

## The Health Foundation's response to DCMS and Home Office consultation on the Online Harms White Paper

*July 2019*

### About the Health Foundation

The Health Foundation is an independent charity committed to bringing about better health and health care for people in the UK. Our aim is a healthier population, supported by high quality health care that can be equitably accessed. We learn what works to make people's lives healthier and improve the health care system. From giving grants to those working at the front line to carrying out research and policy analysis, we shine a light on how to make successful change happen.

### The Health Foundation's response

The Health Foundation emphasises the impact that the proposed approach to regulating online harms may have on health and wellbeing within the UK. We welcome efforts to combat illegality online and ensure that providers have a duty to ensure the safety of their products. There are substantial fears from both parents and the media around the impact of online harms to health, particularly for young people. This includes both societal harms to health, for example the spread of misinformation relating to vaccines as well as individual harms, such as those relating to mental health. This is of concern given that rates of emotional problems and self-harm are rising, particularly amongst young women and girls ([Morgan et al 2017](#), [Sadler et al 2018](#), [McManus et al 2019](#)). However, the existing evidence base demonstrating online harms to health is limited ([UK CMO commentary](#), [House of Commons Science and Technology Committee report](#)), and cannot consider any longer term consequences. At the same time, the benefits of being online are often overlooked, particularly as voices of young people are not considered. For example, young people with Special Education Needs (SEN) and disabilities may be more likely to experience cyberbullying, but also emphasize the value of support from online communities ([Anti-bullying alliance](#); [Caton et al 2016](#)). The Health Foundation is concerned that without proper attention paid to the experiences of these users, new regulation will fail to minimise the risk of harm for all, and unintentionally limit access to the benefits of online technologies.

To develop an effective, evidence-based approach to regulation there is an urgent need to better understand the impact of being online. This requires strong action from the proposed regulator to ensure much greater transparency from providers. This should include access to information regarding user activity, content moderation and the way in which technologies are designed, which can lead to the algorithmic amplification of potentially harmful content. Information and data should be shared by providers with independent researchers to allow evaluation of risks and benefits that can be used to guide action from the regulator. To ensure a nuanced approach, this research, and any further work to develop the evidence base for policy, should be done in collaboration with users, particularly young people.

Without this input, the regulator cannot develop a set of standards which appropriately reflect the experiences and needs of the users they are aiming to protect.

**Question 1: This government has committed to annual transparency reporting. Beyond the measures set out in this White Paper, should the government do more to build a culture of transparency, trust and accountability across industry and, if so, what?**

The Health Foundation welcomes the commitment to transparency that includes sharing data with academics for research purposes. We think that this is needed to ensure the possible impacts of the growth in online technology on the health of the population, particularly children and young people, is understood. As has been highlighted in a number of reports, this is an area where there is a lack of available evidence ([UK CMO commentary](#), [House of Commons Science and Technology Committee report](#)). Recent data sharing agreements between major social media companies and social science researchers, brokered through a commission, [social science one](#), provide a framework which could be adapted for the UK context and used to examine questions regarding health and wellbeing.

To build a culture of transparency, it is important that as well as data on user activity, companies make information on platform algorithms, which shape user experience and influence behaviour online, available to the regulator and trusted independent researchers. Particularly in the case of children and young people, the duty of care placed on providers must extend to ensuring that algorithms do not encourage exposure to inappropriate or extreme content online. The need for this is acknowledged in the white paper, however the availability of data to evaluate the risks that this amplification of content pose is still lacking.

A second distinct issue related to trust and accountability of data-driven web technology is bias. As the white paper notes, online technologies play an increasingly important role in our lives. But there is increased attention and awareness around the potential for technology to magnify and cement existing inequalities, for example through biased algorithms that reinforce stereotyping or underrepresentation of often marginalised groups (including black and minority ethnic individuals, sexual minorities, women, and those from more deprived communities) ([Crawford 2018](#)). Therefore, information is needed to allow researchers and regulators to identify such bias, allow monitoring, and enable regulators to hold industry accountable.

A key component of building trust and accountability is to ensure that the public are aware of how their data are used and could be used in the future. This is not currently the case, which undermines both public trust and the potential to realise the benefits of online technologies, including in the sphere of health and health care. There may be different approaches that can be adopted to build this trust, for example, the Open Data Institute is piloting the concept of a '[Data Trust](#)', which is about developing models of control that aren't just technical but are also organizational in nature too.

**Question 8: What further steps could be taken to ensure the regulator will act in a targeted and proportionate manner?**

Evidence of rising rates of emotional problems and self-harm, particularly amongst young women and girls (Morgan et al 2017, Sadler et al 2018, McManus et al 2019), has fuelled a great deal of media attention given to the potential impact of digital technology on mental health. However, there is a lack of strong evidence demonstrating these risks (UK CMO commentary, House of Commons Science and Technology Committee report, Orben and Przybylski 2019). To ensure the regulator is acting in an evidence-based and proportionate way, it is critical that data is made available and high-quality research into this issue is funded as a priority.

It is also important that work to understand the impact of online activity addresses both potential benefits as well as potential harms. We are concerned that regulation developed without attention to the benefits of being online will be poorly targeted and may have unintended consequences.

For example, work from the Health Foundation's Q Labs on peer support networks has highlighted the value and growing demand for online peer support networks, as well as their value in correcting misinformation. Platforms which offer peer-support online are likely to fall into the proposed scope of the regulatory framework. This includes the Health Foundation funded, clinically supported forum "The Pill Or What", which aims to encourage more informed contraceptive choices, as well as the moderated, supportive online Elefriends community from Mind, the mental health charity. Whilst it is of course necessary that these online communities are known to be safe for users, clarity and support around any new regulation for organisations acting in this important space is invaluable to ensure that they can comply with their duty of care responsibilities and remain able to support their users.

To capture the needs of users, particularly young people, collaboration is essential. In work from the Health Foundation's Young people's future health inquiry, we have highlighted the importance of including young people throughout the process of mapping out impacts on their health and wellbeing. For example, when examining influences on mental health, the young people involved in the project emphasised the beneficial impacts of social media, in contrast with clinicians and other professionals who were more likely to focus on potentially harmful effects of being online.

Finally, to ensure that the regulator is successful in improving safety, more work is needed to establish effective methods by which to reduce harms. For example, in cases where content may be inappropriate for children, evidence suggests internet filtering technology fails to achieve its aim of reducing exposure (Przybylski & Nash 2017, Przybylski & Nash 2018). The regulator should work to encourage the development a range of approaches and avoid overreliance on purely technological solutions to issues of safety.

**Question 16: What, if any, are the most significant areas in which organisations need practical guidance to build products that are safe by design?**

The Health Foundation is concerned that without careful guidance, efforts to improve safety may have unintended detrimental impacts. For example, efforts to identify and remove potentially harmful content relating to self-harm may have the unintended consequence of increasing stigma and isolation for patients seeking help and support online (Brown 2019, Branley & Covey 2017). But it is not only users that may be at risk of harm: the potential impacts of exposure to extreme content by content moderators working to create a safer online environment also needs to be considered. Therefore, organisations need practical guidance on key principles for developing safe environments and avoiding unintended harms, particularly for these groups.

To ensure a transparent approach to harms and build public trust, clear standards are needed on how to evaluate safety efforts, including the capture of unintended consequences. Clarity of expectations from the regulator will also be key to establishing whether small organisations (including charities offering peer support platforms) can provide the resources necessary to sustain their services.

The new regulator should develop guidance with the collaborative input from those with relevant expertise – including independent experts in the design of online platforms and in the case of harms relating to health, both clinicians and patients, to ensure that efforts to improve safety are effectively targeted.

**Question 18: What, if any, role should the regulator have in relation to education and awareness activity?**

The Health Foundation considers education and awareness activities to be an important component in working to reducing health-related online harms, within a framework of appropriate regulation. Doteveryone research suggests that there is low level of public understanding around digital technologies. As the white paper emphasises, equipping users with the tools to navigate the online world effectively and safely not only helps to reduce the likelihood of harms, but also enables them to benefit from the opportunities that new technologies offer. Nevertheless, we emphasize that education should not be used to compensate for an unsafe environment. The responsibility for safety remains with providers and a cornerstone to this needs to be improving transparency, with technology companies being held accountable for making their products understandable to users.

Education activities need to be appropriate to reduce harms and be supported by research. For example, targeted education efforts to tackle disinformation online may be effective (van der Linden et al 2017) and could be adopted to reduce the impact of hostile actors, who are promoting anti-vaccine messages to intentionally mislead, undermine public health efforts and reduce trust. In contrast, overreliance on restrictive efforts to keep children safer, including the use of technology to filter content (Wisniewski et al 2017, Przybylski & Nash

2017, Przybylski & Nash 2018) and calls to ban or limit screen time (Przybylski & Weinstein 2017) may be ineffective in reducing online harm, and limit the ability of users to develop strategies to protect themselves from online risks.

We highlight the need for a better understanding of practices which are effective in reducing harms, moving beyond an overreliance solely on content moderation to a more comprehensive approach. This includes building greater awareness around these issues, signposting to well-evidenced resources for users and parents of users (UK CMO commentary; RCPCH; Parent Zone). It is important that these education efforts are appropriately targeted, particularly given evidence that certain groups of users may be more likely to experience online harms. Different groups may have greater exposure to these unsafe environments (Craemer et al 2018, Nuffield Foundation), and there is a risk of exacerbating inequality if education efforts do not address these differences.

**For further information:**

Emily Eldridge

[Emily.eldridge@health.org.uk](mailto:Emily.eldridge@health.org.uk)

[www.health.org.uk](http://www.health.org.uk)