

Written evidence submitted by Logically (OSB0094)

Introduction

Logically welcomes the opportunity to provide written evidence to the Joint Pre-legislative Scrutiny Committee on the Draft Online Safety Bill. As specialists in mis-and disinformation, our submission focuses solely on elements of the draft bill relevant to this specific problem space.

We are supportive of the Government's aims to make the UK the safest place to be online, and recognise the importance of both protecting freedom of speech and guarding against over-censorship in the online environment. The UK is far from the only country grappling with the very complex challenge of state intervention in the online space, and we recognise that it is very hard to apply clear definitions to mis-and disinformation, in particular what kind of misleading content could reasonably cause physical or psychological harm.

However, while acknowledging that this bill is attempting to resolve very challenging issues, we believe there are a number of elements of the bill as currently drafted that could have unintended consequences or fail to achieve the stated objectives without further consideration.

Definitions of categories of priority harm

We note that it is the Government's intention not to define and include specific categories of priority harm on the face of the bill, instead opting to propose these in secondary legislation.

While we recognise the Government wants to leave flexibility to respond and adapt to evolving harms, the uncertainty and lack of clarity about the kinds of content that will be in scope, limits and delays actions they could be taking in the short to medium term. If social media companies were given a clear sense of categories of priority harm covered in the bill, they could begin to make changes to their systems and content moderation policies now, rather than potentially wait another 2-3 years before the bill and then secondary legislation comes into effect.

When the time does come to define specific categories of priority harm as it relates to mis-and disinformation, we ask that the proposed Advisory Committee of Misinformation and Disinformation, to be established by Ofcom, plays a key advisory role in setting these definitions. Expert, multi-stakeholder input will be essential to ensure that the right kinds of issues are included.

Protections for democratic content

While we recognise what the government is intending to achieve with this exemption, we believe it is currently sufficiently ill-defined that it could create a significant loophole for the sharing of harmful misinformation. To create and apply this exemption as broadly as currently drafted in the bill risks providing cover to misinformation being included in party or candidate literature, no matter how fringe, without fear of penalty.

This also presents the possibility of conspiracy groups or bad actors pushing misinformation specifically targeting politicians or campaigners to amplify their messages, in the safe knowledge it wouldn't be challenged. Our [investigation into the Hart Group](#) is just one example of how those pushing misinformation target legitimate public figures and media outlets to amplify and endorse their content. Without thoughtful safeguards in place, there is a clear risk we could see more of this kind of activity, particularly around elections and political campaigns.

Effectiveness of the duty of care and the focus on individual vs societal harm

The duty of care does have the potential to be an effective approach to ensuring the safety of individual users. However, a duty of care applied solely to promoting safety and preventing harm at an individual level risks omitting from scope an enormous amount of content that does significant harm at a societal level.

A focus on mitigating the direct harm caused by individual content to individual people risks allowing the more pervasive *indirect* harms caused to individuals through the wider degradation of the civic information environment. To take the anti covid-vaccination movement as an example, it is difficult to capture the harm caused by even the most egregious anti covid vaccination messages in terms of their direct harm to individuals. The risk of harm to individuals from anti-vaccination misinformation comes not directly from individual pieces of misinformative content, but indirectly from the support such content lends to the erroneous view that covid vaccination is unnecessary or harmful.

Similarly, repeated exposure to mis-and disinformation can lead to some vulnerable audiences becoming more susceptible to extremist views and content, even becoming radicalised to take part in illegal activity. Again, this would be difficult to define as a specific harm to an individual as it is something that happens over time and not related to one piece of content, and therefore may be missed in the scope of the bill. We urge the committee to recommend that the government rethinks its decision for the bill to focus solely on individual harms.

A duty of care focused on preventing individual harm will also not address the systemic problems of echo chambers or filter bubbles, nor precipitate any improvements to the wider information environment. One way of addressing this would be for the bill to also introduce an obligation on major platforms to demonstrate regard for protecting and preserving civil discourse. The bill currently asks this of freedom of expression and privacy, so there is scope for this to be expanded to include having regard for protecting civil discourse reflected in their terms of service. [Recent moves by Facebook](#) to reduce the visibility of political content in news feeds show that it is possible for platforms to treat political and social discourse differently. We believe the government should use this bill as an opportunity to encourage all major platforms to have a similar approach in their system design.

The role of safety by design and algorithmic recommendations

Social media platforms can make significant improvements by making changes at a system or platform level. Recent changes announced by a number of major companies ahead of the implementation of the Age Appropriate Design Code in September 2021 demonstrate the power of thoughtful regulation in making positive changes to user experience. Similarly, it should be possible for changes to be made to algorithmic recommendations to mitigate against echo chambers, by making certain classes of content or topics such as misinformation or conspiracy theories inaccessible through recommender systems. YouTube made [encouraging steps](#) towards this a couple of years ago, but a more concerted effort is needed to make it more effective and replicated across all platforms.

Transparency reports

We welcome the introduction of standardised, comprehensive transparency reporting obligations for social media platforms. While the major platforms have made good progress in recent years in improving the amount of detail they share about problematic content on their services, there is significant room for improvement. Given the substantial negative societal impact caused by mis- and disinformation, we recommend that Ofcom requires companies to report on this type of content specifically. This should include metrics for volume, removals, reach and engagement, assessment of effectiveness of content moderation policies, and other contextual information that provides a full picture of the scale of the problem. Only with this kind of full and detailed reporting will we be able to truly measure progress and impact of various interventions and policies.

We also strongly recommend introducing independent third party auditing of companies' transparency reports by qualified researchers in this space, in addition to oversight by Ofcom, to improve the quality of both the data and outcomes of the reporting process.

Media literacy

We welcome the prominent inclusion of media literacy throughout the bill. However, any media literacy efforts - either by Ofcom or by the platforms themselves - need to be rigorous, evidence-based and accompanied by a full impact assessment so we can be sure they are actually delivering positive results. There should be robust reporting and evaluation frameworks for any media literacy initiatives included as part of the transparency reports for companies, and in Ofcom's own annual report. For social media companies, this should include reporting on changes made to platform and system design, in addition to investment in partnerships and programmes.

About Logically

Logically combines advanced AI with a world class open source intelligence (OSINT) team and one of the world's largest dedicated fact-checking teams to help government bodies and social media platforms uncover and address harmful misinformation and deliberate disinformation at scale. Logically is an award-winning international team of 120 data scientists, engineers, analysts, developers and investigators united by the company's mission to enhance civic discourse, protect democratic debate and process, and provide access to trustworthy information.

Logically has developed a suite of products and services to reduce and eventually eliminate the harm caused by the spread of misinformation and targeted disinformation campaigns. These include Logically Intelligence - Logically's sophisticated threat intelligence platform; AI-assisted fact-checking; and a consumer app and browser extension which provide access to our fact-checking team and gives users the ability to rate the credibility of online news sources.

Logically is a verified signatory of the International Fact-Checking Network's code of principles. We have a dedicated fact checking and editorial team who produce frequent fact checks as well as detailed analysis and reports into specific misinformation actors and trends. Our published fact checks can be found on our website [here](#) and a selection of our deep dive investigations into specific conspiracy theory trends can be found [here](#).

We are currently working with government agencies around the world to combat mis- and disinformation, including protecting election integrity and vaccine rollouts. We also have dedicated fact-checking partnerships in the UK with TikTok and Facebook.

September 2021