by political forces or economic interests. Specific rules should be avoided which place regulatory authorities under the influence of political power.

648 Written evidence from Dr Damian Tambini (Distinguished Policy Fellow and Associate Professor at London School of Economics and Political Science) (OSB0066)

649 Written evidence from Department of Digital, Culture, Media & Sport (OSB0243); Q 16

650 Written evidence from Department of Digital, Culture, Media & Sport (OSB0243); Q 16

651 Written evidence from Department of Digital, Culture, Media & Sport (OSB0243); Q 16

the Government's Online Media Literacy Strategy, which is designed to complement the media literacy duties in the Online Safety Bill by exploring how, in practice, the duty can be met.[652] We also welcome that the draft Bill expands on the Communications Act 2003 to give greater detail on Ofcom's duties in relation to media literacy.

380. We heard from 5Rights that, under the draft Bill, Ofcom does not have to set minimum standards for what an initiative aimed at improving media literacy must include.[653] The content is left to the initiative provider. The Government's Online Media Literacy Strategy encourages a landscape where initiative providers can be anyone from a civil society organisation, to a news provider, to a service provider of an online platform.[654] We heard the concern that leaving these organisations to produce media literacy initiatives without oversight and guidance from Ofcom could allow them to distribute an "educational" resource that is biased, self-serving or factually incorrect. 5Rights gave the example of Google and Facebook, who both offer educational resources to schools around the world: "but teach children to accept certain service design elements as 'unavoidable' risks when in fact they could and should be tackled at a design level by those very same companies."[655] Mr Steyer told us that: "The idea that the industry will do high-quality media literacy or digital literacy and citizenship is crazy."[656]

381. **If the Government wishes to improve the UK's media literacy to reduce online harms, there must be provisions in the Bill to ensure media literacy initiatives are of a high standard. The Bill should empower Ofcom to set minimum standards for media literacy initiatives that both guide providers and ensure the information they are disseminating aligns with the goal of reducing online harm.**

382. ***We recommend that Ofcom is made responsible for setting minimum standards for media literacy initiatives. Clause 103 (4) should be amended to include "(d) about minimum standards that media literacy initiatives must meet."***

Ofcom's duty to improve media literacy

383. Under the draft Bill, Ofcom alone is given a duty to improve the media literacy of members of the public, though this can be undertaken through organisations other than themselves. The LSE questioned whether, given that the: "scope of the regulator's role and power … [is] to determine business practice", it is: "ideally placed to take responsibility for public education in relation to media literacy." [657] We received evidence that organisations other than Ofcom that play an important role in media literacy should be given a duty to improve the media literacy of certain groups. For example, the APPG Coalition suggested that: "given the focus on children in the draft Bill, there is surprisingly little insight into the role of teachers, Ofsted, and the Department of Education in developing and delivering a media literacy programme in schools."[658] Mr Steyer made a similar observation and

652 Department for Culture, Media and Sport, *Online Media Literacy Strategy*, (July 2021): https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1004233/DCMS_Media_Literacy_Report_Roll_Out_Accessible_PDF.pdf [accessed 15 November 2021]

653 Written evidence from 5Rights Foundation (OSB0096)

654 Written evidence from Department for Culture, Media and Sport, *Online Media Literacy Strategy*, (July 2021): p 5: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1004233/DCMS_Media_Literacy_Report_Roll_Out_Accessible_PDF.pdf [accessed 15 November 2021]

655 Written evidence from 5Rights Foundation (OSB0096)

656 Q 152

657 Written evidence from LSE Department of Media and Communications (OSB0001)

658 Written evidence from APPG Coalition (OSB0202)

thought that media literacy training should "[belong] in the Education Department."[659] The House of Lords Communications and Digital Committee has recommended that Ofcom should be a co-ordinating body, bringing together the work of Government, civil society, the private sector and academia, and has set out detailed recommendations on what a cross-government digital literacy programme might look like.[660]

384. We heard that service providers might also play a useful role in improving media literacy, given that they have direct access to and engagement with people who use their services. Carnegie UK Trust argued that media literacy should be built into risk assessments as a mitigation measure, which would compel service providers to ensure it is delivered to users.[661]

385. ***We recommend that the Bill reflects that media literacy should be subject to a "whole of Government" approach, involving current and future initiatives of the Department of Education in relation to the school curriculum as well as Ofcom and service providers. We have heard throughout this inquiry about the real dangers that some online content and activity poses to children. Ofsted already assesses how schools manage online safety as part of their safeguarding policies. We recommend that Ofsted, in conjunction with Ofcom, update the school inspection framework to extend the safeguarding duties of schools to include making reasonable efforts to educate children to be safe online***

386. ***Ofcom should require that media literacy is built into risk assessments as a mitigation measure and require service providers to provide evidence of taking this mitigation measure where relevant.***

Media literacy and a focus on individual rather than societal harms

387. The draft Bill gives Ofcom the duty to improve the media literacy of "members of the public". This is a change in wording from the Communications Act 2003, where the duty was to improve "public" awareness of the media. This seems to reflect the draft Bill's focus on individual rather than societal harms. For example, the definition of media literacy involves an understanding of how material is published and accurate it is, how personal information may be protected, and how someone might control what material they receive.[662] Prof Edwards said: "If media literacy is deployed only as a mode of self-protection from exploitation or harm, its potential for supporting our deliberative and democratic capacities could be severely weakened."[663] Glitch told us that media literacy needed to involve "digital citizenship", which "is respecting and championing the human rights of all individuals online" to reduce cases of online abuse.[664]

388. ***We recommend that Clause 103(11) is amended to state that Ofcom's media literacy duties relate to "the public" rather than "members of the public", and that the definition of media literacy is updated to incorporate learning about being a good digital citizen***

659 Q 152

660 Communications and Digital Committee, *Free for all? Freedom of expression in the digital age* (1st Report, Session 2021–22, HL Paper 54), para 293–296; Communications and Digital Committee, *Breaking News? The Future of UK Journalism* (1st Report, Session 2019–21, HL Paper 176), para 87–89

661 Written evidence from Carnegie UK (OSB0095)

662 See Draft Online Safety Bill, CP 405, May 2021, Clause 103

663 Lee Edwards, *'Media literacy in the Online Safety Bill: Sacrificing citizenship for resilience?':* https://blogs.lse.ac.uk/medialse/2021/11/09/media-literacy-in-the-online-safety-bill-sacrificing-citizenship-for-resilience/ [accessed 16 November 2021]

664 Written evidence from Glitch (OSB0097)

and about platform design, data collection and the business models and operation of digital services more broadly.

Use of technology warning notices

389. The draft Bill does not require service providers to use technology to identify and remove CSEA or terrorism content. However, under Clause 63, Ofcom can issue a use of technology warning notice if they have reasonable grounds to believe that a service provider is failing to comply with their safety duties relating to CSEA and/or terrorism content. The purpose of the warning notice is to alert a service provider that Ofcom is considering requiring it to use the technology specified in the notice to identify and remove terrorist content on public channels and/or CSEA content on private and/or public channels.

390. The Bill allows Ofcom to compel a service to use technology to detect CSEA and terrorism content on private and public channels and CSEA content on private communication channels and to "swiftly take down that content" (64(4)(b)). Some children's advocacy groups have welcomed this clause.[665] We heard concerns from others that "these steps would significantly undermine individual privacy and be incompatible with end-to-end encrypted services."[666] Facebook pointed out that the safeguards around these provisions were limited when compared to the powers established in other regimes: there is no judicial oversight, ability to appeal, or "explicit requirement to consider the privacy impact of any use of technology notice, including in the public interest in the integrity and security of the underlying services."[667]

391. The ability to require the use of automated moderation technology has received considerable criticism, not least the because of the inability of automated technology to understand images in context.[668] CSEA content is always illegal, but images or videos used by extremists in one context may be used for educational, journalistic or other legitimate purposes elsewhere.[669] There were also concerns about the inability of mandated technology to keep up with technological advances (e.g. livestreaming), and the risk that it might "lock providers into using tools that have been 'gamed' by bad actors".[670]

392. Another issue arising in this area is how Ofcom can gather enough evidence to justify mandating the use of technology. There are challenges in gathering evidence from private channels, from end-to-end encrypted channels, and if a service is not already using CSEA detection technology or is using it ineffectively. Short of responding to a user report, there is currently no way for a service provider to detect CSEA content on an end-to-end encrypted channel without compromising encryption.[671] Ofcom stated:

> "the bar for Ofcom to be able to require the use of these technologies should be high. But if we are given these powers, we will need to be able to use them effectively. In this regard, we need to avoid a catch-22 whereby it is only through

665 Written evidence from: the Office of the Children's Commissioner (OSB0019), p 9; NSPCC (OSB0109) p 4
666 Written evidence from Facebook (OSB0147), p 23; See also Tech Against Terrorism (OSB0052), p 66
667 Written evidence from: Facebook (OSB0147); Microsoft (OSB0076)
668 Written evidence from: Ms. Daphne Keller (Director, Program on Platform Regulation at Stanford Cyber Policy Center) (OSB0057); See also British & Irish Law, Education & Technology Association (OSB0073), para 6.3.2
669 Written evidence from Ms. Daphne Keller (Director, Program on Platform Regulation at Stanford Cyber Policy Center) (OSB0057).
670 Written evidence from: International Justice Mission (OSB0025); Google (OSB0175)
671 Written evidence from: Internet Watch Foundation (IWF) (OSB0110); Facebook (OSB0147)

the deployment of these technologies that we are able to generate a threshold of evidence that justifies our requiring their use."[672]

393. In the current drafting, Ofcom may issue a use of technology warning notice based on evidence demonstrating "the prevalence" and "the persistent prevalence" of terrorism and/or CSEA content on a service. This evidence could come from independent investigations, from news reports, civil society, whistle-blowing, or academic studies. The Children's Charities Coalition on Internet Safety called this wording "ambiguous"; International Justice Mission said that "any amount of CSEA content is unacceptable" and if "a regulated service is not actively trying to detect and prevent CSEA, there should still be consequences if this Bill is to truly hold services accountable".[673] They question whether only "prevalent" and "persistent" CSEA content should mark the threshold for triggering enforcement powers, and suggest instead that there should be multiple thresholds for triggering Ofcom's enforcement powers.[674]

394. ***The highest risk services, as assessed by Ofcom, should have to report quarterly data to Ofcom on the results of the tools, rules, and systems they have deployed to prevent and remove child sexual exploitation and abuse content (e.g. number and rates of illegal images blocked at upload stage, number and rates of abusive livestreams terminated, number and rates of first- and second- generation images and videos detected and removed).***

395. ***Ofcom should have the power to request research and independent evaluation into services where it believes the risk factors for child sexual exploitation and abuse are high.***

396. ***Ofcom should move towards a risk factors approach to the regulation of child sexual exploitation and abuse material. It should be able to issue a Use of Technology notice if it believes that there is a serious risk of harm from child sexual exploitation and abuse or terrorism content and that not enough is being done by a service to mitigate those risks. The Bill should be amended to clarify that Ofcom is able to consider a wider range of risk factors when deciding whether to issue a Use of Technology notice or take enforcement action. Risk factors should include:***

a) ***The prevalence or the persistent prevalence of child sexual exploitation and abuse material on a service, or distributed by a service;***

b) ***A service's failure to provide and maintain adequate tools, rules, and systems to proactively prevent the spread of child sexual exploitation and abuse content, and to provide information on those tools, rules, and systems to Ofcom when requested;***

c) ***A service's failure to provide adequate data to Ofcom on the results of those tools, rules, and systems (e.g., number and rates of illegal images blocked at upload stage, number and rates of abusive livestreams terminated, number and rates of first- and second- generation images and videos detected and removed);***

d) ***The nature of a service and its functionalities;***

672 Written evidence from: Ofcom (OSB0021); see also NSPCC (OSB0109)
673 Written evidence from International Justice Mission (OSB0025).
674 Written evidence from: Mr John Carr (Secretary at Children's Charities' Coalition on Internet Safety) (OSB0167); International Justice Mission (OSB0025)

e) ***The user base of a service;***

f) ***The risk of harm to UK individuals (and the severity of that harm) if the relevant technology is not used by the service;***

g) ***The degree of interference posed by the use of the relevant technology with users' rights to freedom of expression and privacy; and***

h) ***The safety by design mechanisms that have been implemented.***

9 Transparency and oversight

397. We concluded in Chapter 2 that many service providers' current transparency measures are not sufficient for users or researchers, who feel that service providers' systems and decision-making processes are akin to a black box.[675] Service provider transparency is also crucial for Ofcom to effectively fulfil their function as a regulator, as discussed in Chapter 8. At present there is no requirement on providers to produce transparency reports, leading to greatly varying levels of transparency, access and understanding.

Transparency for users

398. Service providers can be inconsistent in enforcing their terms and conditions, handling complaints, and taking enforcement decisions. We heard that they are often not transparent about these decisions:

> "You could have a particular phrase or word used in one context and it is reported and deleted, but in another context, or in a slightly different post, tweet or whatever you want to say, it is allowed under the terms of service." [676]

399. In some instances, inconsistency may be intentional. Documents released by the Wall Street Journal, showed that Facebook had a category of "whitelisted" users under a program called XCheck ("cross check"). A Facebook internal document explained that this meant: "for a select few members of our community [those whitelisted] we are not enforcing our policies and standards".[677] Unlike the rest of our community, these people can violate our standards without any consequences."[678] Twitter has a similar policy which exempts some users from being subject to their typical moderation processes.[679] Twitter said that this policy is intended to preserve content that is in the public interest,[680] whilst Facebook have explained that their intention was "to create an additional step so we can accurately enforce policies on content that could require more understanding".[681]

400. Inconsistencies also arise in moderation activity, some of which have been attributed to algorithmic or human biases resulting in over-moderation of legitimate content or obstacles to those seeking redress. For example, the removal of Palestinian content during

675 Q 136; Q 146; Written evidence from: Ada Lovelace Institute (OSB0101); ITV (OSB0204); Q 72

676 Q 53

677 This was 5.3 million members according to the Wall Street Journal reports (the Facebook Files). 'Facebook Says Its Rules Apply to All. Company Documents Reveal a Secret Elite That's Exempt' *Washington Post (13 September 2021):* https://www.wsj.com/articles/facebook-files-xcheck-zuckerberg-elite-rules-11631541353 [accessed 8 December 2021]

678 'Facebook Says Its Rules Apply to All. Company Documents Reveal a Secret Elite That's Exempt' *Washington Post (13 September 2021):* https://www.wsj.com/articles/facebook-files-xcheck-zuckerberg-elite-rules-11631541353 [accessed 1 December 2021]; Written evidence from Glitch (OSB0097)

679 Twitter, 'Defining public interest on Twitter': https://blog.twitter.com/en_us/topics/company/2019/publicinterest [accessed 1 December 2021]

680 Twitter, 'Defining public interest on Twitter': https://blog.twitter.com/en_us/topics/company/2019/publicinterest [accessed 1 December 2021]

681 'Facebook oversight board to review system that exempts elite users' *The Guardian* (22 September 2021): https://www.theguardian.com/technology/2021/sep/21/facebook-xcheck-system-oversight-board-review [accessed 1 December 2021]

conflicts in Israel and Palestine in May 2021,[682] Arabic-language content being erroneously removed from social media sites,[683] and the suppression of LGBTQ+ content on YouTube and other websites.[684]

401. Lack of transparency of service providers also means that people do not have insight into the prevalence and nature of activity that creates a risk of harm on the services that they use. Ms Haugen told us:

> "Facebook's own reports say that it is not just that Instagram is dangerous for teenagers; it is actually more dangerous than other forms of social media." [685]

Until Ms Haugen shared this information with the Securities and Exchange Commission and it was reported in the Wall Street Journal and other media outlets, it was not available to the public. We heard that people should be able to make an informed choice about the services that they are using by having an insight into the prevalence and nature of activity that creates a risk of harm on those services, services' terms and conditions, and how those terms and conditions are enforced.[686]

402. We heard from DMG Media that users often don't understand bias in search algorithms, "and imagine that when they search for news on politics, health, business, or any number of other topics, Google's emphasis on relevance and expertise means the content they are shown has been picked because it gives the most reliable and useful information."[687] They describe how these algorithms work as "the company's most closely-guarded secret", that these are likely influenced by commercial interest or bias, and have real implications for media plurality.[688] Sky told us that it was important the provisions protecting journalism were clear, "to ensure a plurality of views is upheld under the regime"[689], a point also made by the NUJ[690]. Transparency will be vital in ensuring that there is no detriment to media plurality from the application of the safety duties.

Provisions on transparency in the draft Bill

403. The draft Bill primarily aims to improve transparency by requiring service providers to produce annual transparency reports for each of their services, with the information included in those transparency reports to be determined by Ofcom in a notice given to the

682 BBC News, 'Israel-Palestinian Facebook posts needed 'bias' review': https://www.bbc.co.uk/news/technology-58558982 [accessed 18 November 2021]; Wired, 'Facebook's censorship-by-algorithm silenced Palestinian voices. Can its biases ever be fixed?': https://wired.me/business/big-tech/facebook-content-moderation-palestine/ [accessed 18 November 2021]; Oversight Board, 'Oversight Board overturns original Facebook decision. Case 2021–009-FB-UA':https://oversightboard.com/news/389395596088473-oversight-board-overturns-original-facebook-decision-case-2021–009-fb-ua/ [accessed 18 November 2021]

683 Project on Middle East Political Science, 'Digital Orientalism: #SaveSheikhJarrah and Arabic content': https://pomeps.org/digital-orientalism-savesheikhjarrah-and-arabic-content-moderation [accessed 18 November 2021]

684 Transthetics, 'YouTube's moderation process is failing the LBGT community. Can we fix this?': https://transthetics.com/YouTubes-moderation-process-is-failing-the-lgbt-community/; [accessed 18 November 2021] Talking Influence, 'LGBTQ+ Creators' Law Suit Against YouTube's Alleged Algorithm Discrimination Sits With Judge': https://talkinginfluence.com/2020/06/04/lgbtq-creators-lawsuit-YouTube-discrimination/ [accessed 18 November 2021]; Written evidence from LGBT Foundation (OSB0191)

685 Q 166

686 Q 68; Q 88

687 Written evidence from DMG Media (OSB0133)

688 Written evidence from DMG Media (OSB0133)

689 Written evidence from Sky (OSB0165)

690 Written evidence from The National Union of Journalists (NUJ) (OSB0166)

provider.[691] The draft Bill also contains other mechanisms that may enhance transparency, giving Ofcom duties to prepare a report about researchers' access to information[692] and to make arrangements for research about people's experiences of regulated services.[693]

Box 4: Summary of information which Ofcom can require from service providers in annual transparency reports

- the incidences of illegal and harmful content, and how many users have encountered such content
- how illegal and harmful content is disseminated on the service
- how terms of service or policies and procedures are applied
- the systems and processes for users to report illegal content, harmful content, or other content which breaches the terms of service or policies and procedures
- the systems and processes used to deal with illegal and harmful content, take it down, or prevent it being encountered in or via search results
- functionalities to help users manage risks relating to harmful content
- steps which a provider is taking to fulfil their various duties
- how the provider cooperates with government, regulatory, or other public sector bodies in the UK
- the systems and processes that the provider uses to assess the risk of harm to individuals from the presence of illegal content or harmful content
- the systems and processes a provider has in place to direct users to information about how they can protect themselves from harm in relation to illegal and harmful content
- the steps a provider is taking to provide a higher standard of protection for children than for adults
- the steps a provider is taking to improve media literacy, and evaluate the effectiveness of these steps
- any other steps the provider is taking relating to online safety matters

Source: Draft Online Safety Bill, Part 2, Chapter 6, Subsection 49(4)

404. Ofcom welcomed the transparency measures in the draft Bill describing them as a "step change" where service providers could be "truly accountable for the first time", resulting in "a significant improvement in transparency and accountability for internet users, the public and Parliament".[694]

691 Draft Online Safety Bill, CP 405, May 2021, Part 3, Chapter 1, Clause 49
692 Draft Online Safety Bill, CP 405, May 2021, Part 4, Chapter 7, Clause 101
693 Draft Online Safety Bill, CP 405, May 2021, Part 4, Chapter 7, Clause 99
694 Written evidence from Ofcom (OSB0021)

Current problems underlying transparency reporting

405. Many service providers currently produce transparency reports, but we heard that these are often not informative due to some service providers' choice of metrics.[695] For example, we were told by Mr Ahmed about "missing statistics",[696] where some service providers give self-selected statistics that do not transparently answer the question that is being asked.[697]

406. We asked Facebook how effective their algorithms were in detecting hate speech. Ms Davis told us that she was aware that some Facebook engineers had said that their AI removes only 3–5 per cent of hate speech, but that she was: "also aware that we have put out a transparency report that indicates that the prevalence of hate speech on our platform has been reduced to 0.05 per cent." She told us that they had "submitted the methodology that we used for that [metric]to an independent audit to verify[it]."[698] Ms Davis was unable to give us the information we requested.

407. Some providers use absolute values in their transparency reports, such as the number of takedowns per quarter. These metrics have limited value and do not give contextual information which could be important to understand them fully:

> " … we know that the volume of reports sent to us is not a useful metric, as many reports are about content which is not harmful …
>
> Another figure that is regularly discussed is the amount of content removed by a platform, but this number too has limited value in isolation: if it goes down, is that because the platform became worse at removing harmful content, or because less harmful content was posted in the first place?"[699]

408. Proportional metrics, such as the percentage of all content which violates a service's policy, also do not necessarily give an accurate picture of the scale of harm on a service. Google, for example, told us that that "removed videos represent a fraction of a percent of total views on YouTube" and that they "work continuously to shrink this even further through improved detection and enforcement".[700] Whilst removed videos may represent a very small proportion of the videos on YouTube, this could still mean a large number of hours spent engaging with policy-violating content that creates a risk of harm, with YouTube reporting in 2019 that they had reached 1 billion hours of viewing time a day globally.[701]

409. Where service providers use different metrics to report the nature and prevalence of activity that creates a risk of harm on their services, it is difficult to compare them. This prevents people from making an informed choice about which services they use. If people can compare the risks of harm on different services, this could encourage the

695 Q 107
696 Q 14
697 Written evidence from Ada Lovelace Institute (OSB0101)
698 Q 222
699 Written evidence from Facebook (OSB0147)
700 Written evidence from Google (OSB0175)
701 OBERLO, '10 YouTube stats every marketer should know in 2021 [Infographic]': https://www.oberlo.com/blog/YouTube-statistics [accessed 18 November 2021]

development of a competitive marketplace where successfully mitigating the risk of harm attracts users and becomes a competitive advantage.[702]

410. ***We recommend that Ofcom specify that transparency reports produced by service providers should be published in full in a publicly accessible place. Transparency reports should be written clearly and accessibly so that users and prospective users of the service can understand them, including children (where they are allowed to use the service) and disabled people.***

411. ***We recommend that the Bill require transparency reporting on a regular, proportionate basis, with the aim of working towards standardised reporting as the regulatory regime matures. The Bill should require minimum standards of accuracy and transparency about how the report was arrived at and the methodology used in research. For providers of the highest risk services, the outcome of the annual audits recommended in paragraph 340 should be required to be included in the transparency report.***

412. ***We agree with the list of information that Ofcom can require as part of its transparency reporting powers and recommend that it should have the clear power to request any other information. We recommend that transparency reporting should aim to create a competitive marketplace in respect of safety, where people can reasonably compare, using robust and comparable information, performance of services as they operate for UK users. We suggest Ofcom also be able to require information be published in transparency reports including (but not limited to):***

a) ***Safety by design features;***

b) ***Most viewed/engaged with content by month;***

c) ***Most recommended content by month by age group and other demographic information (where that information is collected);***

d) ***Their terms and conditions;***

e) ***Proportion of users who are children;***

f) ***Proportion of anonymous users;***

g) ***Proportion of content breaching terms and conditions;***

h) ***Proportion of content breaching terms and conditions removed;***

i) ***Proportion of appeals against removal upheld;***

j) ***Proportion of appeals against removal, by both recognised news publishers and other users on the grounds of public interest, upheld; and***

k) ***Time taken to deal with reports.***

413. ***In addition to transparency reporting, Ofcom should be empowered to conduct its own independent research with the aim of informing the UK public about the comparative performance of services in respect of online safety.***

702 Q 61; Q 248

Access for independent researchers

414. The draft Bill requires Ofcom to prepare a report about researchers' access to information and to publish this within two years of the Bill being enacted into legislation.[703] This report must describe "how, and to what extent, persons carrying out independent research into online safety matters are currently able to obtain information from providers of regulated services to inform their research."[704]

415. Dr Moore told us that "without … research and external scrutiny, we will be unable to properly assess the extent of problematic content and behaviour on these platforms, or assess the harms committed."[705] This position was supported by a number of other witnesses, who told us that lack of transparency from service providers and limited access to information for independent researchers hinders much-needed scientific progress towards understanding the prevalence, impact, causes, and dynamics of online activity that creates a risk of harm.[706] This, in turn, hinders the ability to make policy decisions and dampens innovation, leaving us "working in the dark".[707]

416. Witnesses told us that greater transparency could allow for more scrutiny of service providers, and consequently, increased accountability.[708] Demos urged that the Bill give greater priority for independent researcher access to service providers' data about the service:

> "We would recommend that greater priority be given than is in the current Bill to facilitating independent researcher access to platform data, with appropriate privacy safeguards, so that platform action can be better scrutinised and [to] improve accountability for any failures to take meaningful measures to reduce risks of harm."[709]

417. Many other witnesses, as well as participants in our 3 November roundtable, called for the highest risk services to share data more openly with vetted researchers.[710] Reset called for the transparency powers in the Bill to "include a requirement for platforms to share relevant data with accredited researchers studying online harms/safety" as this would "give academia a much clearer picture of how harmful content is generated and promoted online and what impact it has on fundamental rights and the greater public good."[711] Reset points out that this would "align the Bill with the Digital Services Act" and redress the current transparency arrangements which operate "at the whim of platforms".[712]

418. Ms Edelson told us that service providers are currently developing technological solutions to activity that creates a risk of harm on their own without sharing information,

703 Draft Online Safety Bill, CP 405, May 2021, Part 4, Chapter 7, Clause 101
704 Draft Online Safety Bill, CP 405, May 2021, Clause 101(1)(a)
705 Written evidence from Dr Martin Moore (Senior Lecturer at King's College London) (OSB0063)
706 Carnegie UK (OSB0095); Q 99; Q213; Who Targets Me (OSB0086); Dr Amy Orben (College Research Fellow at Emmanuel College, University of Cambridge) (OSB0131)
707 Written evidence from Reset (OSB0138); Q 62
708 Written evidence from: Ofcom (OSB0021); Q271; Twitter (OSB0072); 5Rights Foundation (OSB0206); Written evidence from Ada Lovelace Institute (OSB0101); Q 63
709 Written evidence from Demos (OSB0159)
710 Written evidence from: Who Targets Me (OSB0086), point 6; Logically (OSB0094); Carnegie UK (OSB0095); Dr Amy Orben (College Research Fellow at Emmanuel College, University of Cambridge) (OSB0131); Catch 22 (OSB0195).
711 Written evidence from Reset (OSB0138), 20.
712 Written evidence from Reset (OSB0138), 20–21.

data, or knowledge. This lack of sharing hinders scientific progress on understanding and developing algorithms:

> "We just do not have enough data. Ideally, we would develop a taxonomy of the variety of harmful content that spreads online and there would be research saying, 'We have developed a classifier and it is X per cent effective at identifying self-harm content'. Someone else would come out with a better one. That is the normal process of scientific research, but we just do not have the data to do that … I do not want to say that it is useless to take the platforms' research without seeing the data that backed it, but it does not advance science about what is going on in these platform… If we just make public data on platforms available to researchers … We can go through the scientific process of understanding various areas of harmful content and how we can avoid promoting them."[713]

Ms Edelson is a member of the Ad Observatory project at New York University, which collected volunteers' Facebook data to study the targeting of users with political advertisements and misinformation on Facebook. Despite collecting data only from consenting volunteers, Facebook shut down the personal accounts and research tools of members of the Ad Observatory in August 2021 for breaching its privacy rules.[714]

419. We heard from Dr Amy Orben, College Research Fellow at Emmanuel College, University of Cambridge, that lack of access to data is "making it impossible for good quality and independent scientific studies to be completed on topics such as online harms, mental health, or misinformation."[715] We heard that researchers have been misled by data that has been shared with them.[716] Where researchers do have access to information, it lacks "detail and richness" that is important for researchers.[717]

420. Where data is available, service providers sometimes restrict access to it. In one case, an entire research programme was disrupted because independent researchers' access to data was revoked by the service provider.[718]

421. We heard there is evidence that social media usage can cause psychological harm to children, but that platforms prevent research in this area from being conducted or circulated. Professor Jonathan Haidt, Professor of Ethical Leadership at New York University Stern School of Business, told us that in 2013 and 2014: "Something happened that started sending girls in particular to hospitals, and the suicide rate greatly increased."[719] Haidt argued that the rise of social media at this time was the cause. Similarly, we heard about The Wall Street Journal's work that found Facebook conducted internal research that revealed young girls were psychologically harmed as a result of using Instagram.[720] Common Sense told us that this example highlighted how important it is that independent

713 QQ 98–99
714 Center for Cybersecurity, 'Facebook Disables Ad Observatory; Academicians and Journalists Fireback': https://cyber.nyu.edu/2021/08/21/facebook-disables-ad-observatory-academicians-and-journalists-fire-back/ [accessed 1 December 2021]; Q 213
715 Written evidence from Dr Amy Orben (College Research Fellow at Emmanuel College, University of Cambridge) (OSB0131); Q 185
716 Q 196
717 Q 63
718 Written evidence from: Reset (OSB0138); Ada Lovelace Institute (OSB0101)
719 Q 148 (Professor Jonathan Haidt)
720 'Facebook knows Instagram Is Toxic For Teen Girls, Company Documents Show', *The Wall Street Journal* (14 September 2021): https://www.wsj.com/articles/facebook-knows-instagram-is-toxic-for-teen-girls-company-documents-show-11631620739 [accessed 2 December 2021]

researchers have access to data from platforms, as internal research "tends to be biased and, if [platforms] do not like the results, they simply will not share it with the public."[721]

422. We heard in other evidence that researchers require further data to greater explore the link between social media usage and psychological harm in children. Dr Orben and Dr Andrew K. Przybylski, Director of Research at the Oxford Internet Institute, stressed to us that: "Industry are not sharing that data and the resources to support independent work have not been allocated. Without these steps being taken, online harms and thresholds of harm cannot have a scientific basis."[722]

423. Facebook told us that it wanted to share data with independent researchers but had unresolved concerns about protecting users' privacy:

> "One of the things that is a particular challenge in the area of research is how we can provide academics who are doing independent research with access to data really to study these things more deeply. We are currently working with some of the leading academic institutions to figure out what the right rules are to allow access to data in a privacy protective way. One thing that we are quite supportive of, in terms of some of the legislation that we are here to talk about today, is working with regulators to set some parameters around that research that would enable that research and would enable people to have trust in the research that is done with access to our data in a privacy-protected way."[723]

In written evidence, Facebook said it would welcome legislation that will address this issue, that it has "a long history of seeking to make privacy-protected data available to support research" and that it was supportive of a solution which accelerates independent researchers' access to information: "Ofcom and the ICO should begin work on their report into researchers' access to information immediately, and not wait until two years after the Bill has passed"[724] Twitter and Google were also supportive of independent researchers having access to data.[725]

424. **Independent researchers currently have limited access to the information needed to conduct research. This hinders progress in understanding online activity that creates a risk of harm, the way that services' systems work, and how services' systems could be improved to mitigate the risk of harm. It also limits the ability to scrutinise service providers and hold them accountable. This issue must be addressed urgently.**

425. **The transparency powers in the Bill are an important opportunity to encourage service providers to share relevant data with external researchers studying online safety and allied subjects.**

426. ***The draft Bill requires that Ofcom produce a report on access to data for independent researchers. We recommend work on this report starts as soon as possible. We recommend***

721 Q 148 (Jim Steyer)

722 Written evidence from Professor Andrew Przybylski (Associate Professor, Senior Research Fellow at University of Oxford) (OSB0193). See also witten evidence from Dr Amy Orben (College Research Fellow at Emmanuel College, University of Cambridge) (OSB0131)

723 Q 200; also oral evidence taken before the Democracy and Digital Technologies Committee, 17 March 2020 (Session 2019–20), Q 298-99 (Karim Palant)

724 Written evidence from Facebook (OSB0147)

725 Written evidence from Twitter (OSB0072); Q 229

that Ofcom be given the powers in the Bill to put into practice recommendations from that report.

427. ***Ofcom should have the power i) to audit or appoint a third-party to audit how services commission, surface, collate and use their research; ii) to request a) specific internal research from services; b) research on topics of interest to the Regulator.***

428. ***Ofcom should commission an independent annual assessment, conducted by skilled persons, of what information should be provided by each of the highest risk services to advance academic research.***

429. ***We recommend that the Bill should require service providers to conduct risk assessments of opening up data on online safety to independent researchers, with some pre-defined issues to comment on, including a) privacy; b) risk of harm to users; c) reputational risks (for the service provider) and; d) financial cost***

430. ***We recommend that Ofcom should require service providers to conduct an annual formal review of using privacy-protecting technologies and enable them to share sensitive datasets.***

Role and value of a Joint Committee on Digital Regulation

431. A Joint Committee to oversee online safety and digital regulation more broadly has been recommended by numerous parliamentary committees. [726] Secretary of State Ms Dorries and Mr Philp told us that they were supportive of the proposal of an ongoing Joint Committee of both Houses.[727]

432. The call for a Joint Committee on Digital Regulation was recently reiterated by the House of Lords Communications and Digital Committee in their in *Digital regulation: joined-up and accountable*.[728] In their report, they highlighted that regulators are increasingly being given "broad powers to address complex and evolving challenges", which brings risks, making sustained attention from Parliament imperative "to ensure both that regulators have the powers they need and … that regulators are using those powers appropriately and effectively.[729] They identified seven different permanent parliamentary committees with remits relating to digital regulation, but none with a remit to focus on digital regulation. Just as digital regulation "needs to be cross-sectoral, so too must be the process of holding regulators to account." A Joint Committee on Digital Regulation could "ensure coherence and draw on the full range of expertise in Parliament".[730]

726 Democracy and Digital Technologies Committee, *Digital Technology and the Resurrection of Trust* (Report of Session 2019–21, HL Paper 77), Recommendation 15; Communications and Digital Committee, *Free for all? Freedom of expression in the digital age* (1st Report, Session 2021–22, HL Paper 54), Recommendation 16; Communications Committee, *Regulating in a digital world* (2nd Report, Session 2017–19, HL Paper 299), Recommendation 33

727 QQ 275–276; Q 294

728 Communications and Digital Committee, *Digital regulation: joined-up and accountable* (3rd Report, Session 2021–22, HL Paper 126)

729 Communications and Digital Committee, *Digital regulation: joined-up and accountable* (3rd Report, Session 2021–22, HL Paper 126)

730 Communications and Digital Committee, *Digital regulation: joined-up and accountable* (3rd Report, Session 2021–22, HL Paper 126)

433. An ongoing Joint Committee would serve numerous critical functions:

a) Oversight and accountability of digital regulators in respect of the Bill: The draft Bill gives Ofcom a wide range of powers for enforcement[731] with some arguing that they are too great.[732] A Joint Committee would provide a greater level of democratic accountability for Ofcom and other digital regulators over how they are using their powers and this could be beneficial in alleviating the concerns raised by witnesses.[733]

b) Scrutiny of the Secretary of State in respect of the Bill: We have heard substantial concerns regarding the "exceptional"[734] powers of the Secretary of State.[735] Prof Wilson told us that: "The question we should always ask of legislation is, 'Would I like this in the hands of my political opponents?' because one day they will come to power."[736]A Joint Committee of both Houses with proportional representation from different political parties could serve a valuable function in scrutinising the Digital Regulation work of the Secretary of State and the way that they use their powers. For example, a Joint Committee could review the priority content that the Secretary of State designates under the Online Safety Bill.

c) Monitoring Ofcom's independence: Ofcom and numerous other witnesses have said that the powers given to the Secretary of State in the draft Bill may undermine Ofcom's independence and their ability to show clear and evidence-based decision-making.[737] Ms Denham told us that Ofcom's independence as a regulator is "critically important".[738] A Joint Committee could monitor the independence of Ofcom and make recommendations to safeguard it where necessary.

d) Look across the digital regulation landscape: Digital regulation is a complex and evolving landscape[739] and the internet is already regulated by multiple independent regulators.[740] Oversight of the digital regulation landscape by a Joint Committee of both Houses could support the ongoing development of regulation and legislation and assess regulatory coherence in this area.[741] The Joint Committee could also maintain an overview of international efforts in digital regulation.

731 Written evidence from Ofcom (OSB0021)

732 Written evidence from: Dr Martin Moore (Senior Lecturer at King's College London) (OSB0063); Virgin Media O2 (OSB0127)

733 Written evidence from: The Age Verification Providers Association (OSB0122); Dr Mikolaj Barczentewicz (Senior Lecturer in Law at University of Surrey) (OSB0152)

734 Carnegie UK, 'The draft Online Safety Bill gives too many powers to the Secretary of State over too many things': https://www.carnegieuktrust.org.uk/blog-posts/secretary-of-states-powers-and-the-draft-online-safety-bill/ [accessed 18 November 2021]

735 Q 138; Written evidence from Professor Damian Tambini (Distinguished Policy Fellow and Associate Professor at London School of Economics and Political Science) (OSB0066); QQ 70–72; Q 77

736 Q 138

737 Written evidence from Ofcom (OSB0021), Q 138, Written evidence from LSE Department of Media and Communications (OSB0001), Q 126, Q 266, Written evidence from: Snap Inc. (OSB0012); Vodafone UK (OSB0015); Global Partners Digital (OSB0194); Confederation of British Industry (CBI) (OSB0186)

738 Q 90

739 Department for Digital, Cultre, Media and Sport, *Digital Regulation: Driving Growth and Unlocking Innovation* (July 2021): https://www.gov.uk/government/publications/digital-regulation-driving-growth-and-unlocking-innovation/digital-regulation-driving-growth-and-unlocking-innovation [accessed 18 November 2021]

740 Q 86; The Digital Regulation Cooperation Forum, *Information about the Digital Regulation Cooperation Forum (DRCF), established to ensure greater cooperation on online regulatory matters* (March 2021): https://www.gov.uk/government/collections/the-digital-regulation-cooperation-forum [accessed 18 November 2021]

741 Communications and Digital Committee, *Free for all? Freedom of expression in the digital age* (1st Report, Session 2021–22, HL Paper 54)

e) Horizon scanning: Digital technologies are complex and rapidly evolving. Legislation will face challenges as new technologies and new risks of harm emerge, and we have heard concerns about how the Bill can handle these challenges.[742] A Joint Committee of both Houses could look to the future to identify newly emerging risks or technologies that represent a challenge to the regulatory and legislative landscape.

f) Generate solutions to current issues: Numerous issues raised throughout our inquiry are complex and as yet unresolved. A Joint Committee of both Houses could be instrumental in helping to generate solutions to ongoing policy issues such as how to accurately identify disinformation and misinformation online.

434. ***We agree with other Committees that it is imperative that digital regulation be subject to dedicated parliamentary oversight. To achieve this, we recommend a Joint Committee of both Houses to oversee digital regulation with five primary functions: scrutinising digital regulators and overseeing the regulatory landscape, including the Digital Regulation Cooperation Forum; scrutinising the Secretary of State's work into digital regulation; reviewing the codes of practice laid by Ofcom any legislation relevant to digital regulation (including secondary legislation under the Online Safety Act); considering any relevant new developments such as the creation of new technologies and the publication of independent research or whistleblower testimonies; and helping to generate solutions to ongoing issues in digital regulation.***

435. ***We fully support the recommendation of the House of Lords Communications and Digital Committee in their report on Digital Regulation that, as soon as possible, full Digital Regulation Cooperation Forum membership should be extended to statutory regulators with significant interests and expertise in the digital sphere, and that partial membership should be extended to non-statutory regulators and advisory bodies with subject specific knowledge to participate on issues particular to their remits.***

436. ***We recommend that, in addition to any other reports the Committee chooses to make, the Joint Committee produces an annual report with recommendations on what could or should change, looking towards future developments. We anticipate that the Joint Committee will want to look at the definition of disinformation and what more can be done to tackle it at an early stage.***

Protections for whistleblowers

437. Whistleblowers like Ms Haugen and Ms Zhang, who both gave evidence to us, have greatly helped to increase understanding of the systems and processes of large service providers and their services. They have set out the challenges that exist with creating and enforcing effective content moderation systems, understanding known harms and emerging threats to user safety. Ms Haugen has also set out the kind of research Facebook conducts on the impact of its services on the welfare of its users and the decisions it has made when safety concerns might conflict with overall engagement with the service.

438. The Public Interest Disclosure Act 1998 provides protection to whistleblowers who make disclosures to their employer or other relevant person of potential criminal offences, the endangering of health and safety of another person or people, or other forms of protected disclosure. Such protection includes against detriment at work and being penalised by

742 Q 66, Q 77, Q 126, Q 190, Q 244, Q 255; Written evidence from: NSPCC (OSB0109); Sky (OSB0165)

non-disclosure agreements. The "relevant persons" to whom such a disclosure can be made are set out in the Public Interest Disclosure (Prescribed Persons) Order 2014. The Order does not currently have a prescribed person relating to online safety.

439. ***We recommend that whistleblowers' disclosure of information to Ofcom and/or the Joint Committee on Digital Regulation, where that information provides clear evidence of non-compliance with the Online Safety Bill, is protected under UK law.***

10 Redress

Redress and reporting mechanisms for in-scope providers

440. As explored in Chapter 9, there is currently little transparency about decision-making or outcomes when users report issues to service providers.[743]

441. We heard compelling evidence from Prof McGlynn and others that this Bill "provides a valuable opportunity to strengthen individual protections against online violence and abuse" but that it currently falls short of what it might achieve in this area.[744] Online violence and abuse can take many forms, and individuals who have been abused often find that their options to gain redress from service providers or from the courts are limited.[745] We heard from Refuge that tech abuse, which is a form of domestic or intimate partner violence, can entail "hundreds of abusive messages … from perpetrators, often across multiple platforms", each of which has to be flagged to the service provider individually.[746] Survivors of tech abuse can wait weeks or months even to receive acknowledgement of their report from service providers.[747] Those who have been subjected to image-based sexual abuse can find themselves failed by "out-of-date, confusing, piecemeal" laws, which can lead police officers to use "informal resolutions" which fail to capture image-based sexual abuse as a sexual offence and therefore the available criminal justice response.[748]

442. Under the draft Bill, services in scope must operate a complaints process that "provides for appropriate action to be taken by the provider of the service" and which is "easy to access", "easy to use (including by children)" and "transparent" (15). Many large services already provide users with complaints procedures, alongside mechanisms to appeal takedown decisions.[749] As such, it seems the clause is designed to improve service providers' existing procedures, rather than to establish new ones. In that respect it differs, for example, from the EU Digital Services Act, which seeks to oblige digital marketplaces to set up complaint and redress mechanisms and out-of-court dispute settlement mechanisms.[750] Yet, the Bill does not currently seek to establish minimum quality standards for the operation of any of these processes.[751] Nor does it establish how Ofcom will assess how easy to access, easy to use, or transparent a service provider's reporting or complaints process is.

743 Written evidence from Ms. Daphne Keller (Director, Program on Platform Regulation at Stanford Cyber Policy Center) (OSB0057); Q 53

744 Written evidence from Professor Clare McGlynn (Professor of Law at Durham University) (OSB0014), p 7. See also written evidence from: Professional Players Federation (OSB0035), p 3; Centenary Action Group, Glitch, Antisemitism Policy Trust, Stonewall, Women's Aid, Compassion in Politics, End Violence Against Women Coalition, Imkaan, Inclusion London, The Traveller Movement (OSB0047) p 6; RSA (Royal Society for the Encouragement of Arts, Manufactures and Commerce) (OSB0070), p 6

745 Written evidence from Refuge (OSB0084), pp 8–9

746 Written evidence from Refuge (OSB0084), pp 8–9

747 Written evidence from Refuge (OSB0084), pp 8–9

748 Written evidence from Professor Clare McGlynn (Professor of Law at Durham University) (OSB0014), p 8

749 See, for example: Facebook, 'How do I appeal the removal of content on Facebook for copyright reasons?': https://en-gb.facebook.com/help/194353905193770 [accessed 1 December 2021]; Twitter, 'Appeal an account suspension or locked account': https://help.twitter.com/forms/general [accessed 1 December 2021]; Instagram, 'I don't think Instagram should have taken down my post': https://www.facebook.com/help/instagram/280908123309761 [accessed 1 December 2021]

750 Written evidence from Electrical Safety First (OSB0100), 6.6

751 Written evidence from 5Rights Foundation (OSB0096), point 4

443. ***The Bill should establish proportionate minimum standards for the highest risk providers' reports, complaints, and redress mechanisms as set out in a mandatory code of practice prepared by Ofcom.***

444. ***We recommend a requirement on the face of the Bill for Ofcom to set out: i) how they will assess the a) ease of use; b) accessibility and c) transparency of a service's complaints process for d) adults; e) children; and g) disabled people f) vulnerable adults; ii) what steps Ofcom will be able to take if it finds any of these processes wanting; and iii) how Ofcom will ensure that requirements to operate complaint, reporting and redress mechanisms are proportionate for smaller in-scope providers.***

445. ***Clause 15 (3)(c) should be amended so that it reads "is easy to access, including for disabled people and those with learning difficulties".***

446. ***Providers of the highest risk services should have to give quarterly statistics to Ofcom on:***

i) ***Number of user reports;***

ii) ***User reports broken down by the reason the report was made;***

iii) ***Number of actionable user reports;***

iv) ***Actionable user reports broken down by the reason the report was made;***

v) ***How long it took the service provider to respond to i) all user reports; ii) actionable user reports;***

vi) ***What response was made to actionable user reports;***

vii) ***Number of user complaints received;***

viii) ***Number of actionable user complaints;***

ix) ***How long it took the service provider to respond to i) all user complaints; ii) actionable user complaints;***

x) ***What response was made to actionable user complaints;***

xi) ***How many pieces of user content were taken down;***

xii) ***How many pieces of content that were taken down were later reinstated;***

xiii) ***The grounds on which content that was reinstated was reinstated;***

xiv) ***How long it took the service provider to reinstate a piece of content that was later reinstated.***

External redress for individuals

447. The Bill mandates that user-to-user services must provide opportunities for individuals to make complaints and provides service providers with the opportunity to appeal decisions made by Ofcom (Clause 104), Ofcom notices (Clause 105) and to make super-complaints (Clause 106).

448. The explanatory notes which accompanied the draft Bill explained that "a body representing the interests of UK users of regulated services, or members of the public can make a super-complaint to Ofcom about any feature of one of more regulated services, or the conduct of one or more providers of such services."[752] The super-complaints measure is useful for reporting multiple and widespread suspected breaches, for example the widespread bullying or abuse of a class or group. It also supports the transparency objective by enabling group complaints against powerful companies. However, a super-complaint may only be made if "the complaint is of particular importance" or if "the complaint relates to [or] impacts on a particularly large number of users of the service or members of the public (106(2)). As such, there is no right for an individual to seek external redress under the Bill as it is currently drafted.

449. We heard four principal arguments in favour of a new external appeals mechanism for individuals once internal routes have been exhausted. Firstly, that it would provide a source of redress to victims of online abuse of the kind described above once they have exhausted a service provider's internal complaints process.[753] Secondly, we heard arguments for the addition of an appeals mechanism on the grounds of consumer protection.[754] Thirdly, that it would empower users if they were consulted on mechanisms for redress, whilst addressing the "current imbalance between democratic 'people' power and the power of platforms".[755] Finally, we received submissions which argued there should be an external appeals process to prevent over-enforcement and to protect the freedom of expression of individuals who feel their content has been unfairly removed or demoted.[756]

450. **Our proposed external redress process would not replace service providers' internal processes or run concurrently to them, nor would it address individual complaints about individual pieces of content or interactions. Rather, for a victim of sustained and significant online harm, someone who has been banned from a service or who had their posts repeatedly and systematically removed, this new redress mechanism would give them an additional body to appeal those decisions after they had come to the end of a service provider's internal process.**

451. **In order for an external redress process to work, clear direction is needed in the Bill about Ofcom's responsibility to set quality standards for service provider's internal complaints procedures, and in relation to complaints about failures to meet those standards. We hope that the Government will consider our recommendations in this area, and that by improving the quality of service providers' internal complaints procedures, any system of external redress will be needed only rarely and for the most serious cases.**

452. Ofcom has said that if they were the body designated to take individual complaints "that could not just overwhelm us in volume but conflict a bit with the role that the Bill gives us, which is a strategic one looking at overall systems and processes" and that they

752 Explanatory Notes to the draft Online Safety Bill [Bill CP 405-EN]

753 Written evidence from Refuge (OSB0084), p 28.

754 Written evidence from Parent Zone (OSB0124), p 4. See Written evidence from Competition and Markets Authority (OSB0160), point 5, which argues that the Draft Bill "risks inadvertently setting a lower standard of consumer protection on platforms for economic and financial harms than that already envisaged by current law, and established by the CMA's enforcement work".

755 Written evidence from LSE Department of Media and Communications (OSB0001)

756 Written evidence from: RSA (Royal Society for the Encouragement of Arts, Manufactures and Commerce) (OSB0070), p 6; Demos (OSB0159), 48; Ms. Daphne Keller (Director, Program on Platform Regulation at Stanford Cyber Policy Center) (OSB0057).

would prefer individual complaints to be handled by a specially appointed Ombudsman.[757] The ICO were similarly cautious about making Ofcom responsible for handling individual complaints: "if individual complaints could come to a different organisation, that might be a way to go, and then Ofcom could learn from the experience of those individuals."[758]

453. Our hope is that by improving the quality of service providers' complaints procedures, the burden on any external redress process will be lessened, but a stronger internal process is no substitute for the rigour of independent oversight. Nevertheless, one of the primary challenges to establishing a redress mechanism for individuals is the high number of potential claimants. The ICO cautioned: "imagine the millions of complaints for take-down requests that might go to an organisation such as Ofcom."[759] There are over 48.5 million Facebook users in the UK, and around 28.8 million on Instagram.[760] However, we note that many external redress systems only apply once the internal process has been exhausted. For example, Ofcom only intervenes once a complainant has exhausted the BBC's internal process.[761] Furthermore, we note that the advent of the General Data Protection Regulation (GPDR) did not bring about an overwhelming surge of cases for breach of data protection law. This may be for a number of reasons including companies preferring to settle claims, that potential litigants are unaware of their rights, and/or because they are put off by the prospect of making a claim for a relatively low level of damages. We envisage the external complaints process as a last resort for users who have suffered serious harm on services. It is an important step towards greater transparency and clarity in service provider's moderation decisions which greatly outweighs the potential for misuse by users.

454. We note with interest South West Grid for Learning's Report Harmful Content, which offers online users an opportunity to report harmful online content, as well as an impartial dispute resolution service for users who have exhausted a service's internal complaint procedures.[762] We suggest that the Department look to Report Harmful Content as a potential model for what such an Ombudsman could look like.[763]

455. The Secretary of State was reluctant to commit to introducing an Ombudsman:

> "I know that [Dame] Melanie at Ofcom raised a point about an ombudsman, which is a slow and onerous process. We do not want to get into that. We want to get into making platforms behave responsibly as quickly as possible, under a legal framework, and that is what we are focused on."[764]

In a letter to the Committee, the Government stated that:

> "Although Ofcom will not investigate or arbitrate on individual complaints (owing to the likelihood of becoming overwhelmed by sheer volume), it will

757 Q 253
758 Q 89
759 Q 89
760 Statista, 'United Kingdom: Facebook users 2021, by age group': https://www.statista.com/statistics/1030055/facebook-users-united-kingdom [accessed 1 December 2021]; Statista, 'United Kingdom: monthly Instagram users 2018–2021': https://www.statista.com/statistics/1018494/instagram-users-united-kingdom [accessed 1 December 2021]
761 Ofcom, 'Complain about the BBC': https://ofcomforms.secure.force.com/formentry/SitesFormBBCIntroductory?complaintType=SitesFormBBCOnlineMaterial [accessed 1 December 2021]
762 Written evidence from SWGfL (OSB0054).
763 'Report Harmful Content': https://reportharmfulcontent.com/?lang=en [accessed 1 December 2021]
764 Q 284

be possible for individuals to submit complaints to Ofcom. Ofcom will use aggregate data from user complaints to inform its horizon scanning, research supervision and enforcement activity."[765]

456. **We support the Government's ambition to make service providers behave responsibly, and by agreeing our recommendations the requirements of the Bill will bring about better responses from service providers to user complaints. However, the fact remains that service providers' user complaints processes are often obscure, undemocratic, and without external safeguards to ensure that users are treated fairly and consistently. It is only through the introduction of an external redress mechanism that service providers can truly be held to account for their decisions as they impact individuals.**

457. ***The role of the Online Safety Ombudsman should be created to consider complaints about actions by higher risk service providers where either moderation or failure to address risks leads to significant, demonstrable harm (including to freedom of expression) and recourse to other routes of redress have not resulted in a resolution. The right to complain to this Ombudsman should be limited to users to those i) who have exhausted the internal complaints process with the service provider against which they are making their complaint and ii) who have either a) suffered serious or sustained harm on the service or b) had their content repeatedly taken down. There should be an option in the Bill to extend the remit of the Ombudsman to lower risk providers. In addition to handling these complaints the Ombudsman would as part of its role i) identify issues in individual companies and make recommendations to improve their complaint handling and ii) identify systemic industry wide issues and make recommendations on regulatory action needed to remedy them. The Ombudsman should have a duty to gather data and information and report it to Ofcom. It should be an "eligible entity" to make super-complaints.***

Liability in the civil courts

458. A duty of care in negligence law implies a right to take anyone who fails in that duty of care to court and seek compensation. As we discussed in paragraph 54 above, the duties of care in the Bill do not create individual liability between the user and the service provider which would allow users to sue for negligence.[766] We asked our witnesses how users might be able to seek redress through the courts. Mr Perrin told us Carnegie UK Trust had rejected the idea of a statutory tort[767] "because we felt it would not lead to a good regulatory outcome; it favours people who have resources to sue, and the courts do not work terribly quickly and are rather overloaded at the moment."[768] Dr Harbinja told us that the approach of the courts to negligence was not in legal terms a good fit with the duties in the Bill, and ran the risk of unintended consequences.[769] Prof Wilson saw some benefits in an avenue of redress through the courts "because the civil courts have a history and an experience of evaluating emotional and psychological harm and awarding damages on that basis. In a sense, they are trained to do that."[770]

765 Written evidence from Department of Digital, Culture, Media & Sport (OSB0243)
766 Draft Online Safety Bill, CP 405, May 2021, Clause 6 and Clause 18
767 Right to bring a claim for breach of the duties in the Bill
768 Q 71
769 Q 71
770 Q 70

459. The UK GDPR allows an individual to take a data controller to court if they believe there has been a breach of data protection law in the handling of their personal data.[771] The court may order compensation for emotional distress without the need for other damage, although a monetary award for other loss may be made[772] including loss of earnings as a result of exacerbation of a pre-existing serious mental health condition.[773]

460. ***We believe that this Bill is an opportunity to reset the relationship between service providers and users. While we recognise the resource challenges both for individuals in accessing the courts and the courts themselves, we think the importance of issues in this Bill requires that users have a right of redress in the courts. We recommend the Government develop a bespoke route of appeal in the courts to allow users to sue providers for failure to meet their obligations under the Act.***

Access to data in cases of bereavement

461. Mr Russell, who campaigns to reduce suicide and self-harm in young people, told us of his distressing experiences in trying to access the digital data of his 14-year-old daughter Molly Russell, who tragically died in 2017. Digital data does not form part of a deceased person's estate, and next of kin who want access have to apply to each tech company which each have their own processes for such requests. Mr Russell told us that a "typical tech company response" was:

- You must be the Administrator of the Estate, or Legal Personal Representative to make this request.
- The following documents are required to move forward with this option:

  (1) The death certificate.

  (2) A court document that confirms you are the legal personal representative of the decedent.[774]

Mr Russell's distressing experiences are sadly not unique. We have also heard about the distress experienced by the bereaved parents of Frankie Thomas in trying to access their deceased daughter's data.

462. Mr Russell also raised concerns about access to digital data for coroners and other investigatory and regulatory authorities. He asked the Committee to consider measures to ensure digital data is available to investigators and, above all, that other bereaved parents do not have to experience what his family has gone through in accessing the social media of their children.

463. ***Bereaved parents who are looking for answers to the tragic deaths of their children in their digital data should not have to struggle through multiple, lengthy, bureaucratic processes to access that data. We recognise that an automatic right to a child's data would raise privacy and child safety concerns. At the same time, we believe there is more than could be done to make the process more proportionate, straightforward and***

771 Data Protection Act 2018, section 167(1)

772 Data Protection Act 2018, section 168(1) and Fieldfischer, 'Article 82 UK GDPR': https://ukgdpr.fieldfisher.com/chapter-8/article-82-gdpr/ [accessed 1 December 2021]

773 *Grinyer v Plymouth Hospitals NH Trust* [2012] EWCA Civ 1043

774 Ian Russell, Molly Rose Foundation (OSB0233); "decedent" is as in the original communication with Mr Russell.

humane. We recommend that the Government undertake a consultation on how the law, and service's terms and conditions, can be reformed to give access to data to parents when it is safe, lawful and appropriate to do so. The Government should also investigate whether the regulator could play a role in facilitating co-operation between the major online service providers to establish a single consistent process or point of application.

464. ***We also recommend Ofcom, the Information Commissioner and the Chief Coroner review the powers of coroners to ensure that they can access digital data following the death of a child. We recommend the Government legislate, if it is required, to ensure that coroners are not obstructed by service providers when they require access to digital data. We recommend that guidance is issued to coroners and regulatory authorities to ensure they are aware of their powers in dealing with service providers and of the types of cases where digital data is likely to be relevant. Our expectation is that the Government will look to implement the outcomes of these consultations in the Bill during its parliamentary passage.***

11 Conclusion

465. We welcome the Government creating this Joint Committee and subjecting the draft Bill to pre-legislative scrutiny. As we and our witnesses have reiterated time and again, online safety is one of the most complex and most fundamental policy issues of our age. We hope that even those who do not agree with our conclusions will accept that the Bill will be the better for these issues being aired in the collaborative environment of a Joint Select Committee, prior to consideration on the floors of both Houses.

466. The major online services have become central to many of our lives, but as their power has grown, there has been no meaningful increase in their public accountability. The harmful user experiences which have emerged on many platforms have been allowed to run unchallenged and unchecked for too long. Now is the time to see platforms held to account for harms which arise from the decisions they make about their systems and processes, the operation of their services, and their corporate governance structures. These decisions are not neutral, nor are they immutable. Online services can and should be accountable to the Regulator, and thus ultimately to the public, about the impact of the decisions that they make. Our hope is that this Bill will make this kind of meaningful accountability possible, and in turn that more accountability will lead to better systems and processes, better operation of services, better corporate governance structures at the major platforms, and a better experience for users who can make informed decisions about the services that they use. A safer internet is possible, and this Bill is a major step towards achieving it.

467. Our inquiry into the draft Bill combined internal expertise from Members of the Committee, many of whom have been part of previous public policy and parliamentary efforts in this space across many years, with contributions from expert witnesses from a wide range of sectors. We heard from over 50 witnesses across 11 meetings, held four roundtables, and received over 200 pieces of written evidence. The result was a wide-ranging, robust, at times challenging debate and discussion, during which we sought to establish what the Bill meant in practice for platforms and, most crucially, for the people who use them. We came away from those discussions certain that people's rights must be protected against an ever-growing onslaught of online harms. This can be achieved through the shaping of a Bill which is practical, implementable, and which will empower the Regulator to take decisive action against platforms which neglect their safety duties. We wholeheartedly support the ambition of the Online Safety Bill—to make the United Kingdom the safest place in the world to be online—and we trust that our recommendations will bring the final Act closer to achieving that aim.

468. We are very grateful for the co-operation that has been shown by the Department and by public bodies, including Ofcom and the ICO, to us during the scrutiny of the draft Bill. We are particularly grateful for the positive way that new ministers have engaged with us and expressed themselves open to our findings. At the same time, much of the draft Bill depends on definitions of types of content, codes and thresholds set by secondary legislation. Some indicative examples published "without prejudice" alongside the draft Bill would have greatly facilitated scrutiny. We also regret that the Government was not willing to publish an ECHR memorandum for the draft Bill.

469. **This Report must be understood as a whole document, comprising a cohesive set of recommendations working in tandem to produce a new vision of the Online Safety**

Act. The Government should not seek to isolate single recommendations without understanding how they fit into the wider manifesto laid out by the Committee. Taken as a whole, our recommendations will ensure that the Bill holds platforms to account for the risks of harm which arise on them and will achieve the Government's ultimate aim of making the United Kingdom the safest place in the world to be online.

Conclusions and recommendations

Chapter 2: Objectives of the Online Safety Bill

1. Self-regulation of online platforms has failed. Our recommendations will strengthen the Bill so that it can pass successfully into legislation. To achieve success, the Bill must be clear from the beginning about its objectives. These objectives must reflect the nature of the harm experienced online and the values of UK society. Online services are not neutral repositories for information. Most are advertising businesses. Service providers in scope of the Bill must be held liable for failure to take reasonable steps to combat reasonably foreseeable harm resulting from the operation of their services. (Paragraph 51)

2. *We recommend the Bill is restructured. It should set out its core objectives clearly at the beginning. This will ensure clarity to users and regulators about what the Bill is trying to achieve and inform the detailed duties set out later in the legislation. These objectives should be that Ofcom should aim to improve online safety for UK citizens by ensuring that service providers:*

    a) *comply with UK law and do not endanger public health or national security;*

    b) *provide a higher level of protection for children than for adults;*

    c) *identify and mitigate the risk of reasonably foreseeable harm arising from the operation and design of their platforms;*

    d) *recognise and respond to the disproportionate level of harms experienced by people on the basis of protected characteristics;*

    e) *apply the overarching principle that systems should be safe by design whilst complying with the Bill;*

    f) *safeguard freedom of expression and privacy; and*

    g) *operate with transparency and accountability in respect of online safety.* (Paragraph 52)

3. The draft Bill creates an entirely new regulatory structure and deals with difficult issues around rights and safety. In seeking to regulate large multinational companies with the resources to undertake legal challenges, it has to be comprehensive and robust. At the same time, a common theme in the evidence we received is that the draft Bill is too complex, and this may harm public acceptance and make it harder for those service providers who are willing to comply to do so. (Paragraph 59)

4. *We recommend that the Bill be restructured to contain a clear statement of its core safety objectives—as recommended in paragraph 52. Everything flows from these: the requirement for Ofcom to meet those objectives, its power to produce mandatory codes of practice and minimum quality standards for risk assessments in order to do so, and the requirements on service providers to address and mitigate reasonably foreseeable risks, follow those codes of practice and meet those minimum standards. Together,*

these measures amount to a robust framework of enforceable measures that can leave no doubt that the intentions of the Bill will be secured. (Paragraph 60)

5. *We believe there is a need to clarify that providers are required to comply with all mandatory Codes of Practice as well as the requirement to include reasonably foreseeable risks in their risk assessments. Combined with the requirements for system design we discuss in the next chapter, these measures will ensure that regulated services continue to comply with the overall objectives of the Bill—and that the Regulator is afforded maximum flexibility to respond to a rapidly changing online world.* (Paragraph 61)

Chapter 3: Societal harm and the Role of Platform Design

6. *We recommend that references to harmful "content" in the Bill should be amended to "regulated content and activity". This would better reflect the range of online risks people face and cover new forms of interaction that may emerge as technology advances. It also better reflects the fact that online safety is not just about moderating content. It is also about the design of platforms and the ways people interact with content and features on services and with one another online.* (Paragraph 68)

7. We heard throughout our inquiry that there are design features specific to online services that create and exacerbate risks of harm. Those risks are always present, regardless of the content involved, but only materialise when the content concerned is harmful. For example, the same system that allows a joke to go viral in a matter of minutes also does the same for disinformation about drinking bleach as a cure for COVID-19. An algorithm that constantly recommends pictures of cats to a cat-lover is the same algorithm that might constantly recommend pictures of self-harm to a vulnerable teenager. Tackling these design risks is more effective than just trying to take down individual pieces of content (though that is necessary in the worst cases). Online services should be identifying these design risks and putting in place systems and process to mitigate them before people are harmed. The Bill should recognise this. Where online services are not tackling these design risks, the regulator should be able to take that into account in enforcement action. (Paragraph 81)

8. *We recommend that the Bill includes a specific responsibility on service providers to have in place systems and processes to identify reasonably foreseeable risks of harm arising from the design of their platforms and take proportionate steps to mitigate those risks of harm. The Bill should set out a non-exhaustive list of design features and risks associated with them to provide clarity to service providers and the regulator which could be amended by Parliament in response to the development of new technologies. Ofcom should be required to produce a mandatory Safety by Design Code of Practice, setting out the steps providers will need to take to properly consider and mitigate these risks. We envisage that the risks, features and mitigations might include (but not be limited to):*

    a) *Risks created by algorithms to create "rabbit holes", with possible mitigations including transparent information about the nature of recommendation algorithms and user control over the priorities they set, measures to introduce diversity of content and approach into recommendations and to allow people to deactivate recommendations from users they have not chosen to engage with;*

b) *Risks created by auto playing content, mitigated through limits on auto-play and auto-recommendation;*

c) *Risks created by frictionless cross-platform activity, with mitigations including warnings before following a link to another platform and ensuring consistent minimum standards for age assurance;*

d) *Risks created through data collection and the microtargeting of adverts, mitigated through minimum requirements for transparency around the placement and content of such adverts;*

e) *Risks created by virality and the frictionless sharing of content at scale, mitigated by measures to create friction, slow down sharing whilst viral content is moderated, require active moderation in groups over a certain size, limit the number of times content can be shared on a "one click" basis, especially on encrypted platforms, have in place special arrangements during periods of heightened risk (such as elections, major sporting events or terrorist attacks); and*

f) *Risks created by default settings on geolocation, photo identification/sharing and other functionality leading to victims of domestic violence or VAWG being locatable by their abusers, mitigated through default strong privacy settings and accessible guidance to victims of abuse on how to secure their devices and online services.* (Paragraph 82)

9. *We recommend that the Bill include a requirement for service providers to co-operate to address cross-platform risks and on the regulator to facilitate such co-operation.* (Paragraph 83)

10. Anonymous abuse online is a serious area of concern that the Bill needs to do more to address. The core safety objectives apply to anonymous accounts as much as identifiable ones. At the same time, anonymity and pseudonymity are crucial to online safety for marginalised groups, for whistleblowers, and for victims of domestic abuse and other forms of offline violence. Anonymity and pseudonymity themselves are not the problem and ending them would not be a proportionate response. The problems are a lack of traceability by law enforcement, the frictionless creation and disposal of accounts at scale, a lack of user control over the types of accounts they engage with and a failure of online platforms to deal comprehensively with abuse on their platforms. (Paragraph 91)

11. *We recommend that platforms that allow anonymous and pseudonymous accounts should be required to include the resulting risks as a specific category in the risk assessment on safety by design. In particular, we would expect them to cover, where appropriate: the risk of regulated activity taking place on their platform without law enforcement being able to tie it to a perpetrator, the risk of 'disposable' accounts being created for the purpose of undertaking illegal or harmful activity, and the risk of increased online abuse due to the disinhibition effect.* (Paragraph 92)

12. *We recommend that Ofcom be required to include proportionate steps to mitigate these risks as part of the mandatory Code of Practice required to support the safety by design requirement we recommended in paragraph 82. It would be for them to decide*

what steps would be suitable for each of the risk profiles for online services. Options they could consider might include (but would not be limited to):

a) *Design measures to identify rapidly patterns of large quantities of identical content being posted from anonymous accounts or large numbers of posts being directed at a single account from anonymous accounts;*

b) *A clear governance process to ensure such patterns are quickly escalated to a human moderator and for swiftly resolving properly authorised requests from UK law enforcement for identifying information relating to suspected illegal activity conducted through the platform, within timescales agreed with the regulator;*

c) *A requirement for the largest and highest risk platforms to offer the choice of verified or unverified status and user options on how they interact with accounts in either category;*

d) *Measures to prevent individuals who have been previously banned or suspended for breaches of terms and conditions from creating new accounts; and*

e) *Measures to limit the speed with which new accounts can be created and achieve full functionality on the platform.* (Paragraph 93)

13. *We recommend that the Code of Practice also sets out clear minimum standards to ensure identification processes used for verification protect people's privacy—including from repressive regimes or those that outlaw homosexuality. These should be developed in conjunction with the Information Commissioner's Office and following consultation with groups including representatives of the LGBTQ+ community, victims of domestic abuse, journalists, and freedom of expression organisations. Enforcement of people's data privacy and data rights would remain with the Information Commissioner's Office, with clarity on information sharing and responsibilities.* (Paragraph 94)

14. We recognise the difficulties with legislating for societal harms in the abstract. At the same time, the draft Bill's focus on individuals potentially means some content and activity that is illegal may not be regulated. We discuss this further in Chapter 4. (Paragraph 106)

15. The viral spread of misinformation and disinformation poses a serious threat to societies around the world. Media literacy is not a standalone solution. We have heard how small numbers of people are able to leverage online services' functionality to spread disinformation virally and use recommendation tools to attract people to ever more extreme behaviour. This has resulted in large scale harm, including deaths from COVID-19, from fake medical cures, and from violence. We recommend content-neutral safety by design requirements, set out as minimum standards in mandatory codes of practice. These will be a vital part of tackling regulated content and activity that creates a risk of societal harm, especially the spread of disinformation. For example, we heard that a simple change, introducing more friction into sharing on Facebook, would have the same effect on the spread of mis- and disinformation as the entire third-party fact checking system. (Paragraph 107)

16. Later in this report we also recommend far greater transparency around system design, and particularly automated content recommendation. This will ensure the regulator and researchers can see what the platforms are doing, assess the impact it has and, in the case of users, make informed decisions about how they use platforms. Online services being required to publish data on the most viral pieces of content on their platform would be a powerful transparency tool, as it will rapidly highlight platforms where misinformation and disinformation is drowning out other content. (Paragraph 108)

17. Many online services have terms and conditions about disinformation, though they are often inconsistently applied. We recommend later a statutory requirement on service providers to apply their terms and conditions consistently, and to produce a clear and concise online safety policy. Later, we identify two areas of disinformation—public health and election administration—which are or will soon be covered in the criminal law and that we believe should be tackled directly by the Bill. (Paragraph 109)

18. *As a result of recommendations made in this report, regulation by Ofcom should reduce misinformation and disinformation by:*

    - *Requiring a consistent enforcement of the providers' own terms and conditions to address user content that is in breach of those terms of service (see Chapter 11);*
    - *Working with the Advertising Standards Authority to address paid content that is in breach of ASA rules (see Chapter 6);*
    - *Use the Safety by Design Code of Practice set out in paragraph 82 to address the spread of misinformation by recommendation algorithms, frictionless sharing of content at scale, use of fake accounts and bots to share malign content and other features that make content viral;*
    - *Publishing codes of practice on Regulated Activity; and*
    - *Improvements to the responsiveness of the complaints processes operated by service providers.*

    The Joint Committee that we recommend later in this report should take forward work to define and make recommendations on how to address other areas of disinformation and emerging threats. (Paragraph 110)

19. *Disinformation and misinformation surrounding elections are a risk to democracy. Disinformation which aims to disrupt elections must be addressed by legislation. If the Government decides that the Online Safety Bill is not the appropriate place to do so, then it should use the Elections Bill which is currently making its way through Parliament.* (Paragraph 111)

20. *The Information Commissioner, Elizabeth Denham, has stated that the use of inferred data relating to users' special characteristics as defined in data protection legislation, including data relating to sexual orientation, and religious and political beliefs, would not be compliant with the law. This would include, for example, where a social media company has decided to allow users to be targeted with content based on their data*

special characteristics without their knowledge or consent. Data profiling plays an important part in building audiences for disinformation, but also has legitimate and valuable uses. Ofcom should consult with the Information Commissioner's Office to determine the best course of action to be taken to investigate this and make recommendations on its legality. (Paragraph 112)

Chapter 4: Safety duties relating to adults

21. We have received a large amount of evidence in our inquiry but very little of it takes issue with the regulation of illegal content. This seems to us to point to a self-evident truth, that regulation of illegal content online is relatively uncontroversial and should be the starting point of the Bill. (Paragraph 118)

22. We believe the scope of the Bill on illegal content is too dependent on the discretion of the Secretary of State. This downplays the fact that some content that creates a risk of harm online potentially amounts to criminal activity. The Government has said it is one of the key objectives of the Bill to remove this from the online world. (Paragraph 126)

23. *We recommend that criminal offences which can be committed online appear on the face of the Bill as illegal content. This should include (but not be limited to) hate crime offences (including the offences of "stirring up" hatred), the offence of assisting or encouraging suicide, the new communications offences recommended by the Law Commission, offences relating to illegal, extreme pornography and, if agreed by Parliament, election material that is disinformation about election administration, has been funded by a foreign organisation targeting voters in the UK or fails to comply with the requirement to include information about the promoter of that material in the Elections Bill.* (Paragraph 127)

24. Implementation of the Law Commission's recommendations on reforming the Communications Offences and Hate Crime will allow the behaviour covered by the new offences to be deemed illegal content. We believe this is a significant enhancement of the protections in the Bill, both for users online but also for freedom of expression by introducing greater certainty as to content that online users should be deterred from sharing. We discuss how to address concerns about ambiguity and the context-dependent nature of the proposed harm-based offence through a statutory public interest requirement in Chapter 7. (Paragraph 135)

25. *We endorse the Law Commission's recommendations for new criminal offences in its reports, Modernising Communications Offences and Hate Crime Laws. The reports recommend the creation of new offences in relation to cyberflashing, the encouragement of serious self-harm, sending flashing images to people with photo-sensitive epilepsy with intent to induce a seizure, sending knowingly false communications which intentionally cause non-trivial emotional, psychological, or physical harm, communications which contain threats of serious harm and stirring up hatred on the grounds of sex or gender, and disability. We welcome the Secretary of State's intention to accept the Law Commission's recommendations on the Communications Offences. The creation of these new offences is absolutely essential to the effective system of online safety regulation which we propose in this report. We recommend that the Government bring in the Law Commission's proposed Communications and Hate*

Crime offences with the Online Safety Bill, if no faster legislative vehicle can be found. Specific concerns about the drafting of the offences can be addressed by Parliament during their passage. (Paragraph 136)

26. *The Government must commit to providing the police and courts with adequate resources to tackle existing illegal content and any new offences which are introduced as a result of the Law Commission's recommendations.* (Paragraph 138)

27. *We recommend that Ofcom be required to issue a binding Code of Practice to assist providers in identifying, reporting on and acting on illegal content, in addition to those on terrorism and child sexual exploitation and abuse content. As a public body, Ofcom's Code of Practice will need to comply with human rights legislation (currently being reviewed by the Government) and this will provide an additional safeguard for freedom of expression in how providers fulfil this requirement. With this additional safeguard, and others we discuss elsewhere in this report, we consider that the test for illegal content in the Bill is compatible with an individual's right to free speech, given providers are required to apply the test in a proportionate manner that is set out in clear and accessible terms to users of the service.* (Paragraph 144)

28. *We recommend that the highest risk service providers are required to archive and securely store all evidence of removed content from online publication for a set period of time, unless to do so would in itself be unlawful. In the latter case, they should store records of having removed the content, its nature and any referrals made to law enforcement or the appropriate body.* (Paragraph 145)

29. *We recommend that the Secretary of State's power to designate content relating to an offence as priority illegal content should be constrained. Given that illegal content will in most cases already be defined by statute, this power should be restricted to exceptional circumstances, and only after consultation with the Joint Committee of Parliament that we recommend in Chapter 9, and implemented through the affirmative procedure. The Regulator should also be able to publish recommendations on the creation of new offences. We would expect the Government, in bringing forward future criminal offences, to consult with Ofcom and the Joint Committee as to whether they should be designated as priority illegal offences in the legislation that creates them.* (Paragraph 148)

30. Clause 11 of the draft Bill has been widely criticised for its breadth and for delegating the authority of the state to service providers over the definition of content that is harmful and what they should do about it. We understand its aims and that the Government intended it primarily as a transparency measure over something companies are already doing. As drafted, however, it has profound implications for freedom of speech, is likely to be subject to legal challenge and yet may also allow companies to continue as they have been in failing to tackle online harm. (Paragraph 174)

31. We agree that the criminal law should be the starting point for regulation of potentially harmful online activity, and that safety by design is critical to reduce its prevalence and reach. At the same time, some of the key risks of harm identified in our evidence are legislated for in parts of the offline world, but not online, where the criminal law is recognised as needing reform, or where drafting that makes sense

in the context of determining individual guilt would allow companies to challenge attempts to make them act. A law aimed at online safety that does not require companies to act on misogynistic abuse or stirring up hatred against disabled people, to give two examples, would not be credible. Leaving such abuse unregulated would itself be deeply damaging to freedom of speech online. (Paragraph 175)

32. *We recommend that Clause 11 of the draft Bill is removed. We recommend that it is replaced by a statutory requirement on providers to have in place proportionate systems and processes to identify and mitigate reasonably foreseeable risks of harm arising from regulated activities defined under the Bill. These definitions should reference specific areas of law that are recognised in the offline world, or are specifically recognised as legitimate grounds for interference in freedom of expression. For example, we envisage it would include:*

- *Abuse, harassment or stirring up of violence or hatred based on the protected characteristics in the Equality Act 2010 or the characteristics for which hatred may be an aggravating factor under Crime and Disorder Act 1998 and section 66 of the Sentencing Act 2020;*
- *Content or activity likely to cause harm amounting to significant psychological distress to a likely audience (defined in line with the Law Commission offence);*
- *Threatening communications that would lead a reasonable person to fear that the threat might be carried out;*
- *Knowingly false communications likely to cause significant physical or psychological harm to a reasonable person;*
- *Unsolicited sending of pictures of genitalia;*
- *Disinformation that is likely to endanger public health (which may include anti-vaccination disinformation);*
- *Content and activity that promotes eating disorders and self-harm;*
- *Disinformation that is likely to undermine the integrity and probity of electoral systems.* (Paragraph 176)

33. *As with the other safety duties, we recommend that Ofcom be required to issue a mandatory code of practice to service providers on how they should comply with this duty. In doing so they must identify features and processes that facilitate sharing and spread of material in these named areas and set out clear expectations of mitigation and management strategies that will form part of their risk assessment, moderation processes and transparency requirements. While the code may be informed by particular events and content, it should be focused on the systems and processes of the regulated service that facilitates or promotes such activity rather than any individual piece of content. We envisage that this code would include (but not be limited to):*

- *the moderation of user generated content to cover the use of AI for moderation;*
- *the appropriate thresholds for human oversight;*

- *the level of expertise needed for human moderation;*
- *dedicated teams for election periods and involve relevant bodies–with planned circuit breakers;*
- *the use of fact checking in proportion to reach and risk;*
- *a transparency requirement on the top 20 viral messages, published on a monthly basis;*
- *user control over their curation, including being joined to groups without permission; and,*
- *targeting through protected characteristics and or political affiliation.* (Paragraph 177)

34. Accepting these recommendations would create a narrower, but stronger, regulatory requirement for service providers to identify and mitigate risks of harm in the online world that may not necessarily meet the criminal thresholds, but which are based on the same criteria as those thresholds, indicating that society has recognised they are legitimate reasons to interfere with freedom of speech rights. It would place these areas on the face of the Bill and remove the broad delegation of decisions on what is harmful from service providers. (Paragraph 178)

35. We recognise that the broad power to define new types of content that is harmful to adults in secondary legislation was a key concern with Clause 11. We recognise that there will need to be the ability to amend what is covered by this proposal to ensure that the Bill is futureproofed. At the same time, it needs to be tightly proscribed and subject to active parliamentary scrutiny and review. (Paragraph 179)

36. *We recommend that additions to the list of content that is harmful should be by statutory instrument from the Secretary of State. The statutory instrument should be subject to approval by both Houses, following a report from the Joint Committee we propose in Chapter 9. Ofcom, when making recommendations, will be required by its existing legal obligations to consider proportionality and freedom of speech rights. The Joint Committee should be specifically asked to report on whether the proposed addition is a justified interference with freedom of speech rights.* (Paragraph 180)

37. The original Clause 11 in the draft Bill, in common with the other safety duties, required providers to produce clear and accessible terms of service and enforce them consistently in relation to content harmful to adults. While we have recommended a narrower but stronger regulatory requirement for service providers to identify and mitigate risks of harm, the requirements for transparency, clarity and consistency are vital to ensuring users are well informed about how platforms promote content to them and what protections they can expect. Clear, concise and fully accessible terms will allow users to make informed choices. (Paragraph 183)

38. *We recommend that the Bill mandates service providers to produce and publish an Online Safety Policy, which is referenced in their terms and conditions, made accessible for existing users and made prominent in the registration process for new users. This Online Safety Policy should: explain how content is promoted and recommended to users, remind users of the types of activity and content that can be illegal online and*

provide advice on what to do if targeted by content that may be criminal and/or in breach of the service providers' terms and conditions and other related guidelines. (Paragraph 184)

39. *The Online Safety Policy should be produced in an accessible way and should be sent to all users at the point of sign up and, as good practice suggests, at relevant future points. "Accessible" should include accessible to children (in line with the Children's Code), where service providers allow child users, and accessible to people with additional needs, including physical and learning disabilities. Ofcom should produce a Code of Practice for service providers about producing accessible and compliant online safety policies and on how they should make them available to users read at appropriate intervals in line with best practice (for example, when the user is about to undertake an activity for the first time or change a safety-relevant setting).* (Paragraph 185)

40. *We welcome the inclusion of fraud and scams within the draft Bill. Prevention must be prioritised and this requires platform operators to be proactive in stopping fraudulent material from appearing in the first instance, not simply removing it when reported. We recommend that clause 41(4) is amended to add "a fraud offence" under terrorism and child sexual exploitation and abuse offences and that related clauses are similarly introduced or amended so that companies are required to proactively address it. The Government should consult with the regulatory authorities on the appropriate offences to designate under this section. The Government should ensure that this does not compromise existing consumer protection regulation.* (Paragraph 194)

41. *The Bill must make clear that ultimate responsibility for taking action against criminal content remains with the relevant regulators and enforcement bodies, with Ofcom reporting systemic issues relating to platform design and operation—including in response to "super complaints" from other regulators. The Bill should contain provisions requiring information-sharing and regulatory cooperation to facilitate this.* (Paragraph 195)

Chapter 5: Protection of Children

42. The test the Law Commission arrived at for their harm-based offence was "likely to cause harm to a likely audience". We believe this is a better way of ensuring that service providers consider those who may be harmed or impacted by content or activity on a platform than the "person of ordinary sensibilities" test in the draft Bill. Having a single test for a key category of illegal content and for regulated content and activity harmful to children reduces regulatory burden and improves consistency. Online providers generally have a good understanding of their audience. Where their platform allows users to target content at particular people it would require service providers to consider how the design of their systems might be used to create or mitigate harm. (Paragraph 201)

43. *Recognising the key objective of offering a higher level of protection for children than adults, we support the inclusion of a broad definition of content that is harmful to children. At the same time, we believe the definition should be tightened. We recommend that Clauses 10(3) to (8) are revised. Content and activity should be within this section if it is specified on the face of the Bill, in regulations or there is a reasonably*

foreseeably risk that it would be likely to cause significant physical or psychological distress to children who are likely to encounter it on the platform. (Paragraph 202)

44. *As with other duties, we recommend that key, known risks of harm to children are set out on the face of the Bill. We would expect these to include (but not be limited to) access to or promotion of age-inappropriate material such as pornography, gambling and violence material that is instructive in or promotes self-harm, eating disorders or suicide, and features such as functionality that allows adults to make unsupervised contact with children who do not know them, endless scroll, visible popularity metrics, live location, and being added to groups without user permission.* (Paragraph 203)

45. *We recognise the concerns that, without proper guidance, service providers might seek to place disproportionate age assurance measures in place, impacting the rights of both children and adults. We recommend that Ofcom be required to develop a mandatory Code of Practice for complying with the safety duties in respect of children. Ofcom should be required to have regard to the UN Convention on the Rights of the Child (in particular, General Comment No. 25 on children's rights in relation to the digital environment), the Information Commissioner's Office's Age Appropriate Design Code, and children's right to receive information under the ECHR when drawing up that Code.* (Paragraph 204)

46. *We recommend that the "likely to be accessed by children" test in the draft Online Safety Bill should be the same as the test underpinning the Age Appropriate Design Code. This regulatory alignment would simplify compliance for businesses, whilst giving greater clarity to people who use the service, and greater protection to children. We agree that the Information Commissioner's Office and Ofcom should issue a Joint Statement on how the two regulatory systems will interact once the Online Safety Bill has been introduced. They should be given powers to cooperate on shared investigations, with appropriate oversight.* (Paragraph 211)

47. Easy, often unwitting or unintended, access by children to pornography was one of the largest online concerns raised with us during our scrutiny of the draft Bill. It is evident to us that the credibility of the Bill will be undermined if the largest online pornography providers simply remove user-to-user elements from their sites and continue showing extreme content and content that creates a risk of harm to children. (Paragraph 221)

48. Whilst there is a case for specific provisions in the Bill relating to pornography, we feel there is more to be gained by further aligning the Bill with the Age Appropriate Design Code. Whilst we understand the concerns over scope and regulatory burden, this provision would only bring within the scope of the Bill services already covered by the scope of the Age Appropriate Design Code. Both regulatory systems are risk-based and require the regulator to act proportionately. This step would address the specific concern around pornography, requiring all such sites to demonstrate that they have taken appropriate steps to prevent children from accessing their content. It would also bring other sites or services that create a risk of harm into scope whilst bringing us closer to the goal of aligned online regulation across data protection and online safety. We believe that our proposal on expanding the role of risk profiles, discussed later in this report, will be key to ensure that the Bill's provisions

impact the riskiest services and are not disproportionate on those at lower risk. (Paragraph 222)

49. *All statutory requirements on user-to-user services, for both adults and children, should also apply to Internet Society Services likely to be accessed by children, as defined by the Age Appropriate Design Code. This would have many advantages. In particular, it would ensure all pornographic websites would have to prevent children from accessing their content. Many such online services present a threat to children both by allowing them access and by hosting illegal videos of extreme content.* (Paragraph 223)

50. There is currently no single regulatory or statutory code in the UK that sets out rules for age assurance. We believe that existing codes, and the duties outlined in the draft Bill, cannot be implemented properly without a statutory system of regulation of age assurance, that is trusted, effective and preserves privacy. We believe that an independent, privacy-protecting age assurance sector operating to a set of minimum standards appropriate for different methods of age assurance in different circumstances is key to any system that aims to protect children from harm online. Such a system:

    a) should be for independent commercial providers as well those built by the service providers themselves;

    b) should impose standards appropriate to the content and age of the user and be compatible with existing law, including international treaties such as the UN Convention of the Rights of the Child, to provide necessary protections for privacy and data protection; and

    c) should provide a route of redress for users to challenge specific conclusions reached on age.

    A binding Code of Practice would provide a clear basis for service providers whose risk assessment identifies their content as likely to be accessed by children to put in place mitigations in the form of a rigorous system of age assurance. (Paragraph 235)

51. *We recommend that the Bill require Ofcom to establish minimum standards for age assurance technology and governance linked to risk profiles to ensure that third-party and provider-designed assurance technologies are privacy-enhancing, rights-protecting, and that in commissioning such services providers are restricted in the data for which they can ask. Ofcom should also require that service providers demonstrate to them how they monitor the effectiveness of these systems to ensure that they meet the minimum standards required.* (Paragraph 236)

52. *The Government should ask Ofcom to prioritise the development of a mandatory age assurance technology and governance code as a priority ahead of the Bill becoming law and, in doing so, set out risk profiles so that the use of such systems is clearly proportionate to the risk. The code must bear in mind that children have rights to freedom of association, participation, and information, as well as the right to protections. We expect this to be in place within three to six months of the Bill receiving Royal Assent.* (Paragraph 237)

Chapter 6: Scope of the draft Bill

53. *We recommend that the categorisation of services in the draft Bill be overhauled. It should adopt a more nuanced approach, based not just on size and high-level functionality, but factors such as risk, reach, user base, safety performance, and business model. The draft Bill already has a mechanism to do this: the risk profiles that Ofcom is required to draw up. We make recommendations in Chapter 8 about how the role of the risk profiles could be enhanced. We recommend that the risk profiles replace the "categories" in the Bill as the main way to determine the statutory requirements that will fall on different online services. This will ensure that small, but high risk, services are appropriately regulated; whilst guaranteeing that low risk services, large or small, are not subject to unnecessary regulatory requirements.* (Paragraph 246)

54. We recognise that search engines operate differently from social media and that the systems and processes required to meet the separate duties that the draft Bill places on them are different. The codes of practice drawn up by Ofcom will need to recognise the specific circumstances of search engines to meet Ofcom's duties on proportionality. Search engines are more than passive indexes. They rely on algorithmic ranking and often include automatic design features like autocomplete and voice activated searches that can steer people in the direction of content that puts them or others at risk of harm. Most search engines already have systems and processes in place to address these and comply with other legislation. It is reasonable to expect them to come under the Bill's requirements and, in particular, for them to conduct risk assessments of their system design to ensure it mitigates rather exacerbates risks of harm. We anticipate that they will have their own risk profiles. (Paragraph 251)

55. The Government needs to provide more clarity on how providers with encrypted services should comply with the safety duties ahead of the Bill being introduced into Parliament. (Paragraph 257)

56. *We recommend that end-to-end encryption should be identified as a specific risk factor in risk profiles and risk assessments. Providers should be required to identify and address risks arising from the encrypted nature of their services under the Safety by Design requirements.* (Paragraph 258)

57. The exclusion of paid-for advertising from the scope of the Online Safety Bill would obstruct the Government's stated aim of tackling online fraud and activity that creates a risk of harm more generally. Excluding paid-for advertising will leave service providers with little incentive to remove harmful adverts, and risks encouraging further proliferation of such content. (Paragraph 268)

58. *We therefore recommend that clause 39(2) is amended to remove "(d) paid-for advertisements" to bring such adverts into scope. Clause 39(7) and clause 134(5) would therefore also have to be removed.* (Paragraph 269)

59. Ofcom should be responsible for acting against service providers who consistently allow paid-for advertisements that create a risk of harm to be placed on their platform. However, we agree that regulating advertisers themselves (except insofar as they come under other provisions of the Bill), individual cases of advertising that

are illegal, and pursuing the criminals behind illegal adverts should remain matters for the existing regulatory bodies and the police. (Paragraph 270)

60. *We recommend that the Bill make clear Ofcom's role will be to enforce the safety duties on providers covered by the online safety regulation, not regulate the day-to-day content of adverts or the actions of advertisers. That is the role of the Advertising Standards Authority. The Bill should set out this division of regulatory responsibility.* (Paragraph 271)

61. We recognise that economic harms other than fraud, such as those impacting consumers, and infringement of intellectual property rights, are an online problem that must be tackled. However, the Online Safety Bill is not the best piece of legislation to achieve this. Economic harms should be addressed in the upcoming Digital Competition Bill. We urge the Government to ensure this legislation is brought forward as soon as possible. (Paragraph 275)

Chapter 7: Freedom of speech requirements, journalism, and content of democratic importance

62. We propose a series of recommendations throughout this report to strengthen protection for freedom of expression. These include greater independence for Ofcom, routes for individual redress beyond service providers, tighter definitions around content that creates a risk of harm, a greater emphasis on safety by design, a broader requirement to be consistent in the applications of terms of service, stronger minimum standards and mandatory codes of practice set by Ofcom (who are required to be compliant with human rights law), and stronger protections for news publisher content. We believe these will be more effective than adjustments to the wording of Clause 12. (Paragraph 284)

63. *We recommend that Ofcom be required to produce an annual report on the impact of regulated services on media plurality.* (Paragraph 291)

64. *We recommend that the news publisher content exemption is strengthened to include a requirement that news publisher content should not be moderated, restricted or removed unless it is content the publication of which clearly constitutes a criminal offence, or which has been found to be unlawful by order of a court within the appropriate jurisdiction. We recommend that the Government look at how bad actors can be excluded from the concept of news publisher. We suggest that they may wish to exclude those that have been repeatedly found to be in breach of The Ofcom Broadcasting Code, or are publications owned by foreign Governments. Ofcom should also examine the use of new or existing registers of publishers. We are concerned that some consumer and business magazines, and academic journals, may not be covered by the Clause 40 exemptions. We recommend that the Department consult with the relevant industry bodies to see how the exemption might be amended to cover this off, without creating loopholes in the legislation.* (Paragraph 304)

65. The draft Bill already makes a distinction between "news publisher content" and citizen journalism, in recognition that the former is subject to editorial control and there are existing mechanisms for accountability. There is also a clear difference between the categories, as one is based on "who" is sharing the content, and the other

focuses on the purpose of the content, rather than the identity of those behind it. For both citizen journalism and content of democratic importance, the justification for special consideration appears to be that they are in the public interest to be shared. This should therefore be key to any final definition and providers will require guidance as to how to balance the risk of harm with the public interest. It is not, nor is it intended to be, a blanket exemption in the same way as that for news publisher content, but a counterbalance to prevent overzealous moderation, particularly in borderline cases. (Paragraph 305)

66. Our recommendations to narrowly define content that is harmful to adults by way of reference to existing law should provide some of the extra clarity service providers need to help protect freedom of expression. At the same time, journalism and content of democratic importance have long been recognised as vital in a democratic society and should be given specific consideration and protection by providers, who have significant influence over the information we see. We have heard concerns around the definitions used however, and about the ability of the providers to interpret and apply them consistently. We feel that "democratic importance" may be both too broad—creating a loophole to be exploited by bad actors—and too narrow—excluding large parts of civil society. Similarly, we are concerned that any definition of journalistic content that is designed to capture citizen journalism would be so broad it would render the consistent application of the requirement almost impossible, and see the expedited complaints route overwhelmed by people claiming without merit to be journalists in order to have their content reinstated. "Public interest" might be more useful in ensuring that content and activity is judged on its merit, rather than its author. (Paragraph 306)

67. *We recommend that the existing protections around journalistic content and content of democratic importance should be replaced by a single statutory requirement to have proportionate systems and process to protect 'content where there are reasonable grounds to believe it will be in the public interest'. Examples of content that would be likely to be in the public interest would be journalistic content, contributions to political or societal debate and whistleblowing. Ofcom should produce a binding Code of Practice on steps to be taken to protect such content and guidance on what is likely to be in the public interest, based on their existing experience and case law. This should include guidance on how appeals can be swiftly and fairly considered. Ofcom should provide guidance to companies in cases of systemic, unjustified take down of content that is likely to be in the public interest. This would amount to a failure to safeguard freedom of expression as required by the objectives of the legislation.* (Paragraph 307)

Chapter 8: Role of the regulator

68. Robust regulatory oversight is critical to ensuring the ambition of the Online Safety Bill is fully met. Tech companies must not be allowed to snub the Regulator, to act with impunity, to continue to rely on self-regulation, or to abdicate responsibility for the harms which occur through the operation of their services or because of their governance structures. In turn, Ofcom must be able to move at pace to hold providers to account authoritatively to issue substantial fines, and assist the appropriate authorities with criminal prosecutions. The Bill extends substantial

powers to the Regulator, but there are improvements to be made if the Government is to ensure the Bill is enforced effectively. (Paragraph 312)

69. *Ofcom should have the power on the face of the Bill to share information and to co-operate with international regulators at its discretion.* (Paragraph 315)

70. *To help differentiate between the risk assessment undertaken by the regulator and that undertaken by the service providers, Ofcom's risk assessment should be renamed the "Ofcom register of risks of regulated services" (henceforth, register of risks). Ofcom should begin working on this immediately so that it is ready to be actioned when the Bill becomes law.* (Paragraph 317)

71. *The Bill's provision that Ofcom should develop risk profiles based on the characteristics of services should be strengthened. Ofcom should begin drawing up risk profiles immediately so that they are ready to be actioned when the Bill becomes law. Risk profiles should reflect differences in the characteristics of the service. These could include (but are not limited to) risks created by algorithms; risks created by a reliance on artificial intelligence moderation; risks created by unlimited 'one-click' sharing; risks caused by "engagement" maximising design features; risk of unsupervised contact between adults and children which may give rise to grooming; risks caused by surveillance advertising; and such other risks as Ofcom identifies in its overall risk assessment, as well as platform design, risk level, end-to-end encryption, algorithmic design, safety by design measures, and the service's business model and overall corporate aim. Ofcom should also be able to take into account whether a company has been the subject of a super complaint, other legal proceedings or publicly documented evidence of poor performance e.g. independent research, a poor monitoring report in the EU's Code of Conduct for Illegal Hate, or whistleblowers' evidence.* (Paragraph 323)

72. *The Bill should be amended to clarify that Ofcom is able to take enforcement action if it identifies a breach of the safety duties, without requiring a provider to redo a risk assessment.* (Paragraph 325)

73. It should not be possible for a service provider to underestimate the level of risk on their service without fear of sanction. If Ofcom suspects such a breach, it should have the power to investigate, and, if necessary, to take swift action. We are not convinced that the draft Bill as it currently stands achieves this. (Paragraph 332)

74. *Ofcom should be required to set binding minimum standards for the accuracy and completeness of risk assessments. Ofcom must be able to require a provider who returns a poor or incomplete risk assessment to redo that risk assessment. Risk assessments should be carried out by service providers as a response to the Online Safety Act before new products and services are rolled out, during the design process of new features, and kept up to date as they are implemented.* (Paragraph 333)

75. *The required content of service providers' risk assessments should follow the risk profiles developed by Ofcom, which in turn should be based on the differences in the characteristics of the service, platform design, risk level, and the service's business model and overall corporate aim. For example, a provider that does not have an engagement-based service would not need to address irrelevant risks associated with*

virality, whilst a site containing adult content would have to address the higher level of risks associated with children accessing the site. (Paragraph 334)

76. *The Bill should be amended to clarify that risk assessments should be directed to "reasonably foreseeable" risks, to allow Ofcom greater leeway to take enforcement action against a company that conducts an inadequate risk assessment.* (Paragraph 335)

77. *Ofcom should look to the Data Protection Impact Assessment as they come to form their own guidance for minimum standards for risk assessments for regulated services.* (Paragraph 336)

78. *In bringing forward the final Bill, we recommend the Government publish an assessment of the audit powers given to Ofcom and a comparison to those held by the Information Commissioner's Office and the Financial Conduct Authority. Parliament should be reassured that the Bill will give Ofcom a suite of powers to match those of similar regulators. Within six months of the Act becoming law, Ofcom should report to Parliament on how it has used those powers.* (Paragraph 339)

79. *We recommend that the largest and highest-risk providers should be placed under a statutory responsibility to commission annual, independent third-party audits of the effects of their algorithms, and of their risk assessments and transparency reports. Ofcom should be given the explicit power to review these and undertake its own audit of these or any other regulated service when it feels it is required. Ofcom should develop a framework for the effective regulation of algorithms based on the requirement for, and auditing of, risk assessments.* (Paragraph 340)

80. *In taking on its responsibilities under the Bill, Ofcom will be working with a network of other regulators and third parties already working in the digital world. We recommend that the Bill provide a framework for how these bodies will work together including when and how they will share powers, take joint action, and conduct joint investigations.* (Paragraph 346)

81. *We reiterate the recommendations by the House of Lords Communications and Digital Committee in their Digital Regulation report: that regulators in the Digital Regulation Cooperation Forum should be under a statutory requirement to cooperate and consult with one another, such that they must respect one another's objectives, share information, share powers, take joint action, and conduct joint investigations; and that to further support coordination and cooperation between digital regulators including Ofcom, the Digital Regulation Cooperation Forum should be placed on a statutory footing with the power to resolve conflicts by directing its members.* (Paragraph 347)

82. *The draft Bill does not give Ofcom co-designatory powers. Ofcom is confident that it will be able to co-designate through other means. The Government must ensure that Ofcom has the power to co-designate efficiently and effectively, and if it does not, this power should be established on the face of the Bill.* (Paragraph 348)

83. *During the course of its duties, Ofcom will be required to investigate companies for a range of breaches, some of which will relate to suspected or known child sexual exploitation and abuse material. As child sexual exploitation and abuse investigations lie so far outside Ofcom's normal duties, we expect Ofcom to work closely with*

experts like the Internet Watch Foundation, to develop and update the child sexual exploitation and abuse Code of Practice; monitor providers to ensure compliance with the child sexual exploitation and abuse code; and during investigations relating to child sexual exploitation and abuse content. (Paragraph 352)

84. *Ofcom may receive unsolicited child sexual exploitation and abuse material which would constitute an offence under Section 1 of the Protection of Children Act 1978. The Bill should be amended to provide Ofcom with a specific defence in law to allow it to perform its duties in this area without inadvertently committing an offence.* (Paragraph 353)

85. *The Bill should be amended to make clear that Codes of Practice should be binding on providers. Any flexibility should be entirely in the hands of and at the discretion of the Regulator, which should have the power to set minimum standards expected of providers. They should be subject to affirmative procedure in all cases.* (Paragraph 358)

86. *Ofcom should start working on Codes of Practice immediately, so they are ready for enforcement as soon as the Bill becomes law. A provisional list of Codes of Practice, including, but not necessarily limited to, those listed in Box 2 above should be included on the face of the Bill. Some of the Codes should be delegated to co-designated bodies with relevant expertise, which would allow work on multiple Codes to happen simultaneously and thus the entire endeavour to be completed more quickly. Once the Codes of Practice are completed, they should be published.* (Paragraph 359)

87. *The Bill should require that companies' risk assessments be reported at Board level, to ensure that senior management know and can be held accountable for the risks present on the service, and the actions being taken to mitigate those risks.* (Paragraph 367)

88. *We recommend that a senior manager at board level or reporting to the board should be designated the "Safety Controller" and made liable for a new offence: the failure to comply with their obligations as regulated service providers when there is clear evidence of repeated and systemic failings that result in a significant risk of serious harm to users. We believe that this would be a proportionate last resort for the Regulator. Like any offence, it should only be initiated and provable at the end of an exhaustive legal process.* (Paragraph 368)

89. *The Committee welcomes the Secretary of State's commitment to introduce criminal liability within three to six months of Royal Assent and strongly recommends that criminal sanctions for failures to comply with information notices are introduced within three months of Royal Assent.* (Paragraph 369)

90. *The power for the Secretary of State to exempt services from regulation should be clarified to ensure that it does not apply to individual services.* (Paragraph 376)

91. *The powers for the Secretary of State to a) modify Codes of Practice to reflect Government policy and b) give guidance to Ofcom give too much power to interfere in Ofcom's independence and should be removed.* (Paragraph 377)

92. *Exercise of the Secretary of State's powers in respect of national security and public safety in respect of terrorism and child sexual exploitation and abuse content*

should be subject to review by the Joint Committee we propose later in this report. (Paragraph 378)

93. If the Government wishes to improve the UK's media literacy to reduce online harms, there must be provisions in the Bill to ensure media literacy initiatives are of a high standard. The Bill should empower Ofcom to set minimum standards for media literacy initiatives that both guide providers and ensure the information they are disseminating aligns with the goal of reducing online harm. (Paragraph 381)

94. *We recommend that Ofcom is made responsible for setting minimum standards for media literacy initiatives. Clause 103 (4) should be amended to include "(d) about minimum standards that media literacy initiatives must meet."* (Paragraph 382)

95. *We recommend that the Bill reflects that media literacy should be subject to a "whole of Government" approach, involving current and future initiatives of the Department of Education in relation to the school curriculum as well as Ofcom and service providers. We have heard throughout this inquiry about the real dangers that some online content and activity poses to children. Ofsted already assesses how schools manage online safety as part of their safeguarding policies. We recommend that Ofsted, in conjunction with Ofcom, update the school inspection framework to extend the safeguarding duties of schools to include making reasonable efforts to educate children to be safe online* (Paragraph 385)

96. *Ofcom should require that media literacy is built into risk assessments as a mitigation measure and require service providers to provide evidence of taking this mitigation measure where relevant.* (Paragraph 386)

97. *We recommend that clause 103(11) is amended to state that Ofcom's media literacy duties relate to "the public" rather than "members of the public", and that the definition of media literacy is updated to incorporate learning about being a good digital citizen and about platform design, data collection and the business models and operation of digital services more broadly.* (Paragraph 388)

98. *The highest risk services, as assessed by Ofcom, should have to report quarterly data to Ofcom on the results of the tools, rules, and systems they have deployed to prevent and remove child sexual exploitation and abuse content (e.g. number and rates of illegal images blocked at upload stage, number and rates of abusive livestreams terminated, number and rates of first- and second- generation images and videos detected and removed).* (Paragraph 394)

99. *Ofcom should have the power to request research and independent evaluation into services where it believes the risk factors for child sexual exploitation and abuse are high.* (Paragraph 395)

100. *Ofcom should move towards a risk factors approach to the regulation of child sexual exploitation and abuse material. It should be able to issue a Use of Technology notice if it believes that there is a serious risk of harm from child sexual exploitation and abuse or terrorism content and that not enough is being done by a service to mitigate those risks. The Bill should be amended to clarify that Ofcom is able to consider a*

wider range of risk factors when deciding whether to issue a Use of Technology notice or take enforcement action. Risk factors should include:

a) *The prevalence or the persistent prevalence of child sexual exploitation and abuse material on a service, or distributed by a service;*

b) *A service's failure to provide and maintain adequate tools, rules, and systems to proactively prevent the spread of child sexual exploitation and abuse content, and to provide information on those tools, rules, and systems to Ofcom when requested;*

c) *A service's failure to provide adequate data to Ofcom on the results of those tools, rules, and systems (e.g., number and rates of illegal images blocked at upload stage, number and rates of abusive livestreams terminated, number and rates of first- and second- generation images and videos detected and removed);*

d) *The nature of a service and its functionalities;*

e) *The user base of a service;*

f) *The risk of harm to UK individuals (and the severity of that harm) if the relevant technology is not used by the service*

g) *The degree of interference posed by the use of the relevant technology with users' rights to freedom of expression and privacy; and*

h) *The safety by design mechanisms that have been implemented.* (Paragraph 396)

Chapter 9: Transparency and oversight

101. *We recommend that Ofcom specify that transparency reports produced by service providers should be published in full in a publicly accessible place. Transparency reports should be written clearly and accessibly so that users and prospective users of the service can understand them, including children (where they are allowed to use the service) and disabled people.* (Paragraph 410)

102. *We recommend that the Bill require transparency reporting on a regular, proportionate basis, with the aim of working towards standardised reporting as the regulatory regime matures. The Bill should require minimum standards of accuracy and transparency about how the report was arrived at and the methodology used in research. For providers of the highest risk services, the outcome of the annual audits recommended in paragraph 340 should be required to be included in the transparency report.* (Paragraph 411)

103. *We agree with the list of information that Ofcom can require as part of its transparency reporting powers and recommend that it should have the clear power to request any other information. We recommend that transparency reporting should aim to create a competitive marketplace in respect of safety, where people can reasonably compare, using robust and comparable information, performance of services as they operate*

for UK users. We suggest Ofcom also be able to require information be published in transparency reports including (but not limited to):

a) *Safety by design features;*

b) *Most viewed/engaged with content by month;*

c) *Most recommended content by month by age group and other demographic information (where that information is collected);*

d) *Their terms and conditions;*

e) *Proportion of users who are children;*

f) *Proportion of anonymous users;*

g) *Proportion of content breaching terms and conditions;*

h) *Proportion of content breaching terms and conditions removed;*

i) *Proportion of appeals against removal upheld;*

j) *Proportion of appeals against removal, by both recognised news publishers and other users on the grounds of public interest, upheld; and*

k) *Time taken to deal with reports.* (Paragraph 412)

104. *In addition to transparency reporting, Ofcom should be empowered to conduct its own independent research with the aim of informing the UK public about the comparative performance of services in respect of online safety.* (Paragraph 413)

105. Independent researchers currently have limited access to the information needed to conduct research. This hinders progress in understanding online activity that creates a risk of harm, the way that services' systems work, and how services' systems could be improved to mitigate the risk of harm. It also limits the ability to scrutinise service providers and hold them accountable. This issue must be addressed urgently. (Paragraph 424)

106. The transparency powers in the Bill are an important opportunity to encourage service providers to share relevant data with external researchers studying online safety and allied subjects. (Paragraph 425)

107. *The draft Bill requires that Ofcom produce a report on access to data for independent researchers. We recommend work on this report starts as soon as possible. We recommend that Ofcom be given the powers in the Bill to put into practice recommendations from that report.* (Paragraph 426)

108. *Ofcom should have the power i) to audit or appoint a third-party to audit how services commission, surface, collate and use their research; ii) to request a) specific internal research from services; b) research on topics of interest to the Regulator.* (Paragraph 427)

109. *Ofcom should commission an independent annual assessment, conducted by skilled persons, of what information should be provided by each of the highest risk services to advance academic research.* (Paragraph 428)

110. *We recommend that the Bill should require service providers to conduct risk assessments of opening up data on online safety to independent researchers, with some pre-defined issues to comment on, including a) privacy; b) risk of harm to users; c) reputational risks (for the service provider) and; d) financial cost* (Paragraph 429)

111. *We recommend that Ofcom should require service providers to conduct an annual formal review of using privacy-protecting technologies and enable them to share sensitive datasets.* (Paragraph 430)

112. *We agree with other Committees that it is imperative that digital regulation be subject to dedicated parliamentary oversight. To achieve this, we recommend a Joint Committee of both Houses to oversee digital regulation with five primary functions: scrutinising digital regulators and overseeing the regulatory landscape, including the Digital Regulation Cooperation Forum; scrutinising the Secretary of State's work into digital regulation; reviewing the codes of practice laid by Ofcom any legislation relevant to digital regulation (including secondary legislation under the Online Safety Act); considering any relevant new developments such as the creation of new technologies and the publication of independent research or whistleblower testimonies; and helping to generate solutions to ongoing issues in digital regulation.* (Paragraph 434)

113. *We fully support the recommendation of the House of Lords Communications and Digital Committee in their report on Digital Regulation that, as soon as possible, full Digital Regulation Cooperation Forum membership should be extended to statutory regulators with significant interests and expertise in the digital sphere, and that partial membership should be extended to non-statutory regulators and advisory bodies with subject specific knowledge to participate on issues particular to their remits.* (Paragraph 435)

114. *We recommend that, in addition to any other reports the Committee chooses to make, the Joint Committee produces an annual report with recommendations on what could or should change, looking towards future developments. We anticipate that the Joint Committee will want to look at the definition of disinformation and what more can be done to tackle it at an early stage.* (Paragraph 436)

115. *We recommend that whistleblowers' disclosure of information to Ofcom and/or the Joint Committee on Digital Regulation, where that information provides clear evidence of non-compliance with the Online Safety Bill, is protected under UK law.* (Paragraph 439)

Chapter 10: Redress

116. *The Bill should establish proportionate minimum standards for the highest risk providers' reports, complaints, and redress mechanisms as set out in a mandatory code of practice prepared by Ofcom.* (Paragraph 443)

117. *We recommend a requirement on the face of the Bill for Ofcom to set out i) how they will assess the a) ease of use; b) accessibility and c) transparency of a service's complaints*

process for d) adults; e) children; and g) disabled people f) vulnerable adults; ii) what steps Ofcom will be able to take if it finds any of these processes wanting; and iii) how Ofcom will ensure that requirements to operate complaint, reporting and redress mechanisms are proportionate for smaller in-scope providers. (Paragraph 444)

118. *Clause 15 (3)(c) should be amended so that it reads "is easy to access, including for disabled people and those with learning difficulties".* (Paragraph 445)

119. *Providers of the highest risk services should have to give quarterly statistics to Ofcom on:*

    i) *Number of user reports;*

    ii) *User reports broken down by the reason the report was made;*

    iii) *Number of actionable user reports;*

    iv) *Actionable user reports broken down by the reason the report was made;*

    v) *How long it took the service provider to respond to i) all user reports; ii) actionable user reports;*

    vi) *What response was made to actionable user reports;*

    vii) *Number of user complaints received;*

    viii) *Number of actionable user complaints;*

    ix) *How long it took the service provider to respond to i) all user complaints; ii) actionable user complaints;*

    x) *What response was made to actionable user complaints;*

    xi) *How many pieces of user content were taken down;*

    xii) *How many pieces of content that were taken down were later reinstated;*

    xiii) *The grounds on which content that was reinstated was reinstated;*

    xiv) *How long it took the service provider to reinstate a piece of content that was later reinstated.* (Paragraph 446)

120. Our proposed external redress process would not replace service providers' internal processes or run concurrently to them, nor would it address individual complaints about individual pieces of content or interactions. Rather, for a victim of sustained and significant online harm, someone who has been banned from a service or who had their posts repeatedly and systematically removed, this new redress mechanism would give them an additional body to appeal those decisions after they had come to the end of a service provider's internal process. (Paragraph 450)

121. In order for an external redress process to work, clear direction is needed in the Bill about Ofcom's responsibility to set quality standards for service provider's internal complaints procedures, and in relation to complaints about failures to meet those standards. We hope that the Government will consider our recommendations in

this area, and that by improving the quality of service providers' internal complaints procedures, any system of external redress will be needed only rarely and for the most serious cases. (Paragraph 451)

122. We support the Government's ambition to make service providers behave responsibly, and by agreeing our recommendations the requirements of the Bill will bring about better responses from service providers to user complaints. However, the fact remains that service providers' user complaints processes are often obscure, undemocratic, and without external safeguards to ensure that users are treated fairly and consistently. It is only through the introduction of an external redress mechanism that service providers can truly be held to account for their decisions as they impact individuals. (Paragraph 456)

123. *The role of the Online Safety Ombudsman should be created to consider complaints about actions by higher risk service providers where either moderation or failure to address risks leads to significant, demonstrable harm (including to freedom of expression) and recourse to other routes of redress have not resulted in a resolution. The right to complain to this Ombudsman should be limited to users to those i) who have exhausted the internal complaints process with the service provider against which they are making their complaint and ii) who have either a) suffered serious or sustained harm on the service or b) had their content repeatedly taken down. There should be an option in the Bill to extend the remit of the Ombudsman to lower risk providers. In addition to handling these complaints the Ombudsman would as part of its role i) identify issues in individual companies and make recommendations to improve their complaint handling and ii) identify systemic industry wide issues and make recommendations on regulatory action needed to remedy them. The Ombudsman should have a duty to gather data and information and report it to Ofcom. It should be an "eligible entity" to make super-complaints.* (Paragraph 457)

124. *We believe that this Bill is an opportunity to reset the relationship between service providers and users. While we recognise the resource challenges both for individuals in accessing the courts and the courts themselves, we think the importance of issues in this Bill requires that users have a right of redress in the courts. We recommend the Government develop a bespoke route of appeal in the courts to allow users to sue providers for failure to meet their obligations under the Act.* (Paragraph 460)

125. *Bereaved parents who are looking for answers to the tragic deaths of their children in their digital data should not have to struggle through multiple, lengthy, bureaucratic processes to access that data. We recognise that an automatic right to a child's data would raise privacy and child safety concerns. At the same time, we believe there is more than could be done to make the process more proportionate, straightforward and humane. We recommend that the Government undertake a consultation on how the law, and service's terms and conditions, can be reformed to give access to data to parents when it is safe, lawful and appropriate to do so. The Government should also investigate whether the regulator could play a role in facilitating co-operation between the major online service providers to establish a single consistent process or point of application.* (Paragraph 463)

126. *We also recommend Ofcom, the Information Commissioner and the Chief Coroner review the powers of coroners to ensure that they can access digital data following*

the death of a child. We recommend the Government legislate, if it is required, to ensure that coroners are not obstructed by service providers when they require access to digital data. We recommend that guidance is issued to coroners and regulatory authorities to ensure they are aware of their powers in dealing with service providers and of the types of cases where digital data is likely to be relevant. Our expectation is that the Government will look to implement the outcomes of these consultations in the Bill during its parliamentary passage. (Paragraph 464)

Chapter 11: Conclusion

127. This Report must be understood as a whole document, comprising a cohesive set of recommendations working in tandem to produce a new vision of the Online Safety Act. The Government should not seek to isolate single recommendations without understanding how they fit into the wider manifesto laid out by the Committee. Taken as a whole, our recommendations will ensure that the Bill holds platforms to account for the risks of harm which arise on them and will achieve the Government's ultimate aim of making the United Kingdom the safest place in the world to be online. (Paragraph 469)

Appendix 1: Case Studies

In this Appendix we set out briefly how the draft Bill and our recommendations will help address some of the online harms we heard about during our inquiry.

Racist abuse

Racist abuse is as unacceptable online as it is offline. Our recommendations will ensure that tech companies put systems in place to take down racist abuse and to stop its spread.

The prevalence of racist abuse online was brought sharply to public attention this year when Marcus Rashford, Jadon Sancho, and Bukayo Saka faced a wave of abuse on social media after they missed penalties in the Euro 2020 final. In turn, Rio Ferdinand spoke movingly to this Committee about the impact that racist abuse on social media has had on him and his family. But racism online is by no means isolated to high-profile individuals. It is a fact of life for many people of colour online, and it is always unacceptable.

The 1991 Football Offences Act made racist chanting that is 'threatening, abusive or insulting to a person' and offence within football grounds. There should also be enforcement against the same behaviour online as well.[775]

Our recommendations ensure that in addition to encompassing abuse, harassment, and threats on the grounds of race against individuals, online services will also have to address hate crimes such as stirring up racial hatred that may not currently be covered.

Platforms will have a duty to design their systems to identify, limit the spread of, and remove racist abuse quickly following a user report. Ofcom will produce a Code of Practice on system and platform design against which platforms will be held responsible for the way in which such material is recommended and amplified. Online services will be required to take steps to prevent abuse by anonymous accounts and will be required to ensure there are governance processes in place to ensure proper requests from law enforcement are responded to quickly. Where possible service providers should also share information about known offenders with the football authorities so that they can consider whether offences have been committed that would require further penalties, like the imposition of stadium banning orders. Finally, they will be required to address the risks that algorithmic recommendation tools and hashtags may amplify racist abuse.

Online fraud

Too many people are falling foul of online scams, and we want this Bill to help protect them.

85 per cent of financial scams rely on the internet in some way. Fraud is now the most reported crime in the UK. Fraudsters can approach individuals directly or pay for advertising to promote their scams. We heard, for example, how fraudsters pretend to be Martin Lewis, founder of Moneysavingexpert.com, to entice victims.

775 Football (Offences) Act 1991, section 3

The draft Bill includes fraud within "illegal content", with online services required to mitigate risks of harm and swiftly take content down if it has been reported. Paid-for advertisements are exempt.

Under our recommendations, providers will be required also to have systems and processes in place to proactively identify fraudulent content and minimise its impact on their platforms. This will include paid for adverts.

Advertising more widely is already a regulated industry, and the Bill will not step on the toes of existing regulators such as the ASA, though Ofcom will co-designate regulators to regulate parts of the Bill where their areas of responsibility sit side-by-side. The role of Ofcom will be to regulate how companies like Google or Facebook allow and promote adverts. The regulation of the advertisers themselves will remain matters for those regulators, and for the police when criminal offences are committed.

Extreme pornography

The fact that pornography depicting rape and serious violence is freely available on the internet is unacceptable and must be addressed.

We heard shocking evidence about the types of pornography which are easily available on the internet. Depictions of rape and extreme violence are freely available on many widely-used sites, often autoplaying with no age checks. The sale of such abhorrent material offline would often be a criminal offence, and it has no place online.

Under the draft Bill, user-to-user and search engines will be required to prevent children from accessing pornography. The position in relation to adults under the draft Bill is less certain. The Government has said it will designate so-called "revenge pornography" as priority illegal content. Yet, other kinds of extreme pornography may not be covered. While the dissemination of such material is illegal, it is not an offence against specific individuals. Unless the Government were to make it "priority content", it would only be regulated as part of a platform's terms of service.

Our recommendations will mean that sites should be required to prevent children from accessing all pornography, whether or not they host user-to-user content or are a search engine. The requirement on services to be safe by design will prevent people being sent unwanted recommendations of pornographic content. Online services will also be required to take down illegal, extreme pornographic material with speed once reported and take other mitigating measures. These could include warnings for users against the uploading of such content, effective governance to deal with reports, and reports to law enforcement.

Religious hatred and antisemitism

No-one should be abused for their religious faith or identity and tech companies must take steps to prevent the spread of such material and remove it from their platforms.

We heard many moving testimonies about the impact that religious hatred and antisemitism online has on individuals, families and communities. There was a record number of antisemitic incidents in the UK in May–June 2021, many of which were online.

45 per cent of religious hate crime offences in 2020–21 were against Muslims, many of which took place online. Online material can have real-world consequences—the attackers in both the Finsbury Park Mosque attack in 2017 and the 2019 Christchurch Mosque attack are believed to have been radicalised in part online.

Our recommendations ensure that the Bill encompasses abuse, harassment, and threats on the grounds of religion against individuals, as in the draft Bill, and ensure that online services will also have to address hate crimes such as stirring up religious hatred that may not currently be covered. Our report means service providers will have be required to design their systems to identify, limit the spread of, and remove such material quickly following a report. Online services will be required to take steps to prevent abuse by disposable anonymous accounts and will be required to ensure there are governance processes in place to ensure proper requests from law enforcement are responded to quickly. Finally, they will be required to address the risks that algorithmic recommendation tools and hashtags may amplify antisemitic abuse or religious hatred.

Self-harm

Promoting self-harm online should be illegal.

Self-harm, particularly amongst teenagers, is an epidemic. We were horrified to hear how videos instructing people in self-harm could come up in recommendation feeds and be promoted again and again by predatory algorithms to teenagers.

Under the draft Bill, online services would be expected to protect children from viewing content promoting or instructing viewers in self-harm. For adults, they would be required to set terms and conditions and apply them consistently.

Under our recommendations, encouraging or assisting someone to cause themselves serious physical harm would be criminalised in line with the Law Commission's recommendation. Service providers would have to protect adults and children from content or activity promoting self-harm by taking that content down quickly.

Our recommendations would also require online services to address the risk of algorithms creating "rabbit holes", in which self-harm promoting content is consistently recommended to vulnerable individuals and becomes normalised.

We recognise that great care is required in applying these recommendations to avoid barring some vulnerable individuals from social media. Ofcom will produce a Code of Practice helping platforms to identify this and other illegal and harmful content.

Zach's law

Targeting epilepsy sufferers with flashing images online is a despicable practice which should be made illegal.

Zach Eagling became the figurehead for a campaign for online safety when internet trolls targeted his fundraising tweet with flashing images. The Epilepsy Society has highlighted how people with photosensitive epilepsy are regularly targeted with flashing images which can cause a seizure. As well as the serious medical, physical, and psychological implications of such attacks, some people with epilepsy feel driven off social media as a result. This can

have a huge impact on people for whom the internet offers an opportunity to meet other people with epilepsy and build new communities.

Our recommendation would make sending flashing images to a person with epilepsy with the intention of causing a seizure a criminal offence, as recommended by the Law Commission. It would also require platforms to consider safety by design features to mitigate these risks, and Ofcom will be responsible for producing and implementing a Code of Practice on the design of systems and platforms. One example of this might be to create a user setting preventing flashing images from autoplaying or blocking them from showing at all.

Cyberflashing

Targeting people with unsolicited sexual images should be made illegal.

Cyberflashing, or sending unsolicited images of genitalia, is a problem particularly faced by young women and girls. We heard from Professor Clare McGlynn that the Ofsted review of sexual abuse in schools saw a high percentage of girls having to deal with being sent unsolicited penis images on a regular basis.

Our recommendation is that cyberflashing should become illegal once the Government has acted on the Law Commission's recommendations. This means that platforms will have the duty to mitigate and effectively manage the risk of harm to individuals from cyberflashing and remove unsolicited nude images from their platform quickly.

However, even when this bar is not met and the incident does not meet a criminal threshold, the platform will need to include as part of its risk assessment ways of identifying and mitigating the risk of harm arising from the dissemination of such material—for example, by not automatically displaying images when received.

Violence against women and girls (VAWG).

Online abuse targeted at women and girls should be prevented, illegal acts stopped, and systems should not facilitate violence.

We heard about the epidemic of online abuse against women and girls. 62 per cent of women aged 18–34 report having been a victim of online abuse and harassment. This can include stalking, abusive messages, sending unsolicited explicit images or sharing intimate pictures without consent, coercive 'sexting' and the creation and sharing of 'deepfake' pornography. Inevitably, repeated attacks increase the distress felt and harm caused by victims, though we heard how the trauma from some of those harms is often not recognised or is minimised or trivialised.

Many of these abusive acts are illegal and the list of criminal offences in this area continues to grow. Upskirting, for example, is now an offence, and if our recommendations are accepted, cyberflashing will become one. A new harm-based offence for communications could also cover sharing intimate pictures.

Our recommendation is that platforms should systems in place to identify where there is a risk of harm from all such illegal acts and put systems in place to mitigate and effectively

manage risks of harm to individuals, and to remove such material quickly when they become aware of it.

Unlike stirring up racial hatred, there is currently no criminal offence of stirring up hatred against women (though we support the Law Commission's recommendation to create one), nor is misogyny a hate crime. Under the draft Bill, such acts would be regulated as part of the service provider's terms of service and their commitment to act on misogynistic abuse online. We do not believe this should be left to platforms. Where the abuse and harassment of women and girls leads to serious psychological harm, it should be criminalised. We recommend that services should be required to identify and effectively mitigate risks caused by misogynistic abuse resulting from the way their systems and processes operate. We also recommend that they be required to address functionality that could be used in domestic abuse or VAWG, such as geolocation, and act to reduce those risks.

Incitements to violence

Any online attempt to encourage the violent overthrow of the result of a UK parliamentary election will be treated as terrorist content, and tech companies must proactively identify and remove such content.

We heard how the spread of disinformation online has been associated with extensive real-world harm, including riots, mass-killings, and harms to democracy and national security. Frances Haugen described events such as mass-murder in Myanmar and Ethiopia, and the riots at the US Capitol on 6 January as the "opening chapters" if engagement-based ranking is left unchecked and continues to amplify and concentrate extreme content that is divisive and polarising.

Under the draft Bill, any attempt to violently overthrow the UK's Parliament or elected Government would be treated as terrorism content if it threatened serious violence, damage to property or risked the health and safety of the public in trying to advance an ideological cause. Online services would be required to proactively minimise the presence of such content.

Our recommendations also tackle the design features that can lead to the spread of content advocating violence. Platforms will be required to consider safety by design measures that allow them to react quickly to emerging threats and situations, that can create friction around the sharing of illegal content and require effective moderation of groups.

Deepfake pornography

Knowingly false and threatening communications such as deepfake pornography should be made illegal and tech companies should be held responsible for reducing its spread.

The malicious use of deepfake pornography is an issue which is growing in prevalence and which can have a devastating impact on victims. In an adjournment debate on 2 December, Maria Miller MP described the decision to create and share a deepfake or a nudified image as "a highly sinister, predatory and sexualised act undertaken without the consent of the person involved."

Under our recommendations, the Law Commission's new draft offence of sending knowingly false communications likely to cause harm will be implemented concurrently with the Online Safety Act. In this way, platforms would be required to exercise their duty to mitigate for the creation of deepfake pornography.

Platforms which host pornography could reasonably be expected to identify deepfake pornography as a risk that could arise on their services and would therefore need systems and processes in place to mitigate that risk.

The offence of sending knowingly false content on a user-to-user service with malicious intent, could also apply to any known deepfake film.

Foreign interference in elections

Using anonymous accounts to influence elections from the UK or abroad should be treated as a risk by the tech companies

We heard of instances in the UK and other jurisdictions of malicious actors at home and overseas using platforms to manipulate election processes, aggravate divides and generally sow distrust. Sophie Zhang touched on her investigations into a number of co-ordinated campaigns where inauthentic Facebook accounts had been used in this way in Honduras, Brazil and Azerbaijan.

Our recommendation is that platforms which allow anonymous and pseudonymous accounts should be required to include the resulting risks as a specific category in their risk assessment on safety by design. In particular, they might be expected to cover the risk of illegal activity taking place on their platform without law enforcement being able to tie it to a perpetrator, the risk of "disposable" accounts being created for the purpose of undertaking illegal or harmful activity, and the risk of increased online abuse due to the disinhibition effect. In this way, platforms will be required to take steps to prevent abuse by disposable anonymous accounts and will be required to ensure there are governance processes in place to ensure proper requests from law enforcement are responded to quickly.

Campaign activity, including advertising, which is clearly being coordinated from overseas and in breach of election law, should also be treated as illegal content.

The Elections Bill will require online campaign material to display its promoter. Material failing to do so should be treated as illegal content and services should have systems and processes in place to mitigate the resulting risks of harm.

Appendix 2: Glossary

Age assurance. Age assurance refers to any system of age checking and estimation. The Age Verification Providers Association (AVPA) makes the distinction between "age assurance" and "age verification": age assurance is a broad term for different methods of discerning the age or age-range of an online user; age verification is a subset of that with more stringent methods and a higher level of accuracy and confidence in the age or age-range of that user.

Age verification. Age verification is a subset of age assurance, with more stringent methods and a higher level of accuracy and confidence in the age or age-range of that user.

Artificial Intelligence (AI). AI is technology which aims to replicate the problem-solving and decision-making capabilities of the human mind.

Algorithm. An algorithm is a list of rules that must be followed in a particular sequence in order to answer a question, solve a problem or perform a computation.

ASA. Advertising Standards Authority.

BBFC. British Board of Film Classification

Codes of Practice cover a range of authoritative guidance on best practice in different sectors, often by regulators. A statutory Code of Practice has the backing of an Act of Parliament and therefore carries greater weight than, for example, a self-regulating Code of Practice.

Content refers to a range of media, including text, images, memes, audio and videos.

Content moderation is the process, policy and technology used to check and curate content and activity on a service.

CSEA. Child sexual exploitation and abuse.

CMA. Competition and Markets Authority.

DCMS. The Government Department for Digital, Culture, Media and Sport.

Digital services and products. The publication of material or provision of a service through a digital medium, either free of charge or for a price.

Disinformation is factually incorrect content that is created and /or shared with the deliberate intention of misleading or deceiving audiences (in contrast to *misinformation*).

DRCF. Digital Regulation Cooperation Forum.

Duty of care. A legal "duty of care" is a term derived from the common law of negligence. It is a duty on one party not to inflict damage on another carelessly. The duty of care proposed by the 2019 White Paper and the draft Bill is a statutory duty (or series of duties) on service providers to people using their platforms.

End-to-end encryption (E2EE) is a method of secure communication between a sender and recipient which stops third parties from accessing the content of the communication.

'Third parties' includes service providers and Internet service providers. In practice, this means that encrypted services such as WhatsApp cannot view messages sent on their services.

FCA. Financial Conduct Authority.

FOS. Financial Ombudsman Service.

FSA. Financial Services Authority.

Freedom of expression. Article 10 of the European Convention on Human Rights (ECHR) states that 'everyone has the right to freedom of expression'. It includes the 'freedom to hold opinions and to receive and impart information and ideas without interference by public authority and regardless of frontiers'.

Freedom of speech. During the course of their work the Committee heard witnesses refer to freedom of expression and freedom of speech as interchangeable terms. Both free expression and free speech can be subject to restrictions.

Friction is the degree of resistance that users encounter while posting, sharing, viewing and interacting with online content or engaging in online activity. Generally, increasing friction entails adding additional steps that users must undertake before they can act online. For instance, users face very little friction if they can post content on a platform by clicking a single button. They face more friction if they need to tick a consent box and are given a warning before they can post.

Harvesting. In the context of this Report, harvesting is data harvesting, which is the automatic collection of information from online sources, such as websites or databases. Often the purpose of data harvesting is to extract information about individual users.

Harmful content is content—whether legal or illegal—with the potential to cause physical or psychological harm to a group of users (see definition of Content above).

Harmful activity is activity—whether legal or illegal—with the potential to cause people physical or psychological harm. Activity includes, but is not limited to, any behaviour which disseminate or promotes harmful content (see above).

Inferred data is data about an individual's personal attributes, such as their gender or their age, which can be inferred from their online activity. It is distinct from personal data that is explicitly provided by the user.

Microtargeted advertising is the process by which data is used to segment a set of users into smaller groups (typically based on their demographics, interest , outlooks or psychology) in order to send them tailored messages which promote something such as a product, political candidate or organisation.

Misinformation is factually incorrect content that is created and/or shared without the deliberate intention of misleading or deceiving audiences (in contrast to *disinformation*).

News publisher is any organisation that publishes news-related material which has been produced by different people. The material is subject to editorial control, as well as a

standards code. The term "recognised news publisher" is defined in clause 40 of the draft Bill.

Ofcom. Office of Communications.

Pre-legislative scrutiny is the Parliamentary process by which a draft Bill yet to be introduced into the Houses of Parliament is subject to the scrutiny of a Committee of one or both Houses, who produce a Report such as this one, containing a series of recommendations for amendment.

Priority illegal content. Refers to illegal content specified by the Secretary of State for DCMS through regulations. Service providers have to proactively minimise the presence of this content on their platforms.

Regulated activity is the posting and sharing of content, as well as ways of disseminating content or interacting with other users which are regulated by the draft Bill.

Regulated entities are services which fall into one or more of the categories regulated by the draft Bill.

Risk assessment is an assessment by the service provider of individuals who might be harmed by different categories of the regulated activity and how. It also covers what steps the service has taken to control those risks, what further steps need to be taken, by whom and by when. It is an important element of the material submitted to the regulator as part of the audit process.

Risk profile. A description of the specific risks resulting from the features of a service or group of services used to ensure regulatory requirements are proportionate and robust.

A **risk register** is a register of the risks of harm that might be encountered by as a result of certain types of content, activity or design features. Service providers will each be required to produce a risk register. The regulator might also create a register of risks likely to be encountered by different groups of users online as part of its standard-setting guidance to service providers.

Search services. A website that provides a search engine which gathers and reports on the information available on the internet in response to a query from a user.

Service provider. Throughout this Report, the term 'service provider' has been used to describe those entities that fall under the scope of the Bill. In evidence received by the Committee, other terms were used on occasion, including online providers, online services and platforms.

Social media is the collective terms for a range of digital applications that allow people to interact with each other online, as well as with businesses and organisations. Prominent examples are Facebook, TikTok, Instagram and Twitter. Different social media applications are constantly being developed.

System design. In the context of this Report, 'system design' refers to the different ways in which regulated services design their platform in order to enable the creation and dissemination of content, including their use of algorithms.

Transparency reports are the mechanism through which regulated services are required to provide data to the regulator on areas such as the categories and quantity of harmful content and activity (see definitions above) including illegal content. Transparency reports also include data on the number of requests from user for material to be taken down and the speed of the platforms' responses. The precise information to be included in such reports and the regularity of the reports will be determined by the regulator.

A **user** is a person who posts, shares, views or otherwise interacts with content published or hosted by service providers.

A **user-to-user** service provider is a business that hosts or publishes content produced by at least one person in order to be viewed, and engaged with, by at least one other person. Such services differ from commercial websites whose users are businesses.

Virality describes the rapid speed at which content can be spread online to large audiences. Content can spread in a range of ways, including recommendations, sharing and algorithmic amplification.

List of members and declarations of interest

Members

Lord Black of Brentwood	Debbie Abrahams MP
Lord Clement-Jones	Damian Collins MP
Lord Gilbert of Panteg	Darren Jones MP
Baroness Kidron	John Nicolson MP
Lord Knight of Weymouth	Dean Russell MP
Lord Stevenson of Balmacara	Suzanne Webb MP

Declarations of interests (Lords)

Lord Black of Brentwood

Directorships

Deputy Chairman (formerly Executive Director), Telegraph Media Group Limited

Non-financial interests

Director, Advertising Standards Board of Finance Ltd

President, Institute of Promotional Marketing (IPM)

Vice President, News Media Europe (http://www.newsmediaeurope.eu/)

Board Director, Regulatory Funding Company (the principal task of the company is to fund the work of the Independent Press Standards Organisation (ISPO); it also has certain responsibilities in relation to the Editors' Code of Practice Committee)

Special Adviser, Albany Associates International Limited (strategic communications solutions in fragile, conflict and post-conflict states)

Chairman (formerly Member of Council), Royal College of Music

Governor of Brentwood School, Essex

Trustee, Imperial War Museum Foundation

Board Member, WAN-IFRA (World Association of News Publishers protecting the rights of journalists and publishers around the world to operate independent media)

Chairman, Commonwealth Press Union Media Trust

Vice President, The Debating Group

President, London Press Club

President Emeritus (formerly Vice President), Printing Charity

Vice President, The Journalists' Charity

Trustee, Bloomsbury Network (charity supporting people living with HIV in London)

Master of the Guild of St Bride's, Fleet Street

Lord Clement-Jones

Directorships

Non-executive Chair, Ombudsman Services Limited (independent dispute resolution services for consumers in energy, telecoms and other industries)

Remunerated employment, office, profession etc.

Consultant (formerly London Managing Partner), DLA Piper UK LLP (International law firm)

Consultant, Big Innovation Centre, India (consultancy)

Member of KCS Europe Group Advisory Board

Non-financial interests

Advisory Board Member, Corporate Finance Faculty, Institute of Chartered Accountants in England and Wales

Advisory Council Member, Airmic

Consultant, Council of Europe's Ad hoc Committee on AI (CAHAI)

Member, Organisation for Economic Co-operation and Development (OECD) Parliamentary Group on AI

Senior Fellow, Atlantic Council's GeoTech Center

Chair of Council, Queen Mary University of London

Governor, Haileybury College (interest ceased 17 April 2021)

Advisory Board Chair, Institute for Ethical AI in Education (interest ceased 17 April 2021)

Ambassador, Barts Charity

City Ambassador, The Law Society (interest ceased 17 April 2021)

Council Member, Heart of the City

President, Ambitious About Autism (national UK charity for children with autism)

Trustee, Space for Giants

Additional interests declared at the first meeting of the Committee (27 July 2021)

Non-financial interests as Co-Chair of the All-party Parliamentary Group (APPG) on Artificial Intelligence and Vice Chair of the APPG on Digital Democracy.

Lord Gilbert of Panteg

Remunerated employment, office, profession etc.

Consultant, Finsbury (strategic communications) (interest ceased 19 April 2021)

Electoral Commissioner

Gifts, benefits and hospitality

Honorary membership of Carlton Club, given by Carlton Club (London) Limited, St James's Street, London SW1

Non-financial interests (b)

Member, Executive Committee and Chair, Audit and Risk Committee, British Group Inter-Parliamentary Union (BGIPU)

Additional interests declared at the first meeting of the Committee (27 July 2021)

Non-financial interests as a member of the All-party Parliamentary Groups on Publishing and on LGBT+ Rights and a financial interest as an Electoral Commissioner, as set out above.

Baroness Kidron

Directorships

Director, Bodyline Films Ltd

Director, Cross Street Films (Trading) Ltd

Director, Soho Angel Films Limited

Remunerated employment, office, profession etc.

Director and producer for independent film, theatre and television companies, including many major broadcasters and Streaming Video on Demand (SVOD)

Occasional fees from writing articles, making speeches and as a consultant adviser

Person with significant control of a company

Cross Street Films (Trading) Ltd

Soho Angel Films Ltd

The Imagine Workshop Ltd

Shareholdings

Bodyline Films Limited

Cross Street Films (Trading) Ltd

Soho Angel Films Limited

The Imagine Workshop Ltd

Land and property

Office accommodations in London

Sponsorship

Secretarial and other assistance in connection with the member's parliamentary duties may be undertaken by employees of and in the offices of Cross Street Films (Trading) Ltd

Non-financial interests

Member, Advisory Council, Institute for Ethics in AI, University of Oxford

Member, Executive Steering Group, Digital Makers Programme, Born in Bradford

Commissioner, UNESCO Broadband Commission for Sustainable Development

Visiting Professor in Practice, Department of Media and Communications, London School of Economics

Chair, 5Rights Foundation (charity working on digital rights for children)

Member, Council on Extended Intelligence

Member, UNICEF Artificial Intelligence and Child Rights policy guidance group (interest ceased 31 October 2020)

Trustee, Kidron Hall Charitable Trust

Founding Executive Board Member, Data Protection Foundation

Lord Knight of Weymouth

Directorships

Director, Suklaa Ltd (educational consultancy)

President, xRapid France SAS (developer of mobile diagnostic technology)

Remunerated employment, office, profession etc.

Adviser, London School of Commerce (interest ceased 31 May 2021)

Chief Education and External Officer (formerly Chief Education Adviser), TES Global Ltd (interest ceased 31 August 2020)

Person with significant control of a company

Suklaa Ltd

Whole Education Limited (interest ceased 30 June 2021)

xRapid France SAS

Shareholdings

Suklaa Ltd (educational consultancy)

TES Global Ltd (interest ceased 31 August 2020)

xRapid France SAS (developer of mobile diagnostic technology)

Non-financial interests

Director, Climate Subak CIC (community interest company)

Director, Whole Education Limited (private company limited by guarantee without share capital) and Trustee of associated charity

Visiting Professor, London Knowledge Lab, Institute of Education, University of London

Trustee and Director, E-ACT (multi-academy trust)

Trustee, Centre for Accelerating Social Technology (registered charity)

Lord Stevenson of Balmacara

Non-financial interests

Company Secretary, Aughadown Consulting Limited (private limited company owned by member's wife offering legal consulting services; post is unremunerated)

Director, Catalyst Foundation for Universal Education (a Delaware corporation, formerly Catalyst Trust for Universal Education, a UK limited company)

Trustee, The Brown Archive Trust (registered Scottish charity)

Declarations of interest (Commons)

Debbie Abrahams MP

Miscellaneous

Along with the Good Law Project and two other MPs, since 6 October 2020 I have been party to judicial review proceedings crowdfunded through the Good Law Project, 3 East Point High Street, Seal, Sevenoaks, TN15 0EG. Total costs incurred on this action: £98,914.42. £85,000 was paid by the government, leaving the remainder (£13,914.42) payable by GLP direct to Counsel. (Registered 27 October 2020; updated 26 January 2021 and 28 April 2021)

Damian Collins MP

Employment and earnings

From 1 March 2020 until further notice, Member of the Advisory Board of the Author's Licensing and Collecting Society, 1st Floor, Barnard's Inn, 86 Fetter Lane, London EC4A 1EN. I receive remuneration of £1,500 a quarter in return for a time commitment of 12 hrs per year (4 meetings with a commitment of 3 hrs each). (Registered 05 March 2020)

25 September 2020, received £550 from Mace Media Ltd, Unit 3 & 4 Croxted Mews, 286/288 Croxted Road, London SE24 9DA, for two articles for The Mace magazine. Hours: approx. 3 hrs. (Registered 20 October 2020)

Land and property portfolio: (i) value over £100,000 and/or (ii) giving rental income of over £10,000 a year: Flat in London: (i) and (ii).

Miscellaneous

From 31 March 2020 until further notice, unpaid partner of Infotagion LLP, an independent fact checking service. This involves a commitment of 10 hrs per month. (Registered 28 April 2020)

From 20 July 2020 until further notice, unpaid board member of the Center for Countering Digital Hate, a not for profit NGO headquartered in London that seeks to disrupt the architecture of online hate and misinformation. This involves a commitment of 2 hrs per month. (Registered 22 July 2020)

Additional Interests Declared at the first meeting of the Committee (27 July 2021)

Non-pecuniary interests as a member of the "Real Facebook Oversight Board" and Chair of the All-party Parliamentary Groups (APPGs) on Media Literacy and Media Freedom.

Darren Jones MP

Employment and earnings

From 3 October 2017 until 30 October 2020, contracted to provide legal consultancy, via the office of Darren Jones Ltd, to Kemp Little LLP, 138 Cheapside, London EC2V 6BJ. (Registered 19 October 2017; updated 04 November 2020)

22 October 2020, received £1,200 for work undertaken between 3 June and 14 August 2019. Hours: 8 hrs. (Registered 02 November 2020)

30 October 2020, received £262.50 for work undertaken on 28 September 2020. Hours: 1 hr 45 mins. (Registered 02 November 2020)

2. (b) Any other support not included in Category 2(a)

Name of donor: The Joseph Rowntree Reform Trust

Address of donor: The Garden House, Water End, York YO36 6WQ

Amount of donation, or nature and value if donation in kind: £19,000 to be paid in quarterly instalments to Office of Darren Jones Limited. The bulk of this money will be donated to the Labour Party to employ a part time member of staff for Labour Digital, an unincorporated organisation which I run. This support will also meet personal and project expenses incurred by Labour Digital, by me as director, and by its staff and volunteers.

Date received: 5 August 2019

Date accepted: 5 August 2019

Donor status: company, registration 357963

(Registered 12 August 2019)

Name of donor: Plexal (City) Limited

Address of donor: 6th Floor Lansdowne House, Berkeley Square, London W1J 6ER

Amount of donation, or nature and value if donation in kind: A complimentary pass to use a co-working space, valued at £3,168

Date received: 27 October 2020 to 27 October 2021

Date accepted: 27 October 2020

Donor status: company, registration 0012478

(Registered 29 October 2020)

Shareholdings: over 15% of issued share capital

Office of Darren Jones Ltd; legal consultancy. (Registered 05 July 2017)

Miscellaneous

Director, currently unpaid, of the Office of Darren Jones Ltd. (Registered 19 October 2017)

I am the director of the Henacre Charitable Fund (the "Henacre Trust") launched on 8 March 2019. This is a named trust owned and operated by the Quartet Community Foundation on my behalf. The Henacre Trust will make charitable grants to young people from Bristol North West. (Registered 11 March 2019)

From 1 May 2019, Chairman, currently unpaid, of the advisory board of the Institute of Artificial Intelligence (a not for profit organisation working with legislators from across the world on the topic of regulation of artificial intelligence). The Institute of Artificial Intelligence is the "trading name" of the Office for the Regulation of Artificial Intelligence Limited. (Registered 05 June 2019)

Additional declarations made at the first meeting of the Committee (27 July 2021)

Non-pecuniary interests as Chair of the All-party Parliamentary Groups (APPGs) on Technology and National Security and Data Poverty, Co-chair of the APPG, the

Parliamentary Information, Communications and Technology Forum, vice-chair of the APPG on Artificial Intelligence, Parliamentary Champion of the OECD Artificial Intelligence Observatory, Chair of Labour Digital, Chair of the Parliamentary Labour Party's DCMS Backbench Committee, a member of the World Economic Forum, Global AI Action Forum and Chairman of the advisory board of the Institute of Artificial Intelligence as set out above.

John Nicolson MP

Employment and earnings

Payments from News UK, 1 London Bridge Street, London SE1 9GF, for hosting a radio show:

15 July 2020, received £4,800. Hours: 24 hrs. (Registered 16 July 2020)

14 August 2020, received £2,400. Hours: 12 hrs. (Registered 17 August 2020)

11 September 2020, received £3,000. Hours: 15 hrs. (Registered 16 September 2020)

20 November 2020, received £4,800. Hours: 24 hrs. (Registered 23 November 2020)

22 January 2021, received £2,400. Hours: 12 hrs. (Registered 22 January 2021)

12 February 2021, received £3,000. Hours: 15 hrs. (Registered 12 March 2021)

15 April 2021, received £4,800. Hours: 24 hrs. (Registered 26 May 2021)

6. Land and property portfolio: (i) value over £100,000 and/or (ii) giving rental income of over £10,000 a year:

House in London: (i) and (ii). (Registered 09 January 2020)

Dean Russell MP

Employment and earnings

Payments received from EPIFNY Consulting Ltd, a business education provider, of 6 High Street, Wheathamstead AL4 8AA, for consultancy on marketing and training content and for ad-hoc client support: 22 February 2021, received £200 for services provided to the Data and

Marketing Association Ltd, DMA House, 70 Margaret Street, London W1A 8SS. Hours: 4 hrs. (Registered 24 February 2021)

Payment of £600 expected from Hult Ashridge Executive Education, Ashridge House, Berkhamsted HP4 1NS, for services provided on 12 March2021. Hours: 6 hrs. (Registered 08 April 2021)

Payment of £100 expected from the Data and Marketing Association Ltd, DMA House, 70 Margaret Street, London W1A 8SS, for services provided on 14 March 2021. Hours: 2 hrs (Registered 08 April 2021)

Payment of £200 expected from the Data and Marketing Association Ltd, DMA House, 70 Margaret Street, London W1A 8SS, for services provided on 18 May 2021. Hours: 4 hrs. (Registered 03 June 2021)

7 May 2021, received £750 from Gareth Bacon, [private address], for a painting which he commissioned. Hours: 8 hrs. (Registered 03 June 2021)

Shareholdings: over 15% of issued share capital

Win That Pitch Ltd; training consultancy (dormant). (Registered 10 January 2020)

Pretty Square Picture Company Ltd; greeting cards retailer (dormant). (Registered 10 January 2020)

Miscellaneous

On 25 September 2020, Camelot UK Lotteries Ltd made a donation of £30,000 in support of the Watford Mental Health First-Aider project, in which I am a partner. (Registered 13 October 2020)

Additional interests declared at the Committee's first meeting (27 July 2021)

Non-pecuniary interests as the Chair of the All-party Parliamentary Groups (APPGs) on Digital ID, Digital Health, and Film and Production; Co-Chair of the APPG on Mental Health; Vice Chair of the Loneliness APPG; member of the Film and Broader Screen APPG, and Small Business Ambassador for the Conversative Party.

Suzanne Webb

Employment and earnings

Councillor, Birmingham City Council, Victoria Square, Birmingham B1 1BB.

Until further notice, I receive a monthly payment of £1,435.58. Hours: 56 hrs per month. (Registered 08 January 2020)

Shareholdings: over 15% of issued share capital

Newhall Consultancy Ltd; management consultancy. (Registered 08 January 2020))

Full lists of Members' interests are recorded in the Commons Register of Members' Financial Interests: http://www.parliament.uk/business/publications/commons and the Register of Lords' Interests: https://www.parliament.uk/mps-lords-and-offices/standards-and-financial-interests/register-of-lords-interests/

Declarations of interest are also recorded in the formal minutes of the Committee.

Special advisers' interests

Jacquie Hughes

I'm a media, communications and tech sector policy regulator, currently focused on online, emerging technology, AI and XR. I was the Director of Content Policy at Ofcom, where I drew a new operating licence and performance framework for the BBC, and before that I was a Current Affairs journalist and documentary film-maker and Commissioning Editor. My interests in the broadcast policy and regulation space led me to academia, and the exploration of public value and the public interest, media convergence, regulation and standards, children's rights in the online world and ultimately AI. I'm a Board member and Fellow of 'For Humanity' a 300+ strong global community of academics, lawyers, researchers, data scientists, policymakers and auditors, committed to examining ways to mitigate the downside risk of the use of AI and automation, and a member of 'MKAI' and 'All Tech is Human'—also global voluntary communities focused on understanding the impact of artificial intelligence on society and individuals. My recent consultancy work has revolved around Intelligence Support for the Department of Health and Social Care.

Dr Charles Kriel

Founder and principal shareholder in Metrotone Media, a film and television production company making documentaries, primarily on disinformation and radicalisation. We have a film scheduled for release during the time of the Committee, titled Dis/Informed. We are also a music publisher, and do business as Kriel.Agency, primarily in capacity building in developing democracies and fragile states. Our film and television distributor is Abacus Media Rights.

Co-Founder and significant shareholder in Lightful, a digital marketing agency serving the non-profit sector

Collaborator with the Programme on Democracy and Technology, Oxford Internet Institute at the University of Oxford

Associate Fellow at the King's Centre for Strategic Communications, King's College, London

Founder and Principal shareholder in People You May Know Ltd., a special-purpose vehicle limited company, formed for the production of our last film, People You May Know. We are in the process of wrapping up this company.

Dr Bertram Vidgen:

Research Fellow in Online Harms at The Alan Turing Institute, where I lead the Online Safety team (funded by the EPSRC) and the Online Harms Observatory (part-funded by DCMS). Our team has received funding from Ofcom to conduct research into online harms. In addition, I am a Research Associate at the Oxford Internet Institute, University of Oxford, where I completed my PhD. I am co-owner of the hate speech detection machine learning task on Dynabench, which was funded by Facebook AI Research, and a co-organiser of the Workshop on Online Abuse and Harms (other co-organisers include researchers at Google and Facebook). I was previously a visiting fellow at the Open University. I am co-founder and CEO of Rewire Online Limited,

a startup building socially responsible AI for online safety. I have collaborations with industry researchers at Facebook AI Research and Google, who have funded the creation of publicly available datasets for computer science research. I have also worked on the Online Safety Data Initiative, a DCMS-funded project to improve the availability of data for innovation in online safety. My advisory work is paid via my consulting company, Vidgen Consulting Limited.

Formal Minutes

Damian Collins MP, in the Chair

Lord Black of Brentwood	Debbie Abrahams MP
Lord Clement-Jones	Darren Jones MP
Lord Gilbert of Panteg	John Nicolson MP
Baroness Kidron	Dean Russell MP
Lord Knight of Weymouth	Suzanne Webb MP
Lord Stevenson of Balmacara	

Draft Report (*Draft Online Safety Bill*), proposed by the Chair, brought up and read.

Ordered, That the Chair's draft Report be read a second time, paragraph by paragraph.

Paragraphs 1 to 469 read and agreed to.

Appendix 1 read and agreed to.

Appendix 2 read and agreed to.

Summary read and agreed to.

Resolved, That the Report be the Report of the Committee to both Houses.

Ordered, That the Chair make the Report to the House of Commons and that the Report be made to the House of Lords.

Ordered, That embargoed copies of the report be made available, in accordance with the provisions of House of Commons Standing Order No.134.

Witnesses

The following witnesses gave evidence. Transcripts can be viewed on the inquiry publications page of the Committee's website.

	Question number
Thursday 9 September 2021	
Mr Imran Ahmed, CEO and Founder at Center for Countering Digital Hate	QQ 1–51
Sanjay Bhandari, Chair at Kick it Out	
Rio Ferdinand, former professional footballer	
Edleen John, Director of International Relations and Corporate Affairs and Co-partner for Equality, Diversity and Inclusion at The Football Association	
Nancy Kelly, Chief Executive at Stonewall	
Danny Stone, MBE, Director at The Antisemitism Policy Trust	
Monday 13 September 2021	
Matt Harrison, Public and Parliamentary Affairs Manager at Royal Mencap Society	QQ 52–68
Clare Pelham, Chief Executive at Epilepsy Society	
Ian Russell, Chief Executive at Molly Rose Foundation	
Izzy Wick, Director of UK Policy at 5Rights Foundation	
Nina Jankowicz, Director of External Engagement at Alethea Group	
Thursday 23 September 2021	
William Perrin, Trustee at Carnegie Trust UK	Q 69–91
Professor Sonia Livingstone, Professor of Social Psychology at LSE Department of Media and Communications	
Dr Edina Harbinja, Senior Lecturer in Media and Privacy Law at Aston University	
Dr Professor Clare McGlynn, Professor of Law at Durham University	
Jimmy Wales, Founder at Wikipedia	
Elizabeth Denham CBE, Information Commissioner	
Stephen Bonner, Executive Director - Regulatory Futures and Innovation at The Information Commissioner's Office	
Thursday 14 October 2021	
Guillame Chaslot, Founder at Algo Transparency	QQ 92–109
Laura Edelson, Researcher at New York University	
Renee DiResta, Research Manager at Stanford Internet Observatory	

Monday 18 October 2021

Rocio Concha, Director of Policy and Advocacy and Chief Economist at Which	QQ 110–117
Martin Lewis OBE, Founder and Chair at MoneySavingExpert.com and Money and Mental Health Policy Institute	
Mark Steward, Executive Director of Enforcement and Market Oversight at Financial Conduct Authority	QQ 118–127
Michael Grenfell, Executive Director for Enforcement at Competition and Markets Authority	
Guy Parker, Chief Executive at advertising Standards Authority	
T/Commander Clinton Blackburn, National Economic Crime Coordinator at City of London Police	
Sophie Zhang, former Facebook employee	QQ 128–134

Thursday 21 October 2021

Professor Richard Wilson, Associate Dean for Faculty Development and Intellectual Life, Gladstein Chair and Professor of Anthropology and Law at University of Connecticut	QQ 135–142
Matthew D'Ancona, journalist formerly of Index on Censorship, currently Editor at Tortoise Media	
Barbora Bukovska, Senior Director, Law and Policy at Article 19	
Silkie Carlo, Director at Big Brother Watch	
Gavin Millar QC	QQ 143–147
Matt Rogerson, Director of Public Policy at Guardian Media Group, and News Media Association	
Alison Gow, President at Society of Editors	
Peter Wright, Editor Emeritus at DMG Media	
Professor Jonathan Haidt, Professor of Ethical Leadership at New York University Stern School of Business	QQ 148–153
Jim Steyer, CEO at Common Sense Media	

Monday 25 October 2021

Frances Haugen, former Facebook employee	QQ 154–192

Wednesday 27 October 2021

Maria Ressa, CEO at Rappler	QQ 193–199

Thursday 28 October

Antigone Davis, Global Head of Safety at Facebook	QQ 200–222
Chris Yiu, Director of Public Policy for Northern Europe at Facebook	
Leslie Miller, Vice President, Government Affairs and Public Policy at YouTube	QQ 223–232
Markham C. Erickson, Vice President, Government Affairs and Public Policy at Google	
Nick Pickles, Senior Director, Global Public Policy Strategy, Development and Partnerships at Twitter	QQ 233–249
Dr Theo Bertram, Director of Government Relations, Europe at TikTok	

Monday 1 November 2021

Dame Melanie Dawes, Chief Executive at Ofcom	QQ 250–273
Richard Wronka, Director for Online Harms at Ofcom	

Thursday 4 November 2021

Rt Hon. Nadine Dorries MP, Secretary of State at Department for Digital, Culture, Media and Sport	QQ 274–294
Rt Hon. Damian Hinds MP, Minister of State (Minister for Security and Borders) at Home Office	
Chris Philp MP, Parliamentary Under Secretary of State (Minister for Tech and the Digital Economy) at Department for Digital, Culture, Media and Sport	
Sarah Connolly, Director, Security and Online Harms at Department for Digital, Culture, Media and Sport	

Published written evidence

The following written evidence was received and can be viewed on the inquiry publications page of the Committee's website.

OSB numbers are generated by the evidence processing system and so may not be complete.

1 5Rights Foundation (OSB0096), (OSB0206), (OSB0205)

2 5 Sports: The Football Association, England and Wales Cricket Board, Rugby Football Union, Rugby Football League and Lawn Tennis Association, The FA (OSB0111)

3 Action for Primates, Lady Freethinker (OSB0139)

4 Ada Lovelace Institute (OSB0101)

5 Advertising Standards Authority (OSB0213)

6 Advisory Committee For Scotland (OSB0067)

7 The Age Verification Providers Association (OSB0122)

8 Alliance for Intellectual Property (OSB0016)

9 All-Party Parliamentary Group on Commercial Sexual Exploitation (OSB0037)

10 Anti-Defamation League (OSB0030)

11 Antisemitism Policy Trust (OSB0005)

12 APPG Coalition (OSB0202)

13 The Arise Foundation (OSB0198)

14 Association of British Insurers (OSB0079)

15 Aviva Plc (OSB0042)

16 Barclays Bank (OSB0106)

17 Dr Mikolaj Barczentewicz (Senior Lecturer in Law at University of Surrey) (OSB0152)

18 Barnardo's (OSB0017)

19 Baroness Floella Benjamin, DBE (OSB0161)

20 BBC (OSB0074)

21 BBFC (OSB0006)

22 Big Brother Watch (OSB0136)

23 Board of Deputies of British Jews (OSB0043)

24 Dr Emma Briant (Research Associate at Bard College) (OSB0155)

25 British & Irish Law, Education & Technology Association (OSB0073)

26 British Horseracing Authority (OSB0061)

27 British Retail Consortium (OSB0087)

28 Bumble Inc. (OSB0055)

29 Mr Rae Burdon (Director at Reform Political Advertising) (OSB0199), (OSB0226)

30 Andrew Campling, Director of 419 Consulting Ltd. (OSB0172)

31 Care (OSB0085)

32 Carnegie UK (OSB0095)

33 Mr John Carr (Secretary of the Children's Charities' Coalition on Internet Safety) (OSB0167), (OSB0216)

34 Catch 22 (OSB0195)

35 CEASE (Centre to End All Sexual Exploitation) (OSB0104)

36 Centenary Action Group, Glitch, Antisemitism Policy Trust, Stonewall, Women's Aid, Compassion in Politics, End Violence Against Women Coalition, Imkaan, Inclusion London, The Traveller Movement (OSB0047)

37 Center for Countering Digital Hate (OSB0009), (OSB0227)

38 Sarah Champion MP (OSB0208)

39 The Children's Society (OSB0245)

40 CIFAS (OSB0051)

41 Clean up the Internet (OSB0026), (OSB0238), (OSB0239)

42 Cloudflare (OSB0091)

43 Coadec (OSB0029)

44 Coalition for Content Provenance and Authenticity (OSB0240)

45 Rachel Coldicutt (OSB0153)

46 Common Sense (OSB0018)

47 Compassion in Politics (OSB0050)

48 Competition and Markets Authority (OSB0160)

49 Confederation of British Industry (CBI) (OSB0186)

50 Conscious Advertising Network (OSB0180)

51 COST Action CA16207 - European Network for Problematic Usage of the Internet (OSB0038)

52 The Lord Bishop of Oxford, Rt Revd Dr Steven Croft (OSB0212)

53 Crown Prosecution Service (OSB0179)

54 Paul Davis (Director of Fraud at TSB Bank Plc) (OSB0164)

55 Defenddigitalme (OSB0188)

56 Demos (OSB0159)

57 Department of Digital, Culture, Media and Sport, and The Home Office (OSB0011), (OSB0243), (OSB0248)

58 Digital Identity Net U.K. Ltd (OSB0143)

59 Dignify (OSB0196)

60 Direct Line Group (OSB0082)

61 DMG Media (OSB0133), (OSB0220)

62 Mr Downing (OSB0156)

63 Electrical Safety First (OSB0100)

64 Elizabeth Kanter (Director of Government Relations at TikTok) (OSB0219)

65 End the Virus of Racism (OSB0173)

66 Engine Advocacy (OSB0137)

67 Epilepsy Society (OSB0008)

68 Facebook (Meta) (OSB0147)

69 Financial Conduct Authority (OSB0044), (OSB0229)

70 The Football Association, Kick It Out (OSB0234)

71 The Football Association, The Premier League, EFL, Kick It Out (OSB0007)

72 For Humanity (OSB0230)

73 Full Fact (OSB0056)

74 Gambling Related Harm All-Party Parliamentary Group (OSB0151)

75 Girlguiding (OSB0081)

76 Glassdoor (OSB0033)

77 Glitch (OSB0097)

78 Global Action Plan (OSB0027)

79 Global Action Plan, on behalf of the End Surveillance Advertising to Kids coalition, The Mission and Public Affairs Council of the Church of England, Global Witness, New Economics Foundation, Foxglove Legal, Fairplay, 5Rights Foundation, Andrew Simms, New Weather Institute, Dr Elly Hanson, Avaaz (OSB0150)

80 Global Partners Digital (OSB0194)

81 Google (OSB0175)

82 Google UK Limited (OSB0218)

83 Guardian Media Group (OSB0171)

84 Gumtree UK (OSB0185)

85 Hacked Off (OSB0041)

86 Dr Elly Hanson (Clinical Psychologist) (OSB0078)

87 Dr Edina Harbinja (Senior lecturer in law at Aston University, Aston Law School) (OSB0145)

88 Hargreaves Lansdown (OSB0197)

89 Henry Jackson Society (OSB0028)

90 Dame Margaret Hodge (Member of Parliament for Barking and Dagenham at House of Commons) (OSB0201)

91 HOPE not hate (OSB0048)

92 IMPRESS (OSB0092)

93 The Independent Media Association (OSB0064)

94 Independent Schools Council (OSB0187)

95 Index on Censorship (OSB0249)

96 Information Commissioner's Office (OSB0062), (OSB0210), (OSB0211)

97 Innovate Finance (OSB0116)

98 International Justice Mission (OSB0025)

99 Internet Association (OSB0132)

100 Internet Matters (OSB0103)

101 The Internet Service Provider Association (ISPA) (OSB0059)

102 Internet Watch Foundation (OSB0110)

103 The Investment Association (OSB0162)

104 ITV (OSB0204)

105 Ms. Daphne Keller (Director, Program on Platform Regulation at Stanford Cyber Policy Center) (OSB0057)

106 Keoghs LLP (OSB0003)

107 Sara Khan (Former Lead Commissioner at Commission for Countering Extremism); Sir Mark Rowley (Former Assistant Commissioner at Metropolitan Police Service) (OSB0034)

108 Legal to Say, Legal to Type (OSB0049)

109 The LEGO Group (OSB0146)

110 LGBT Foundation (OSB0045), (OSB0046), (OSB0191)

111 Lloyds Banking Group plc (OSB0135)

112 Local Government Association (LGA) (OSB0178)

113 Logically (OSB0094)

114 LSE Department of Media and Communications (OSB0001), (OSB0236), (OSB0247)

115 M&G PLC (OSB0176)

116 Match Group (OSB0053)

117 Minderoo, Centre for Technology and Democracy (OSB0237)

118 Professor Clare McGlynn (Professor of Law at Durham University) (OSB0014), (OSB0244)

119 medConfidential (OSB0010)

120 Mencap (OSB0075)

121 Meta (Facebook) (OSB0224)

122 Microsoft (OSB0076)

123 Gavin Millar QC (OSB0221)

124 Mrs Gina Miller (OSB0112)

125 Mobile UK (OSB0168)

126 Molly Rose Foundation (OSB0149), (OSB0233)

127 Money and Mental Health Policy Institute (OSB0036)

128 MoneySavingExpert (OSB0113)

129 Dr Martin Moore (Senior Lecturer at King's College London) (OSB0063)

130 Mumsnet (OSB0031)

131 The Naked Truth Project (OSB0023)

132 The National Union of journalists (OSB0166)

133 Mr Hadley Newman (OSB0125)

134 News Media Association (OSB0107)

135 NSPCC (OSB0109), (OSB0228)

136 Ofcom (OSB0021), (OSB0223)

137 Office of the Children's Commissioner (OSB0019)

138 Office of the City Remembrancer, City of London Corporation (OSB0148)

139 OnlyFans (OSB0217)

140 Open Rights Group (OSB0118)

141 Dr Amy Orben (College Research Fellow at Emmanuel College, University of Cambridge) (OSB0131)

142 Parentkind (OSB0207)

143 Parent Zone (OSB0124), (OSB0250)

144 Patreon Inc. (OSB0123)

145 Mrs Friese Peach (OSB0002)

146 Peers for Gambling Reform (OSB0114)

147 Professor Andy Phippen (Professor of Digital Rights at Bournemouth University) (OSB0121)

148 PIMFA (OSB0102)

149 Polis Analysis (OSB0108)

150 Premier Christian Communications Ltd (OSB0093)

151 Professional Players Association (OSB0154)

152 Professor Andrew Przybylski (Associate Professor, Senior Research Fellow at University of Oxford) (OSB0193)

153 Publishers Association (OSB0099)

154 Quilter (OSB0024)

155 Reddit, Inc. (OSB0058)

156 Refuge (OSB0084)

157 Reset (OSB0138), (OSB0203), (OSB0232)

158 Revolut (OSB0117)

159 Professor Jacob Rowbottom (OSB0126)

160 RSA (OSB0070)

161 Samaritans (OSB0182), (OSB0251)

162 Schillings International LLP (OSB0183)

163 Shout Out UK (OSB0128)

164 Nathan Silver (OSB0013)

165 Siobhan Baillie MP – Member for Stroud (OSB0242)

166 Sky (OSB0165)

167 Sky, BT, Channel 4, COBA, ITV, NBC Universal, TalkTalk, Virgin Media O2, Warner Media (OSB0177)

168 Snap Inc. (OSB0012)

169 Dr Francesca Sobande (Lecturer in Digital Media Studies at Cardiff University) (OSB0144)

170 Somerset Bridge Group Ltd. (OSB0004)

171 Sport and Recreation Alliance (OSB0090)

172 StepChange Debt Charity (OSB0222)

173 Stonewall (OSB0083)

174 Rt Hon. Mel Stride MP (Chair at House of Commons Treasury Select Committee) (OSB0209)

175 SumOfUs (OSB0068)

176 Suzy Lamplugh Trust (OSB0246)

177 SWGfL (OSB0054)

178 Mr Mark Taber (Consumer Finance Expert, Campaigner & Media Contributor) (OSB0077)

179 TalkTalk (OSB0200)

180 Professor Damian Tambini (Distinguished Policy Fellow and Associate Professor at London School of Economics and Political Science) (OSB0066)

181 Tech Against Terrorism (OSB0052)

182 techUK (OSB0098)

183 TikTok(OSB0181)

184 Transfrom Hospital Group Ltd (OSB0040)

185 Twitter (OSB0072), (OSB0225)

186 UK Finance (OSB0088)

187 UK Interactive Entertainment (OSB0080)

188 UKRI Trustworthy Autonomous Systems Hub (OSB0060)

189 Virgin Media O2 (OSB0127)

190 Vodafone UK (OSB0015)

191 Dr Dimitris Xenos (Lecturer in Law at Cardiff Metropolitan University) (OSB0157)

192 Mr Richard Watts (OSB0231)

193 WebGroup Czech Republic, a.s., NKL Associates s.r.o. (OSB0142)

194 Which? (OSB0115)

195 Who Targets Me (OSB0086)

196 Wikimedia UK (OSB0169)

197 Women in Sport (OSB0105)

198 Work and Pensions Committee (OSB0020)

199 World Parrot Trust (OSB0215)

200 Yoti (OSB0130)

201 Young Epilepsy (OSB0140)

202 Sophie Zhang (OSB0214)

203 Zoom Video Communication (OSB0174)