

Written evidence submitted by Professor Andrew Przybylski Associate Professor, Senior Research Fellow at University of Oxford (OSB0193)

Dear committee members,

As an expert who has been studying the positive and negative influences digital technologies on adults and young people for many years, I would like to thank you for providing the opportunity to provide scientific advice on the Government's draft Bill to establish a new regulatory framework to tackle harmful content online.

Introduction

I will confine my observations to what I know as an expert studying psychological aspects of mental health, well-being and online contexts and only reference the named objectives content and services which I feel I can authoritatively speak to these terms of references below.

I would start by observing that the idea that online platforms including diverse forms of social media have negative, causal, and ultimately harmful effects, on the mental health of users is deeply worrying and should be treated with the upmost seriousness. Widespread media coverage of indirect evidence of harm is disturbing. Unfortunately, this press attention does not help society sort high quality evidence and anecdotal or biased opinion. My own analyses motivated by the desire to get to the bottom of this question have indicated there is no evidence that the associations between adolescents' digital technology engagement and mental health have increased (1).

I believe that transparent and independent analysis of industry data is the best way forward to make the UK the safest place in the world to be online.

My experience working on this topic for more than a decade makes it clear that we are operating in a nearly complete data vacuum insofar as we attempt to define or measure a scientific basis for the concept of "online harms" (2). Understanding that the absence of evidence is not evidence of absence, **I would strongly caution against sweeping changes to law and policy that are not supported and informed by the full weight of reliable, transparent, and replicable scientific evidence.**

Terms of Reference

Will the proposed legislation effectively deliver the policy aim of making the UK the safest place to be online?

No. As far as I can tell the term online harms is not defined well enough to provide the basis of an effective intervention strategy for online platforms. I am concerned that interventions pursued in a haphazard fashion might lend the appearance of a fix but could have a wide range of unintended and unmeasurable consequences and will ultimately face repeal for a lack of scientific support (3).

Are children effectively protected from harmful activity and content under the measures proposed in the draft Bill?

No. We require more reliable data on how young people use the Internet and how specific hot button topics, such as exposure to adult content cause (or do not) cause harm (4–6). If they do, the effectiveness of steps meant to limit exposure to such events should be tested before interventions are put into law so that their longer-term effectiveness might be measured and improved upon (7,8). If changes are made to online spaces by way of legislation and the consequences of these interventions are not meticulously tracked, we risk losing an invaluable opportunity to set a good global example for making the online world safe for young people.

Does the draft Bill make adequate provisions for people who are more likely to experience harm online or who may be more vulnerable to exploitation?

No. There is no scientific basis to flag specific individuals, or groups of individuals, as reliably more or less resilient to the challenges of the online world. Biological age is a poor proxy for maturity or resilience. Conducting this research is critical but it has not been resourced by either The Government or national research councils. A reliable scientifically grounded model of individual and group risk and resilience is a prerequisite to judge the suitability of any such provision (9). We should have this, but we do not.

Is the “duty of care” approach in the draft Bill effective?

No. I believe it is likely mis-specified. It appears conditions on a unrealistically low level of risk we would not adopt in other parts of our society. Introducing a ‘precautionary principle mindset’ begs the question of harm and legitimates solutions without a principled process for understanding the causal dynamics - critical to highlighting mechanisms for intervention - and weighing the costs and benefits of intervention.

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?

It has the potential to. I believe strengthening provisions for regularly released and granular (i.e. not aggregated); industry data sharing will be required to make trying to meet any of these standards successful (or not).

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?

This is the most overlooked aspect of this endeavour from my perspective. Beyond anecdotal accounts, by default, statistical significance of a negative correlation between some form of online activity and a mental health or well-being outcome is taken as evidence of a ‘harm’ by many researchers and charities active in the area (10–12). **This is an exceptionally poor standard to judge the strength of scientific evidence.**

Systematic evidence reviews, and reviews of reviews which grade the former on the quality of evidence also show this is a poor standard for and they do not support the idea that 'online harm' exists (13,14). It should not be the basis in scientific study and cannot be the basis for legislation.

Harm has at least two aspects that need to be grappled with, the magnitude of a link and causal nature of a link. (A) In terms of magnitude: I think the bare minimum standard for harm should be whether an individual can, subjectively reliably report lower levels of psychological health or well-being (12). This threshold, roughly equivalent to shift of one half of a standard deviation in well-being. The existing evidence does not suggest that regular use of online or media platforms is associated with even a small fraction of this difference in the population or within people across time (5,15). (B) In terms of causal nature: I think the bare minimum standard here for harm should be whether we can prove, reliably that an online event causes (or reliably precedes) a negative shift in psychological health or well-being. This is a very basic standard we would expect in other domains such as effectiveness of pharmaceuticals or transportation safety. I have found no evidence for this (16,17). It is entirely absent from both the legislation being proposed and existing evidence base.

Taken together, it is my view we have nothing near the quality of evidence required to set a principled and evidence-based standard for "online harms" which could be deployed by Ofcom or any other UK body. Those who claim otherwise are misunderstanding or misrepresenting how scientific knowledge is produced or evaluated on this topic.

Closing

It is my view this bill in its current form reflects a more fundamental problem in the way that social science scholarship surrounding online platforms is produced and integrated into policy in the UK. Social media companies have the capacity to collect the data needed to inform this process and possible regulations. They do not. The Government has the resources to support large-scale independent team-based science. It does not. 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. This topic will grab headlines, but it will not be informed by reliable evidence. Without addressing the need for independent analysis of industry data I do not see a way out of this trap.

In short, my research field is not capable of producing the scientific evidence the public or The Government needs to have a fruitful policy debate on the topic of online harms. I therefore urge you to consider the need to change the scientific landscape before proceeding to take sweeping policy decisions which might need to be dramatically revised later.

Signed,

Andrew K. Przybylski

Andrew K. Przybylski

References

1. Vuorre M, Orben A, Przybylski AK. There Is No Evidence That Associations Between Adolescents' Digital Technology Engagement and Mental Health Problems Have Increased. *Clin Psychol Sci*. 2021 May 3;216770262199454.
2. Przybylski AK, Etchells P. We won't know if screen time is a hazard until Facebook comes clean. *Wired UK* [Internet]. 2019 Jan 31 [cited 2019 Jul 15]; Available from: <https://www.wired.co.uk/article/screen-time-uk-report-facebook-research-app-data>
3. Im E. Korea to ax games curfew. 2021 Aug 25; Available from: <http://www.koreaherald.com/view.php?ud=20210825000854>
4. Orben A. The Sisyphean Cycle of Technology Panics. *Perspect Psychol Sci*. 2020 Sep;15(5):1143–57.
5. Orben AC, Przybylski A. Social media is nothing like drugs, despite all the horror stories [Internet]. *The Conversation*. 2017 [cited 2017 Oct 2]. Available from: <http://theconversation.com/social-media-is-nothing-like-drugs-despite-all-the-horror-stories-79382>
6. Orben A, Przybylski AK. The association between adolescent well-being and digital technology use. *Nat Hum Behav*. 2019 Jan 14;1.
7. Przybylski AK, Nash V. Internet Filtering Technology and Aversive Online Experiences in Adolescents. *J Pediatr*. 2017 May 1;184:215-219.e1.
8. Przybylski AK, Nash V. Internet Filtering and Adolescent Exposure to Online Sexual Material. *Cyberpsychology Behav Soc Netw*. 2018 Jul;21(7):405–10.
9. Orben A, Przybylski AK. Screens, Teens, and Psychological Well-Being: Evidence From Three Time-Use-Diary Studies. *Psychol Sci*. 2019 Apr 2;095679761983032.
10. Lakens D. How a power analysis implicitly reveals the smallest effect size you care about [Internet]. 2017 [cited 2018 May 14]. Available from: <http://daniellakens.blogspot.com/2017/05/how-power-analysis-implicitly-reveals.html>
11. Lakens D. Equivalence Tests: A Practical Primer for t Tests, Correlations, and Meta-Analyses. *Soc Psychol Personal Sci*. 2017 May 1;8(4):355–62.
12. Anvari F, Lakens D. Using anchor-based methods to determine the smallest effect size of interest. *J Exp Soc Psychol*. 2021 Sep;96:104159.
13. Parry DA, Davidson BI, Sewall CJR, Fisher JT, Mieczkowski H, Quintana DS. A systematic review and meta-analysis of discrepancies between logged and self-reported digital media use. *Nat Hum Behav* [Internet]. 2021 May 17 [cited 2021 Sep 24]; Available from: <http://www.nature.com/articles/s41562-021-01117-5>

14. Dickenson K, Richardson M, Kwan I, MacDowall W, Stansfield C, Brunton G, et al. Screen-based activities and children and young people's mental health: A Systematic Map of Reviews [Internet]. 2018. Available from: https://eppi.ioe.ac.uk/cms/Portals/0/PDF%20reviews%20and%20summaries/Systematic%20Map%20of%20Reviews%20on%20Screen-based%20activities_08.01.19.pdf?ver=2019-01-29-155200-517
15. Orben A, Dienlin T, Przybylski AK. Social media's enduring effect on adolescent life satisfaction. *Proc Natl Acad Sci*. 2019 May 21;116(21):10226–8.
16. Johannes N, Nguyen T, Weinstein N, Przybylski AK. Objective, subjective, and accurate reporting of social media use: No evidence that daily social media use correlates with personality traits, motivational states, or well-being. *Technol Mind Behav* [Internet]. 2021 Aug 12 [cited 2021 Sep 24];2(2). Available from: <https://tmb.apaopen.org/pub/accuracy-in-reporting-social-media-use>
17. Przybylski AK, Nguyen T, Law W, Weinstein N. Does taking a short break from social media have a positive effect on well-being? Evidence from three preregistered field experiments. *J Technol Behav Sci*. In Press;

28 September 2021