Clean Up The Internet—written evidence (FEO0038)

House of Lords Communications and Digital Committee inquiry into Freedom of Expression Online

Introduction

Clean Up The Internet is an independent, UK-based not-for-profit organisation concerned about the degradation in online discourse and its implications for democracy. We campaign for evidence-based action to increase civility and respect online, to safeguard freedom of expression, and to reduce online bullying, trolling, intimidation, and misinformation. We are delighted to have the opportunity to submit evidence to your inquiry.

The large social media platforms are now a central part of the UK public sphere. Questions of who is excluded and what content is removed from these conversations have so far focused on what content is removed, and on high profile "bans" or suspensions from (or indeed of) the platforms, most recently those of Donald Trump and Parler. These decisions certainly do raise important issues, in terms of whether the terms and conditions cited to justify such bans are applied more consistently, and – more fundamentally – whether it can be right for such decisions to be entirely in the hand of unaccountable (and in the case of the UK public sphere, foreign-owned) corporations. Measures within the Online Harms Bill to introduce regulatory oversight of the largest platforms' T&Cs and their implementation, and more routes of redress for users on the receiving end of sanctions, are to be welcomed and should start to address these concerns.

However, this is but one form of exclusion from social media which is currently threatening freedom of expression – and in our view though the most high profile it is not necessarily the most harmful. Individuals from all walks of life, but disproportionately those from marginalised or vulnerable groups, are currently less able to participate and express themselves freely on these same platforms, due to high levels of online intimidation and abuse which is having a significant silencing effect. This restriction on freedom of expression may be less visible than that which arises due to an active decision to remove a piece of content or ban a user – however it is at least as serious, and the evidence suggests it affects more people.

We see the government's Full Response to the Online Harms White Paper Consultation, finally published on 15 Dec 2020, and its forthcoming Online Safety Bill, as an opportunity to protect and enhance freedom of expression online. However because one person's "freedom" can adversely impact on others, we believe that improvements to the current proposals are needed. Large platforms should be required by the regulator to demonstrate that they are ensuring that their platforms are taking adequate action to ensure that all individuals - and in particular those with protected characteristics - are equally able to participate, and aren't being bullied off or silenced. Regulatory action

should seek to ensure that a right to freedom of expression online is enjoyed equally by everyone, and that the shift to an online public sphere doesn't perpetuate, or even exacerbate, existing gaps in who participates in democratic debate.

The large social media platforms are now a critical part of the "public sphere". This means critical decisions about freedom of expression online currently sit with under-regulated, foreign-owned private companies

Theorists of the public sphere and its role in a healthy democracy trace its origins to face-to-face gatherings in the salons of Paris or the coffee houses of London. However the 21st century public sphere has been moving online for some time now, and the Covid-19 pandemic has accelerated this trend.

The vast majority of MPs use social media platforms owned by Twitter, Facebook and Alphabet/Google both to disseminate information and to debate ideas. Journalists rely on social media both to get information and to break their stories. For ordinary UK citizens, these platforms are a major source of political information, and a major forum for political discussion.

The platforms derive significant commercial benefits from their dominant role in the public sphere. Political debates generate content, users, and "eyeballs", which in turn generate advertising revenue. However the platforms are currently not designed or incentivised to support the kind of public sphere which a healthy democracy requires. Bluntly the platforms do not treat users as their customers, to be looked after and to meet their needs, but as a product to be delivered to advertisers. Their increasing role in the public sphere has delivered profits, but has not inculcated a sense of responsibility to ensure their platforms promote healthy, constructive democratic debate. In the absence of effective regulation, decisions about the design and operation of social media platforms have been left to foreign-owned and democratically unaccountable corporations. The result has been a degeneration of the quality of democratic debate, and new threats to freedom of expression.

Left unregulated, the platforms favour a laissez-faire approach because this keeps overheads low and because harmful speech and fake accounts have been found to generate user engagement which in turn generates advertising revenue. This includes taking a minimalist and reactive approach to enforcing their own rules, which tends to occur only sporadically when confronted with a sudden need to be seen to respond in the face of negative publicity. This approach leads to a significant amount of inconsistency, where abusive behaviour and fake accounts are prevalent, but do occasionally get tackled when a particularly egregious example is highlighted by civil society, journalists or politicians.

Clean Up The Internet therefore welcomes the government's commitment to introduce greater oversight of these companies. We believe that a differentiated approach to the largest platforms, termed "Category One" platforms in the

government's proposals, does make sense given their dominance of the public sphere (although regulation of "Category Two" platforms still needs to be sufficiently robust to address the different harms and issues which arise in these smaller spaces).

Decisions about about account suspension and content moderation on public sphere platforms should be subject to regulation

An absence of regulation means important decisions about the acceptable bounds of free speech are left to private companies. Given the importance of such decisions to the health of the UK public sphere, a degree of regulatory oversight by a UK body is urgently required. Clean Up The Internet therefore strongly supports the government plan to include within the Online Safety Bill measures to require greater rigour, transparency, accountability, and consistency in platforms' terms and conditions, including what "legal but harmful" behaviour is and isn't permitted. However, we have a couple of significant concerns about the detail of how they implement this.

Firstly, for this regulatory oversight to be effective in promoting freedom of expression, it will be important that the "priority categories of legal but harmful material" are got right. At present the process for defining categories lacks detail. The government says that it will define "priority categories of legal but harmful material in secondary legislation (e.g. content promoting self-harm, hate content, online abuse that does not meet the threshold of a criminal offence, and content encouraging or promoting eating disorders). Ofcom will be required to provide non-binding advice to the government on what should be included in that secondary legislation."

There's a logic to the detail being left to secondary legislation, and to Ofcom providing expert input, but a lot more explanation is needed on the process. It will be important that the process includes due regard to experts, evidence, and the experience of groups directly affected by harms e.g. groups representing individuals who are targeted by online abuse. It will also be important that the impact of different online harms on freedom of expression is considered in a rounded way, including the impact on levels of self-censorship, silencing and exclusion.

Secondly, the government proposes that the regulator will tackle "legal but harmful" activity on "category one" platforms by holding those platforms to account for implementing "their own terms of conditions''. In some ways this makes sense, in that it focuses regulation at the level of the outcome (that the harm is addressed) rather than dictating to platforms the exact approach they take to deliver that outcome. However, there's a risk that this could in effect enshrine in law a privatisation of responsibility for setting online standards - platforms would no longer be allowed merely to "mark their own homework", but instead would be authorized to set the homework and in return be given a safe harbour.

The government says that the regulator will be able to require that terms and conditions explicitly address the priority harms, but little detail is offered as to how the regulator would assess whether the terms and conditions are adequate, and what sanctions they would have if they assess that they are inadequate. Getting this mechanism right will be critical if this "terms and conditions" approach is to succeed. This should include a requirement that terms and conditions, including moderation policies, consider freedom of expression in a rounded way including considerations of silencing and exclusion.

The accessibility and inclusivity of the public square matters to freedom of expression just as much as what people are able to say once they've entered it

A healthy public sphere depends on civil liberties such as freedom of opinion, expression and assembly. For it to be a truly *public* sphere, these freedoms need to be equally enjoyed by everyone.

Whilst online abuse, trolling and threats can be experienced by anyone, abuse is directed disproportionately either at those in the public eye (such as celebrities and politicians) or at individuals who are already vulnerable and under-represented in political debate. This includes groups with protected characteristics under the Equality Act. One impact of this abuse is to have a silencing effect: people are bullied off social media platforms altogether, or self-censor on certain topics.

A study conducted by Glitch and the End Violence Against Women Coalition during summer 2020 found that abuse directed at women appears to have worsened during the pandemic.[1] An overwhelming majority of respondents said they had modified their behaviour online following incidents of online abuse, with as many as 82.5% of Black or minoritised respondents reporting this impact. Research by Amnesty International in 2018 included polling which found that 78% of British women didn't believe Twitter to be a place where they could share their opinion without receiving violence or abuse.[2] In 2019 several female MPs cited social media abuse as a factor in their decision to step back from politics.[3]

At present the cumulative impact of some users' "free speech", and the way in which the design and operation of platforms (e.g. the design of content prioritisation algorithms; a permissive approach to anonymity and fake accounts) amplifies such speech combine to exclude other individuals from the online public sphere. This is detrimental to the freedom of expression of effected individuals, and exacerbates and perpetuates the under-representation of important sections of the population from democratic debate.

---

1   https://fixtheglitch.org/wp-content/uploads/2020/09/Glitch-COVID-19-Report-final-1.pdf
2   https://www.amnesty.org.uk/press-releases/toxic-twitter-failing-women-letting-online-violence-thrive-new-research
3   BBC News - Women MPs say abuse forcing them from politics - https://www.bbc.co.uk/news/election-2019-50246969

Clean Up The Internet therefore recommends that the forthcoming Online Safety Bill gives Ofcom powers to:

- Take a rounded approach to freedom of expression, which considers issues of silencing and exclusion, and the potential impact of an action or inaction on potential users as well as current users.

- Require social media platforms to report on the diversity of their user base

- Require large social media platforms to have due regard to the inclusion of groups with protected characteristics in decisions about the design and operation of their platforms

- Impose sanctions upon large social media companies which fail to take adequate steps to protect the freedom of expression of users with protected characteristics

A more proactive approach to managing anonymity is required to maximise freedom of expression

The ability to use social media anonymously can be an important safeguard of freedom of expression online for example for whistleblowers, political dissidents, or someone fleeing an abusive partner. However, there is overwhelming evidence that anonymity also fuels harmful behaviour, including hateful disinformation and abuse which undermines the quality of debate and threatens the freedom of expression of other users. Opinion polling which we commissioned, carried out by YouGov, found strong public agreement with the view that anonymity fuels negative online behaviour.[4]

At present the large social media platforms take a laissez-faire and self serving approach to anonymity and identity management. They minimise their overheads and maximise the number of users they are able to claim for the purposes of selling advertising. Clean Up The Internet has conducted extensive research into the ways in which anonymity is used and abused on social media platforms.[5] For example our analysis identified that on Twitter anonymous accounts were disproportionately likely to be promoting the conspiracy theory linking 5G phone masts to Coronavirus. We believe there is considerable

4   https://www.cleanuptheinternet.org.uk/post/new-opinion-poll-83-of-brits-thinks-anonymity-makes-people-ruder-online
5   See for example:
    https://www.cleanuptheinternet.org.uk/post/new-opinion-poll-83-of-brits-thinks-anonymity-makes-people-ruder-online
    https://www.cleanuptheinternet.org.uk/post/new-research-anonymous-twitter-accounts-fuelled-the-spread-of-coronavirus-5g-conspiracy-theories
    https://www.cleanuptheinternet.org.uk/post/a-breakdown-of-election-manifesto-promises-on-laws-to-stop-online-harms
    https://www.cleanuptheinternet.org.uk/post/academic-research-about-anonymity-inauthenticity-and-misinformation

potential for different approaches to managing anonymity which could safeguard its legitimate uses whilst restricting its abuse.

An example of such an approach would be to:

- Give all users the option of robustly verifying their own identity

- Make it transparent which users have verified their identity, and which have chosen not to and so are anonymous or may be using a pseudonym

- Give all verified users the ability to manage the level of interaction which they have with unverified users, including an option to block unverified users *as a category* given the association between anonymity and a higher risk of abusive behaviour

We are concerned that the government's current stated intention to "not put any new limits on online anonymity" may mean opportunities are missed to reduce the role which anonymity plays in harming freedom of expression – whilst also safeguarding legitimate uses.

We recommend that within the Online Safety Bill the regulator should have the power to define risk factors such as anonymity and identity deception. Platforms should be required to demonstrate that their platform has been designed to mitigate such risks. Where platforms fail to demonstrate that they are acting to mitigate abuse of anonymity and identity deception, we suggest that a possible sanction could be that they are made liable, as a kind of "publisher of last resort", for harmful content which cannot be attributed to a user.

Conclusion

To ensure a full range of voices are able to be heard and a full range of opinions are expressed (i.e. maximum freedom of expression for maximum people), a healthy online public sphere needs to be rules-based. Clean Up The Internet sees much to welcome in the government's plans for an Online Safety Bill. We welcome an end to failed self-regulation, and the plan to appoint Ofcom as the independent regulator, underpinned by statute. We welcome, in as far as they go, the specific proposals relating to freedom of expression, including requiring companies to assess the impact of their design on their users' rights, and improved transparency and rights of redress in relation to moderation decisions.

However, for the Online Safety Bill to fulfil its potential to improve Freedom of Expression the proposals need strengthening. Most importantly, for the large, public sphere platforms ("Category One", in the language of the proposals) consideration must be given to the impact of the platform's decisions on the freedom of expression of potential users not just existing users. The regulator's powers to shape the platforms' Terms and Conditions must include the power to require an approach to freedom of expression which includes consideration of the ability of groups with protected characteristics to participate fully and equally in the public sphere. The regulator must also have the power to identify risk

factors which can limit freedom of expression, such as anonymity, and to require platforms to mitigate them.

January 2021