Jeffrey Howard[1]
Online Safety Bill Amendments - Evidence

1. *Summary*: This evidence criticizes the elimination of Category 1 services' adult-safety duties, which ought to be restored. The proposed user-empowerment tools, while valuable, are an inadequate replacement, and can be improved. Further, the amendments undermine the systemic focus of the Bill, reinforcing a leave-up/take-down binary and disincentivizing alternative tools such as algorithmic deamplification.

2. Proposed user-empowerment tools are insufficient because they reflect an unduly narrow view of how speech causes harm. While some speech directly harms its targets (as with racist abuse), other speech causes harm indirectly. User-empowerment tools will do little to mitigate the latter harms. For example, the danger of speech inciting hatred is that it risks inculcating bigoted views in susceptible audiences, potentially even inspiring offline discrimination or violence. *The ultimate victims of such speech need not ever see it, and their decision to opt-out will not protect them.*

3. The user-empowerment tools set the wrong default. Given the low value of the speech targeted by these tools, and the harm it can cause, it would be better for *exclusion* to be the default, such that users must *opt-in* to see such speech. This is likely to reduce substantially the number of people exposed to directly harmful speech such as abusive content and the promotion of self-harm.

4. The elimination of adult-safety duties neglects speech that is legal yet becomes harmful only when amplified (as with some misinformation). For such speech, criminal liability for speakers would be inappropriate, yet platforms have a moral obligation to reduce its reach, or at least announce clear policies. It will be replied that these duties pose a serious threat to freedom of expression. But this is mistaken. Despite widespread misreporting in the press on this point, the Online Safety Bill has never required platforms to remove legal speech that is harmful for adults. Instead past versions would have required platforms to specify, enforce, and risk-assess clear policies for such content—even if their stance is to allow it. Such measures pose no serious threat to users' free speech, while substantially bolstering large platforms' accountability for the governance of online content.

5. Taken together, the amendments reinforce a leave-up/take-down binary, and so undermine a systemic and flexible approach to governing harmful content. For example, the renewed focus on strictly enforcing terms of service overlooks content that is technically allowed by platforms, but is algorithmically deamplified/demoted (e.g., Meta demotes spam, certain misinformation, and "borderline" content). Demotion is an increasingly used tool of content governance, yet it remains seriously unclear how demotion policy fits within the amended architecture of the Bill.

---

[1] Associate Professor of Political Theory and Director of the Online Speech Project, School of Public Policy, University College London. The evidence provided reflects the views of the author.

Perversely, demotion may even be disincentivized by the Bill, depending on whether demoting content counts as "restricting users' access" to it. Requirements to enforce terms of service should be clarified so that platforms enjoy legal latitude to engage in a variety of demotion efforts.

10 December 2022