Written evidence submitted by Full Fact

Draft Online Safety Bill Joint Committee

Summary

- The scope of the draft Bill has been reduced from earlier proposals so that the problems of misinformation and disinformation will not be effectively addressed. This should be revisited so the Bill includes misinformation and tackles the harms to our society and democracy – as well as the harms to individuals – as set out in the government's counter-disinformation strategy.
- There are problems with the definition and process for determining what counts as "content that is harmful" and instead the legislation should clearly identify specific harms it seeks to address and the remedies required on them, and do so with far more democratic oversight.
- The Bill should provide for promoting authoritative information and sources as well as a requirement on Category 1 internet companies (as with radio and television) to include news content so that users are exposed to news as part of a healthy society.
- The draft Bill is not yet one that would strengthen freedom of expression by providing open democratic transparent oversight of both commercial and political decisions which seek to limit ordinary internet users' freedom of expression: safeguards are needed on key decision makers and the risks they pose to freedom of expression.
- Censorship-by-proxy is a problem that needs to be addressed in the Bill, including the addition of a reporting requirement for the government to publish details of all efforts it makes to influence internet company content moderation decisions.
- The Bill should ensure independent scrutiny and testing of the algorithms of in-scope internet companies which restrict what people see and share online. This would enable the regulator to oversee a system that verifies and certifies algorithms as being low risk in relation to direct harm to individuals, discrimination or as having no disproportionate impact on freedom of expression.
- The definitions of democratic content in the draft Bill are not sufficiently clear and must be revisited to avoid serious consequences, as should the proposals around journalism and news publishers which create a loophole for harmful disinformation campaigns and sham news sites to cause harm as well as other problems.
- The Bill should improve democracy and address harms to democracy including protecting against harmful misinformation in elections, and the Government should establish a UK Critical Election Incident Public Protocol to secure public confidence in how elections are protected, given they are vulnerable to interference.
- The powers given to the Secretary of State in the draft Bill are too extensive and need to be narrowed to ensure democratic accountability as well as separation of roles, and the role of Parliament should be strengthened to increase its oversight.
- The government and Parliament's ambition regarding online media literacy should be strengthened from how it is presently set out in the draft Bill, as a key part of citizen-supporting methods of tackling the problems in our information environment.

- The role of the public in tackling harmful misinformation and informing policy in this area should be greatly enhanced under the Online Safety Bill, including citizen voice and participation.

About Full Fact

1. Full Fact fights bad information. We're a team of independent fact checkers, technologists, researchers, and policy specialists who find, expose and counter the harm it does.
2. Bad information damages public debate, risks public health, and erodes public trust. So we tackle it in four ways. We check claims made by politicians, public institutions, in the media and online and we ask people to correct the record where possible to reduce the spread of specific claims. We campaign for systems changes to help make bad information rarer and less harmful, and we advocate for higher standards.
3. Full Fact is a registered charity. We're funded by individual donations, charitable trusts, and by other funders. We receive funding from both Facebook and Google. Details of our funding can be found on our website.[1]

4. Full Fact has long called for the government to take action in this area and the online misinformation that has come with the pandemic has made this need even more evident. We cannot go on relying on the internet companies to make decisions without independent scrutiny and transparency. We welcome that the process of pre-legislative scrutiny of the draft Online Safety Bill is finally going ahead, and that the Joint Committee on the Draft Online Safety Bill is undertaking an inquiry: good legislation and regulation could make a significant difference in tackling dangerous online misinformation.
5. Full Fact's expertise covers online misinformation and public debate. Some areas of the Online Safety Bill including specific harms falls beyond our areas of expertise (e.g., illegal harms around protecting children and relating to terrorism) and this submission reflects this.
6. Full Fact is a member of Ofcom's Making Sense of Media Advisory Panel which brings together experts to debate and inform the development of Ofcom's media literacy research and policy work.
7. Full Fact is a member organisation of the Counter-Disinformation Policy Forum, which is convened by the Department for Digital, Culture, Media & Sport (DCMS) and brings together stakeholders from internet platforms, civil society academia and government to limit the spread and harmful effects of misinformation and disinformation.
8. This submission is in addition to a Full Fact written submission on the draft Online Safety Bill to the DCMS Sub-Committee on Online Harms and Disinformation earlier in September 2021.

---

[1] https://fullfact.org/about/funding/

Objectives

Will the proposed legislation effectively deliver the policy aim of making the UK the safest place to be online? Does the draft Bill make adequate provisions for people who are more likely to experience harm online?

9. No. In its present form the draft Bill cannot be seen as legislation that will make the UK the safest place to be online, for a number of reasons.

10. Bad information ruins lives. We have seen this – and continue to see this – in the pandemic, with Covid-19 misinformation bringing very significant harm undermining safety to life and long-term health, and devastating families across the country.

11. We have seen, within and without the pandemic context, ordinary people being silenced by excessive action and overreaction by internet companies and/or government. The draft Bill, if not changed, could perpetuate this dynamic.

12. The omission of tackling harmful misinformation in the Online Safety Bill itself should be addressed: not putting this on the face of the Bill will not only weaken attempts to address harms but will also undermine protecting freedom of expression.

13. The draft Bill does not make adequate provisions for people who are more likely to experience harm online: it does not read like a Bill that is based on a goal of understanding harms as experienced by people every day and setting out a basis by which proportionate and effective responses are taken forward to address those harms.

14. Whilst there is much in the draft Online Safety Bill that will be very positive to put into place, the Bill should be reoriented around developing proportionate responses to clearly identified harms with far more democratic oversight.

Does the Bill deliver the intention to focus on systems and processes rather than content, and is this an effective approach for moderating content? What role do you see for e.g., safety by design, algorithmic recommendations, minimum standards, default settings?

15. A systems focus is to be preferred, although there are many questions to be addressed on what the consequences may be at content level in the present draft Bill. This is a key reason, for example, why Full Fact is calling for independent scrutiny of the algorithms of in-scope internet companies (see below).

16. In seeking to address harms there is no point in seeing moderation of potentially problematic content as any kind of keystone as a whole ecosystem approach is required. Good information matters and should be addressed in the Bill.

Promoting authoritative information and sources

17. Addressing online harms effectively is often not just about addressing bad information directly, it is also predicated on the supply and dissemination of good information. Despite efforts from some companies such as YouTube to promote 'authoritative news', high-quality news content and robust public information still struggles to get anywhere near the same level of traction as misinformation, as multiple studies have shown.[2] However, during the pandemic internet companies set themselves a high bar for promoting authoritative information, with most pointing users towards WHO and public health sources in well-designed news feed panels and redirecting users to robust sources in search results. The pandemic has shown that identifying authoritative news and information sources can be relatively straightforward, but promotion of robust, non-partisan information is not the norm outside of pandemic-related public health (and elections in some countries).

18. Parliament has previously recognised the need for news as part of a healthy society – it is a required part of radio and television output, for example. As the relative share of attention in legacy media declines, and as audiences fragment, we recognise the erosion of the shared reality that comes from shared access to news. That has consequences for our democracy and society more generally. We believe that Parliament could consider whether a similar requirement to include news content should now be applied to Category 1 internet companies so that internet users are exposed to news in a similar way that broadcast audiences are. It is far better to pre-empt problems of misinformation by making good information readily available than to respond later with measures that restrict freedom of expression, and this may be one way of shifting the balance towards proportionate measures.

Does the proposed legislation represent a threat to freedom of expression, or are the protections for freedom of expression provided in the draft Bill sufficient?

19. We respectfully suggest that this is the wrong question. The status quo is a threat to freedom of expression. We need protections against what the internet companies and government are currently doing of their own accord, and not just a balancing act around the contents of this Bill.

20. A good Online Safety Bill would actually strengthen freedom of expression by providing open democratic transparent oversight of both commercial and political decisions which seek to limit ordinary internet users' freedom of expression.

21. Unfortunately, this is not yet that bill. We will explain the three key groups of powerful decision makers, the risks they pose to freedom of expression, and how the Online Safety Bill should address them in our view.

---

[2] https://www.socialmediatoday.com/news/new-study-shows-that-misinformation-sees-significantly-more-engagement-than/555286/; https://www.wired.com/story/right-wing-fake-news-more-engagement-facebook/

22. The three decision makers are governments, internet companies, and content moderation algorithms. All are actively restricting freedom of expression and shaping public debate online today and all need legal safeguards on their activities, including targeted transparency measures, as outlined below.

23. Bad information ruins lives, and Full Fact believes that we need proportionate action from all three actors above to address clearly identified harms from bad information. But action on specific pieces of content should take freedom of expression as its starting point. We do not believe that an internet company, or anybody else, should necessarily take action just because somebody says something which isn't true. Freedom of expression includes the freedom to be wrong. When action is taken on specific content, the starting point should be giving users information from non-partisan, authoritative sources that helps them make up their own minds about whether to trust what they are seeing, in preference to more restrictive measures. That is not the status quo.

24. Governments can seek to legislate to stop people saying or sharing certain things. However, during the pandemic the UK government has leaned on the internet companies instead, pursuing censorship-by-proxy with little to no political or legal scrutiny. One example is that the government "summoned" internet companies to tell them to remove certain content about 5G mobile networks after harassment of telecoms workers and attacks on facilities. Whether that was a proportionate response deserves debate. Full Fact had warned about the danger of 5G misinformation the year before and called for the free speech response of better public health information. This warning was not heeded and after the situation escalated, perhaps avoidably, the government turned to censorship-by-proxy through the internet companies.

25. We believe that the Bill needs to include a provision to make government interventions in content moderation transparent.

26. Internet companies can overreach on their own initiative too. One admitted example is Facebook's decision to remove posts discussing whether Covid-19 may have come from a lab in response to a significant amount of misinformation around this. That decision was later reversed so the company would "no longer remove the claim that COVID-19 is man-made" in response to the news that the US government was evaluating that possibility.[3] The frightening thing is that this is likely to be the tip of the iceberg of information the internet companies choose to restrict. While those companies hide information about the details and trade-offs of their response and resist independent evaluation, no parliament should rest easy. The only way to protect freedom of expression from the internet companies themselves is to legislate for oversight of their content moderation choices.

27. At the same time, some internet companies have adopted a range of ways of tackling harmful misinformation that leave people free to say and share what they want, including read-before-you-share prompts, Covid-19 information centres, and highlighting independent fact checking. We believe that, in principle, these kinds of responses are preferable to those that restrict freedom of expression and likely to be proportionate in a wider range of circumstances.

---

[3] https://www.politico.com/news/2021/05/26/facebook-ban-covid-man-made-491053

28. So, we welcome the fact that Clause 12 sets out the obligation to carry out an assessment of the impact policies have on freedom of expression (as well as privacy). This is a needed requirement although it is not clear how detailed such assessments should be, whether they should be only retrospective or also forward looking, and also whether it is needed for every existing and new policy. It is welcome that such impact assessments must be kept up to date and be published (12.4.a and b) and that how the company intends to protect users' right to freedom of expression within the law in response to an impact assessment is to be made publicly available (12.5.a).

29. However, there appears to be a strong element of in-scope companies 'marking their own homework' when it comes to adhering to what the draft Bill requires, and this includes how freedom of expression features. Such a system will not work without independent quality control. There needs to be a stronger emphasis on Ofcom being able to set out directions for impact assessments and steps required and a duty to comply with any direction from Ofcom. We also believe that the Bill could set out the requirement for proportionate responses more clearly and enforceably.

30. Finally, content moderation algorithms are of course the responsibility of those who design and operate them. But the reality is that nobody can fully understand their effects when deployed at internet scale, especially given the internet companies' opaque approach. We have seen internet companies deploying algorithms that filter out posts referring to the place "Plymouth Hoe"[4] ; taking down posts from a police force warning about Covid-19 scams[5]; and even banning the internet company's own page after mistaking it for an Australian news site[6]. This is a new challenge for regulation, and impact assessments written by the companies themselves are not adequate to address it.

31. Content moderation algorithms can do real good if they work well, and if they malfunction, they can cause real harm. In this they are like many other safety-critical technologies. However, unlike many safety-critical technologies, the safety consequences of deploying a certain content moderation algorithm are not always obvious. How safe an aeroplane is will ultimately be visible for all to see, despite all the expertise that goes into its engineering. A qualified person can and must test whether an electrical system is safe. The effects of content moderation algorithms are far harder to understand, but their design is subject to no external scrutiny at all. It is no coincidence that some algorithms deployed by internet companies to their billions of users have been shown to directly discriminate on the grounds of race[7].

32. The Online Safety Bill should seek to address these unintended effects by requiring independent testing of safety-critical content moderation algorithms

---

[4] https://www.itv.com/news/westcountry/2021-01-27/facebook-removes-posts-featuring-plymouth-hoe-for-being-offensive

[5] We are aware of this directly from the police force concerned.

[6] Kevin Nguyen of ABC News compiled a list of many pages blocked when Facebook attempted to filter out news content in Australia. Along with Facebook's own page, it included a state government, a domestic charity, and a Fire Department https://twitter.com/cog_ink/status/1362173998696566790

[7] https://www.vox.com/recode/2020/2/18/21121286/algorithms-bias-discrimination-facial-recognition-transparency

(in many ways, just as other safety-critical industries require independent inspections). There would be challenges involved in putting this in place and institutionalising it under the auspices of Ofcom as regulator, keeping in mind factors such as not creating barriers to entry to the industry, or making it excessively costly in operation or to fix problems identified. In our view, these challenges need to be taken on.

33. One of the novel challenges for the Online Safety Bill is that it is no longer just powerful people who can effectively restrict someone's freedom of expression. Online abuse tactics can have much the same effect on their targets.

34. The internet has been a tremendous force for democratising public debate. It has made information more accessible, created a more level playing field, and helped people get together to pursue causes they believe in. But the choices internet companies make can powerfully enhance our ability to impart and receive information, or they can infringe on our freedom of expression.

35. Internet companies make the efforts they do to tackle abuse out of sheer necessity to make their product usable as well as because of external pressure. But they involve the kinds of trade-offs between competing rights that in a democracy should properly be made by elected politicians, or within a democratically accountable framework.

36. Ultimately, we do not understand the argument that freedom of expression would be better protected in the UK if Parliament steps out of the picture and continues to leave these unscrutinised powers in the hands of overseas internet companies.

Content in Scope

The draft Bill specifically places a duty on providers to protect democratic content, and content of journalistic importance. What is your view of these measures and their likely effectiveness?

37. The definition of "content of democratic importance" is not sufficiently clear in relation to what "is or appears to be specifically intended to contribute to democratic political debate" and this has the potential for serious and unintended consequences. Whilst the explanatory notes offer two kinds of examples this raises questions as to what would or would not be such content beyond this. The definition of "a live political issue" is also not clear and neither is it clear who determines that definition.

38. We declare an interest in that Full Fact is a recognised news publisher within the meaning of the draft Bill.

39. There are two main problems with the government's proposed exemption for news publishers and journalists. First, the underlying principle is flawed. If the draft Bill really requires special protections to make journalism possible under its rules, then its restrictions on ordinary internet users go too far. Journalists should not need or have privileged freedom of expression compared to their audiences (as distinct from privileged access to information or protection from interference with newsgathering, which they do sometimes need, and the law provides for).

40. Secondly, it will not work on its own terms and accidentally creates a loophole so big that some of the most reckless and harmful deliberate disinformation campaigns could sail through it. The criteria for a news publisher being a recognised news publisher essentially boils down to setting a standards code, handling complaints, and having an identifiable publisher. It would be trivial for a malicious disinformation campaign to comply with those requirements. During the pandemic at Full Fact we have seen the harm that can be caused by sites purporting to be legitimate news and these sham sites cannot be excluded from the proposed regime.

41. Well-meaning but poorly drafted measures could prove detrimental to individuals and society for years to come. Full Fact believes that democratic speech and journalistic content are among a number of such areas that should be very carefully considered before the final Bill is put to Parliament.

Earlier proposals included content such as misinformation/disinformation that could lead to societal harm in scope of the Bill. These types of content have since been removed. What do you think of this decision?

42. It is important to distinguish between legislating about harms and legislating about content. There are many ways the Bill can help reduce societal harms without restricting content.

43. We strongly believe that the harms caused by misinformation and disinformation should be in the scope of the Bill.

44. The effect of not including this type of content explicitly is to reduce the scope of the draft Bill, so that the problems of misinformation and disinformation will not be effectively addressed, despite the government saying that these "have the potential to cause significant harm to both individuals and society… to influence elections, stoke racial divisions and abuse, and incite violence and rioting."[8]

45. The government's own RESIST counter-disinformation strategy states that: "When the information environment is deliberately confused this can: threaten public safety; fracture community cohesion; reduce trust in institutions and the media; undermine public acceptance of science's role in informing policy development and implementation; damage our economic prosperity and our global influence; and undermine the integrity of government, the constitution and our democratic processes."[9] The draft Bill's narrow focus on harm to individuals will fail to address all of these harms to our society and our democracy.

46. The shift from what was set out in the White Paper has gone alongside a shift from developing proportionate responses to clearly identified harms, to an approach with far less democratic oversight than we would expect, and far

---

[8] 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
[9] RESIST Counter Disinformation Toolkit, 2019 https://gcs.civilservice.gov.uk/publications/resist-counter-disinformation-toolkit/

more emphasis on the control of content that ordinary internet users can see and share.

47. The failure to establish an open, democratic, transparent process for tackling the harms caused by misinformation and disinformation leaves an oversight vacuum and invites overreach as well as risking over-moderation by platforms. The draft Bill's proposal of an advisory committee on misinformation appears to be all the draft Bill presently does in this area and this is utterly inadequate.

48. Meanwhile, it is clear that the government has come to think that its role includes identifying particular bits of legal content on the internet that should not remain online, and pressuring internet companies to remove them.

49. This government's enthusiasm for censorship-by-proxy has been a marked feature of its response to the pandemic. A government press release stated that "Up to 70 incidents a week, often false narratives containing multiple misleading claims, are being identified and resolved."[10] It did not define 'resolved'. A subsequent story briefed by the government to the BBC said: "The culture secretary is to order social media companies to be more aggressive in their response to conspiracy theories linking 5G networks to the coronavirus pandemic."[11] Of course we accept that measures taken with good intentions during an emergency are never likely to be perfect. But instead of establishing open, democratic, transparent methods for responding to harmful false information in future, this draft Bill's silence on misinformation and disinformation risks locking in censorship-by-proxy as the new normal and relying on the good faith and judgement of staff currently in post, rather than limiting potential overreach by less scrupulous future decision-makers.

50. We would be glad to discuss with the committee how the Bill could promote proportionate and open ways of tackling these serious harms.

Democratic harms and election integrity

51. On democratic integrity, the Government appears resistant to tackling societal or collective harms without a definitive clear line to direct individual harm. Yet the government is aware such threats and risks are real. The Government has indicated that foreign state disinformation campaigns during UK elections will be out of scope of the Online Safety Bill.

52. We would urge revisiting this area in relation to elections, and not just its relationship with harms to individuals. We do not believe it is right that the national security and other implications of disinformation campaigns during UK elections are out of the scope of the Bill. Indeed, the Prime Minister assured Parliament earlier this year that the Online Safety Bill that will be presented to the House this year will contain sufficient powers to tackle collective online harms, including threats to our democracy[12].

---

[10] Government cracks down on spread of false coronavirus information online, Mar 2020 https://www.gov.uk/government/news/government-cracks-down-on-spread-of-false-coronavirus-information-online

[11] 'Coronavirus: Tech firms summoned over 'crackpot' 5G conspiracies' BBC News, Apr 2020 https://www.bbc.co.uk/news/technology-52172570

[12] Integrated Review debate, House of Commons, 16 March 2021 https://hansard.parliament.uk/commons/2021-03-16/debates/52D67D49-A516-4598-AC69-

53. The Bill should be strengthened and amended to improve democracy and address harms to democracy, including protecting against harmful misinformation in elections.

The need for a UK Critical Election Incident Public Protocol

54. There may come a time during an election when the public needs to be warned about a specific threat identified by the security services, but at the moment the decision would be up to the government of the day, which would be put in a difficult position and is likely to be seen as conflicted.

55. In Canada, this problem has been solved by setting out a public protocol—the Critical Election Incident Public Protocol (CEIPP)—for handling such situations and to depoliticise a key area where a general election may be vulnerable to interference and requires a solution to protect and defend electoral systems and processes.

56. The UK Government should develop and publish a protocol for alerting the public to incidents or campaigns that threaten the UK's ability to have a free and fair election that is independent of elected politicians. Ideally, the Elections Bill should include a provision requiring such a protocol to be agreed. It is important that The Elections Bill does, as the government has stated, work alongside measures in the Online Safety Bill and Counter-State Threats Bill 'to protect our globally respected UK democracy from evolving threats'. Coherence is required across different regulations and associated practices, including on known and foreseeable risks.

57. The draft Online Safety Bill contains a provision (under Clause 112 Secretary of State directions in special circumstances) enabling the Secretary of State to give Ofcom directions when they consider there is a threat to the health or safety of the public, or to national security. This clause, which requires significant attention in parliamentary scrutiny and wider debate on the draft Bill, does not explicitly mention elections.

58. The present draft Bill clause 112 is largely focused on directing Ofcom to prioritise action to respond to such a specific threat through its media literacy functions and requiring certain internet companies to publicly report on what they are doing to respond to such a threat.

59. Given the Online Safety Bill is about Ofcom and regulation to prevent harm emerging from internet companies' platforms in its scope, if democratic harms were folded into the Bill, as we believe they should be, this could then interlock with other regulators and actors, to address harmful misinformation and disinformation or other incidents threatening free and fair elections. A sensible precautionary provision is required to ensure that the public can be informed of any threats through a predictable and trusted process which can be effectively mitigated through public information.

60. The reality at the moment is that the decision about whether to warn the UK public of a threat to our elections is as likely to be taken in California as Westminster. An election is possible at any moment. If conducted under

current rules or indeed, as the present Elections Bill and draft Online Safety Bill envisages, it will be vulnerable to a serious incident with no protocol in place.

What would be a suitable threshold for significant physical or psychological harm, and what would be a suitable way for service providers to determine whether this threshold had been met?

61. We see three problems with the definition and process for determining what counts as "content that is harmful" set out in clauses 45 and 46.

62. Underlying them all is the central mistake made in the draft Bill of trying to define harm in the abstract and then hand off to Ofcom and internet companies the details of how to identify and tackle it. What is needed instead is legislation that clearly identifies the specific harms it seeks to address and the remedies it wishes to require, as the draft Bill does try to do in respect of specific kinds of illegal content. That legislation would of course need to be regularly revisited by Parliament. The idea of legislating for online safety once and for all is hubristic and ignores the constantly evolving nature of content, etiquette and design of networks online. It would be far better to accept that this is a new area of law that will need to gradually develop with ongoing parliamentary oversight. For context, the UK has passed an immigration act once every 2.5 years this century. The Communications Act is now 18 years old.

63. The three problems we see with the definition and process are as follows. First, as with most of the draft Bill, the definition of harm gives too much power to Ministers. It is impossible to imagine legislation giving Ministers such sweeping powers to designate harmful content offline that, say, newspaper publishers were then required to restrict. The case has not been made that this power is necessary or proportionate. At the very least, this power deserves far more parliamentary scrutiny than the draft Bill allows for. While a time-limited power of this sort might have some value in responding to emergencies, we see no reason why these far-reaching choices should be made by Ministers and set out in unamendable regulations with such limited opportunity for parliamentary debate.

64. Secondly, the ambiguous threshold of "significant" harm can and will be interpreted very differently according to the interests of those involved. The word "significant" can be used to mean anything from "not trivial" to "having serious consequences". Internet companies have every incentive to use the strongest interpretation so as to reduce the work they have to do to meet their obligations. At the very least, this ambiguity will hinder the regulator's work and credibility by leaving them at constant risk of legal challenge for overstepping these blurry boundaries.

65. Thirdly, the draft Bill's definition of harm will make it harder to address the real social harms that the government itself says come from disinformation ("threaten public safety; fracture community cohesion; reduce trust in institutions and the media; undermine public acceptance of science's role in informing policy development and implementation; damage our economic prosperity and our global influence; and undermine the integrity of government, the constitution and our democratic processes."). It will either

require contorted and convoluted attempts to explain how, for example, election interference harms individual people psychologically, or simply leave these harms unaddressed. It is notable that the only part of the draft Bill that in any way responds to the government's own concerns on disinformation, Section 112 ("Secretary of State directions in special circumstances"), does not use the definition of harm used elsewhere but instead refers broadly to threats to "the health and safety of the public, or to national security". Perhaps this shows that the focus on harms to individuals is too narrow to address the harms the government itself is worried about.

66. As we have argued elsewhere in this submission, designating specific harms on the face of the Bill so it is clear what it means to address is the productive way forward. It could also unlock a better way to set out threshold(s) in a more meaningful and actionable way for service providers. Whilst there could be value in retaining a cross-cutting reference, in our view, it is unlikely to be able to work well alone without being seen in combination with a threshold or benchmarks that pertain to a specific category or subcategory of harm. This is because harms are different, and a single threshold such as "having, or indirectly having, a significant adverse physical or psychological impact on an adult of ordinary sensibilities" may be insufficient as a threshold for proportionate action in different contexts.

67. In our research and work on identifying and addressing harms from misinformation[13] we have seen how important it is to understand the different types of harm and the evidence of their impact so as to know whether action is necessary or appropriate, and, if it is, how that can be proportionate.

68. Such a system should not be fixed in time: mechanisms should be available to refresh approaches as learning and evidence emerges on the nature of online harms, their possible overlap and so on. If Ofcom does have the appropriate powers and capabilities, then it can set out these harm-specific benchmarks based on the best evidence base through an open transparent process and subject to democratic oversight. The only available alternative is to leave these decisions in the hands of overseas internet companies and closed-door negotiations with governments.

69. If the Bill sets out the key harms, we believe it would be possible to strengthen the draft legislation for success in addressing those harms in a way that would still allow for the necessary flexibility and changes, including around what appropriate action is. Harm from health misinformation should be in the Bill, for example, as a clear case in point.

Algorithms and user agency

---

[13] For example, *Tackling misinformation in an open society* (2020)
https://fullfact.org/media/uploads/full_fact_tackling_misinformation_in_an_open_society.pdf

What role do algorithms currently play in influencing the presence of certain types of content online and how it is disseminated? What role might they play in reducing the presence of harmful content?
Are there any foreseeable problems that could arise if service providers increased their use of algorithms to fulfil their safety duties? How might the draft Bill address them?

Independent scrutiny of the algorithms of in-scope internet companies

70. There are many current and foreseeable problems that could arise if service providers increase their use of algorithms. In fact, the draft Bill will encourage them to use systems that will have serious damaging unintended consequences which will be largely hidden from scrutiny. Content moderation at internet scale has to be done by machines. They cannot do this accurately, contrary to the claims of many internet companies. Error rates, even when they can be believed, are very significant. When we do get glimpses behind the curtain, the reality is shocking.

71. As described elsewhere in this submission, we have seen internet companies deploying algorithms that filter out content or take down posts that should not have had any action taken. Perversely, in passing this draft Bill Parliament would be requiring internet companies to rely even more heavily on technology that we already know cannot do the job. Worse than that, it doesn't provide for meaningful scrutiny of this technology or how it is used. We expect the result to be significant infringement of individuals' freedom of expression, not by design but by accident.

72. We describe above (see 30. and 31), how content moderation algorithms that work well can do real good, but if they malfunction, they can cause real harm. Yet the effects of content moderation algorithms are not subject to external scrutiny. The draft Bill should be amended to put in place a requirement for independent testing of safety-critical content moderation algorithms to address their unintended effects.

73. Unless there is some sort of independent process to verify algorithms and certify them in some way as being low risk in relation to direct harm to individuals, discrimination or as having no disproportionate impact on freedom of expression, then these harms will be perpetuated in this way.

The role of Ofcom

**Is Ofcom suitable for and capable of undertaking the role proposed for it in the draft Bill? /
Are Ofcom's powers under the Bill proportionate, whilst remaining sufficient to allow it to carry out its regulatory role? Does Ofcom have sufficient resources to support these powers?**

74. If the draft Online Safety Bill can be very significantly improved, then Ofcom could better build capabilities to fulfil its role, increasing its effectiveness over time. At this stage of progress in proposals, it would not make sense to consider a different or new regulator and, as such, every effort needs to be made to

ensure Ofcom can be a successful regulator both through improvements to the proposed legislation and regime as well as Ofcom developing to be institutionally fit for its new remit.

75. The Bill should explicitly give Ofcom responsibility for understanding the harms caused by misinformation and disinformation given that in any field of regulation, proportionate risk-based responses require an evidence base and intelligence. This would build on Ofcom's existing top class research function. Ofcom's allocated powers on obtaining information needed can be deployed to such an end, and done so with strong transparency so information is not only private to Ofcom, but builds understanding on misinformation and disinformation for all stakeholders able to play their part in taking action to address associated harms.

76. The Advisory committee on disinformation and misinformation (Clause 98) should have within its function an explicit remit on such an Ofcom responsibility for understanding the harms caused by misinformation and disinformation. This should be set out in addition to the other three roles envisaged for the Advisory committee in advising Ofcom on: how regulated services deal with disinformation and misinformation (98(4)(a)); transparency reports and disinformation and misinformation (98(4)(b)); and, how the regulator promotes media literacy around disinformation and misinformation (98(4)(c)).

77. None of the internet companies are sufficiently transparent on the action they have taken to prevent misinformation online, including Covid-19 misinformation. Because internet companies can silently and secretly shape public debate, transparency requirements must be clearer on the face of the Bill, so we can better understand what choices in-scope companies are making and why.

78. What is needed is real-time information on suspected misinformation from the internet companies; independent scrutiny of the use of AI by these companies and its unintended consequences (as above), and real-time information on the content moderation actions taken by companies and their effects. It is not practical that Ofcom is only largely asking for information annually, given the speed at which misinformation evolves and the speed at which information crises occur. If there is only a snapshot of the problem months later, then the harm will already have occurred: retrospective reports do not provide timely information when it is needed by Ofcom and stakeholders if online harms are to be addressed.

79. There are many dimensions of the Bill, such as the Codes, that will be dependent on both the powers and the resources that Ofcom has, which will determine whether the regime works well or not.

80. Whilst some indications of additional resourcing for Ofcom are apparent in new roles being hired, for example, it is not yet clear what future capacities are required to address many harms. There is a case for the Secretary of State to be clear as soon as possible on priority harms so that Ofcom can move to put in place capacity and others can align to play their role in effective regulation, although we believe these would better be debated and set out by parliament itself in the Bill. Even indicative priority would be helpful here.

81. Media literacy is a known and critical remit (indeed where Ofcom's role is the very purpose of the Bill alongside Ofcom regulating internet companies), and this is where a very significant uplift in resourcing is needed so that online media literacy is improved to the degree needed (see also media literacy, below, on ambition). This is particularly true given the current context: Online Media Literacy Strategy expenditure by the government is presently not credible. Both Ofcom and government funds on online media literacy need a very significant increase if they are each to fulfil their distinct roles in improving online media literacy. Leveraging or inspiring action by others simply will not be enough (and must be well resourced in itself), but appropriate investment in online media literacy will enable Ofcom to do what is needed and will enable whole of society support to develop).

Are there systems in place to promote transparency, accountability, and independence of the independent regulator? / How much influence will a) Parliament and b) The Secretary of State have on Ofcom, and is this appropriate? / Does the draft Bill make appropriate provisions for the relationship between Ofcom and Parliament? Is the status given to the Codes of Practice and minimum standards required under the draft Bill and are the provisions for scrutiny of these appropriate?

82. We believe that Parliament should have a fuller role than that set out in the draft Bill.

83. A shift to legislation that clearly identifies the specific harms it seeks to address would be far more effective legislation and would need to be regularly revisited by Parliament. If Parliament passes the current skeleton bill it will hand over powers to shape public debate to a politically appointed body subject to control by the government of the day. This is not a healthy possibility.

84. Full Fact believes that the powers envisaged in the draft Bill for the Secretary of State require close examination as they are too extensive and need to be narrowed in several areas.

85. For example, the Secretary of State's power to direct Ofcom to modify its codes of practice to bring them in line with 'government policy' (Clause 33(1)(a)) should be dropped. Ofcom's independence as regulator should not be undermined by any such provision in the Bill.

86. The pivotal regulation power in the Bill is set out in Schedule 4, which allows the Secretary of State to decide what services will be regulated by the Bill. We hope this provision and any future regulations will be carefully scrutinised, but it is properly a matter for primary legislation. The power given to the Secretary of State here is indefensibly broad and makes it impossible for Parliament to understand, debate, and scrutinise the effects of this draft Bill.

87. The unilateral power of the Secretary of State to exempt services based on nothing more than what the government considers (i.e., no advice from Ofcom, no consultation, no specified standard of evidence) is both unnecessary and unreasonable.

88. The power of the Secretary of State to amend the online safety objectives by regulations must make the Joint Committee wonder what point there is in scrutinising the ones in the draft Bill.

89. Whilst many powers for a Secretary of State might be expected in an Act such as this, changes are warranted to ensure separation of roles and that the role of Parliament be strengthened, increasing its oversight and wider role, including in relation to how the framework and regime develops into the future. Every major regulation power in the draft Bill needs scrutiny. Most of these deserve to be replaced with either primary legislation, stronger parliamentary oversight of changes, or more specific conditions for the use of the powers.

90. The Bill should allow the regulator the space to take decisions based on the available evidence. A case in point is the areas referred to in Clause 12, Ofcom should have enough powers to address threats to public safety, public security and national security.

91. If Ofcom is to have such powers, it needs the confidence of all sections of society, and it needs independence baked into its constitution. Many public bodies of similar sensitivity report directly to Parliament and are appointed subject to some form of parliamentary approval. The appointment of the head of the National Audit Office requires agreement from the Chair of the Public Accounts Committee. In the case of the Electoral Commission, the political balance of the Commission itself is subject to rules laid down in legislation, and the appointments process is the responsibility of a Speaker's Committee. Ofcom needs a clear legislative framework and strong protection for its independence and impartiality in law if it is to take on this new role. These examples illustrate that both are very much weaker than for comparable public bodies.

92. Finally, nothing in this draft Bill prevents any future government from bypassing both parliament and Ofcom and leaning on the internet companies directly to change their policies and practices. We have already highlighted how this can lead to censorship-by-proxy without an open democratic transparent process.

93. The Online Safety Bill should include some form of reporting requirement for the government to publish details of all efforts it makes to influence internet company decisions about specific items of content, specified accounts or their terms of service (this could recognise the necessity for a limited time delay in the case of content considered national security sensitive, perhaps subject to review by an appropriate committee).

Are the media literacy duties given to Ofcom in the draft Bill sufficient?

The need for more and better media literacy than set out in the draft Bill

94. Given that the Online Safety Bill is in purpose and scope a law about how the regulator regulates internet companies and how the regulator goes about media literacy, the new regime could be a huge opportunity for transforming media literacy in the UK in the digital era. However, the ambition is presently not sufficiently clear.

95. The draft Online Safety Bill gives Ofcom powers on requiring transparency from service providers on what they are doing to improve the media literacy of their users and how they are evaluating the effectiveness of such action, and Ofcom can provide guidance on those efforts. Some additional transparency on what platforms are doing, plus this guidance, do not amount to a massive step

change and could mean just more of the same: platforms reporting a lot of activity without much evidence that such efforts are solutions commensurate with the problems.

96. The draft Online Safety Bill requires that Ofcom 'carry out, commission or encourage educational initiatives designed to improve the media literacy of members of the public'. Again, either through its own action, or what it leverages and inspires from others, will what Ofcom does be of a scale needed, will it play its full part? The pandemic has reminded us again that providing good information proactively is an effective way to limit the damage bad information can do. With audiences fragmenting, our public bodies need to gear up to do this at a much larger scale than in the past.

97. Ofcom does very good work on media literacy. Yet it is limited in impact: its research on the state of play and what works is very useful, but Ofcom's direct and indirect action needs to translate into accelerated progress—real world difference—and that means resources and initiatives must add up to enough to move the dial.

98. Ofcom's media literacy activity is presently focused on generating an evidence base of UK adults' and children's understanding and use of electronic media and sharing that evidence base within the regulator itself and with external stakeholders. The Bill needs to clarify to what extent Ofcom's research into people's media literacy needs should play a role in shaping public policy or provide other organisations and agencies with evidence that informs what they do in their initiatives. That would also make it easier to envisage what will or needs to be different in the new regime of the eventual Online Safety Act.

99. Ofcom's research must go beyond the state of play to more of what works and setting out a concrete action agenda that emerges from the evidence on interventions. For example, if Ofcom's research released in April this year tells us 24% of UK adults did not consider the potential trustworthiness of online information at all, this is important to know. But even more important would be Ofcom's informed recommendations, for example, about what the internet platforms it will soon regulate could, or indeed should, do to change that, alongside action by others.

100. There are many areas in the media literacy provisions requiring scrutiny, deliberation and changes on the face of the Bill. Overall, to "Promote" and "Improve" media literacy do not necessarily give us a clear notion of the outcomes being sought. What level of improvement and on what key measures are unclear. Without clear ambition from government and authorities such as the regulator it is not apparent what the shared project is. Taking the same Ofcom research figure above, that 24% of UK adults did not consider the potential trustworthiness of online information, what might be the target ambition? Should we not be aiming for that figure to be practically all—something around 3%? By when? And what is the collective assessment of what it might take to help people to do that?

101. It is difficult to envisage internet platforms playing a full part in media literacy without some shared sense of destination at a national level: of what media literacy levels are being worked towards. The Online Safety Bill and the annual plans of the Online Media Literacy Strategy could change that. At present there

is very little substance on what Ofcom's role will be in practice on preparing guidance "about the evaluation of educational initiatives"; "about the evaluation, by providers of regulated services, of any actions taken by them in relation to those services to improve the media literacy of members of the public"; "about the evaluation, by persons [i.e. the platforms] developing or using technologies and systems"; and 'of the effectiveness of those technologies and systems in improving the media literacy of members of the public."

102.	This Bill in its present draft is not ambitious enough on media literacy in the digital era. This risks a situation developing where Ofcom has low ambition, will or lacks leverage to actually improve this nation's media, digital and misinformation literacy. Whilst much of this does not need to be in legislation, some changes are warranted to make sure real progress is made over time.

103.	Strengthening media literacy in this Bill would be part of the freedom of expression, citizen-supporting methods of tackling the problems in our information environment: whether navigating the everyday misinformation that comes with a democracy or more harmful misinformation.

In addition to the above answers to questions the Joint Committee has invited submissions to include, Full Fact raises concerns about the necessity to enhance the role of the public for any resulting Act and regime to succeed:

Enhancing the role of the public under the Online Safety Bill and new regime

104.	Full Fact believes the draft Online Safety Bill and resulting regime would be more effective if the legislation and associated arrangements strengthened the role of the public in various ways.

105.	The process throughout must be inclusive and diverse with citizens and civil society participation both as the legislation is developed and in the resulting regulatory regime.

106.	Significant and continuous debate in Parliament and beyond with the public is needed on these measures given the potential wide-ranging implications and to ensure the resulting regime reflects the concerns of the public, for example in relation to harmful misinformation (see below).

107.	As above, under Transparency requirements, there should be provision to ensure public transparency to citizens by default. Whilst Ofcom should have the legal powers to access information it needs, the draft Bill is not clear on the extent of accountability to the public when it comes to information required.

Citizen voice and participation

108.	Citizens should have an ongoing voice in shaping the new regime so that it can be effective. A number of civil society groups are increasingly engaging with their supporters affected and/or concerned around harms online and it is encouraging that some parliamentarians and committees are also reaching out directly to the public and those that represent many groups affected. There is

more to do in this regard as the legislation is developed. The complexity of the draft Online Safety Bill makes that a very real challenge.

109. There is increasing evidence that the public overwhelmingly wants a better online environment. In forthcoming public opinion research conducted by Ipsos MORI on behalf of Full Fact, 75% of UK adults (18+) are worried about the spread of misinformation and the issue of misinformation is at levels of concern comparable to the Common Market/Brexit/EU/Europe and crime/law and order. The public view social media and video-sharing sites as most to blame for the spread of information around news and current affairs online that is false or misleading and believes these companies, media organisations and politicians or the government should be responsible for tackling the problem. Ipsos MORI research shows that half of the UK public agree misinformation is a problem that can be solved. This further underlines that if the draft Online Safety Bill can be improved, the public will support it being a success in addressing harmful misinformation and welcome that the Government have put in place a new regime to tackle the problems associated with it.

110. UK parliamentarians now have the opportunity with the Online Safety Bill to fulfil this expectation and hope that harmful misinformation will be addressed, and to make sure any approach to tackling misinformation is risk-based and proportionate. In doing so, the Government and Parliament should recognise that the voice and participation of the public is critical to success. Avenues to realise this should be uplifted in dialogue on what is in the Bill, how the regime evolves and the necessary accountabilities that need to flow to people as users of services and as citizens,

111. One example would be around where the draft Bill sees Ofcom setting up an expert advisory group on mis/disinformation. Whilst thoroughly inadequate in itself to address the harms that can result from misinformation, this could be a useful body, especially if it has strong representation from civil society groups and it is positive that representatives of UK users of regulated services as well as experts will be included. This is an area where it is possible to strengthen the provisions so that there is a wider forum to ensure citizens affected by harmful misinformation are active participants in the system and how it operates.

112. If the draft Online Safety Bill can be changed for the better, as it absolutely must in the months ahead and in future iteration, the prospect of a better online environment for the people of the UK can become more of a reality with all the benefits that can follow from that–from improved public debate to a healthier society and democracy.

20 September 2021