

Written evidence from the Carnegie UK Trust (FOE0078)

1. We welcome the Committee's inquiry into Freedom of Expression and the opportunity to submit evidence. Our response is limited to the questions around hate speech and equal speech and draws on the thinking that also informs our work on a statutory duty of care for online harm reduction.
2. We would be happy to provide further information on our work in writing or to discuss it with Committee members at a future evidence session.

About our work

3. In early 2018, Professor Lorna Woods (Professor of Internet Law at the University of Essex and member of the Human Rights Centre there) and former civil servant William Perrin started work to develop a model to reduce online harms through a statutory duty of care, enforced by a regulator. The proposals were published in a series of blogs and publications for Carnegie and developed further in evidence to Parliamentary Committees¹. In April 2019, the government's Online Harms White Paper² proposed a statutory duty of care enforced by a regulator in a variant of the Carnegie model. In December 2019, we published a draft bill³ to implement a statutory duty of care regime, based upon our full policy document of the previous April⁴. Professor Woods also published a comprehensive paper on the statutory duty of care and fundamental freedoms, including freedom of expression.⁵

Hate Speech

4. The Committee has identified, as can well be expected from a Committee with a remit covering all human rights, that a range of interests are in issue. The question of how to strike an appropriate balance between speech and other rights - including the ability of one person's speech amplified through social media to suppress other speech - is at the heart of the current debate.
5. "Hate speech" is a term much used but not well-defined. It can cover a range of speech from the comparatively mild insults based on a group characteristic to speech that dehumanises targeted individuals, threatens their ability to participate in society and undermines fundamental democratic principles; it is not limited to speech criminalised under domestic law. This recognition is significant in the context of freedom of speech because while an interference with hate speech may in principle be justified (whether under Article 10 ECHR or Article 19(3) ICCPR), any such restriction must be proportionate. The imposition of criminal liability, especially for offences that carry a custodial sentence, is harder to justify than civil action; the threshold in terms of the severity of hatred expressed would be higher. It is not open to States to shy away from

¹ Our work can be found here: <https://www.carnegieuktrust.org.uk/project/harm-reduction-in-social-media/>

² <https://www.gov.uk/government/consultations/online-harms-white-paper>

³ <https://www.carnegieuktrust.org.uk/publications/draft-online-harm-bill/>

⁴ https://d1ssu070pg2v9i.cloudfront.net/pex/carnegie_uk_trust/2019/04/08091652/Online-harm-reduction-a-statutory-duty-of-care-and-regulator.pdf

⁵ <https://www.carnegieuktrust.org.uk/publications/doc-fundamental-freedoms/>

this question, as positive obligations under Article 8 ECHR and obligations in relation to the prevention of discrimination require the States to take some action (not necessarily by creating criminal offences⁶) to protect victims' rights.

6. It seems that the current system, which focuses on criminal hate speech, is not functioning. Within society there is no common understanding of what is hateful, as illustrated by the fact the large number of non-crime hate incidents⁷ that are being recorded by the police. This development may be driven by the possible changes in understanding as to what is acceptable, but also because social media gives a platform for groups which might not otherwise come across one another to communicate. Such communications are often retained on the platform and can be shared widely. Although there are civil actions which may provide partial remedies in this area (e.g. defamation; misuse of private information; harassment as a civil action), none target the ill directly. Moreover, relying on victims to assert rights faces considerable difficulties, not least expecting any such victims to have the financial and psychological resources to bring such a case.
7. The Law Commission is reviewing hate speech as a criminal offence and seeking to rationalise the protected characteristics. This is commendable, as there are currently lacunae in the law leaving some groups who receive considerable abuse relatively unprotected. We note that the CPS Guidance on issues in Social Media Offending links to the Violence Against Women and Girls (VAWG) Strategy and Domestic Abuse. These are important and relevant considerations. The Social Media Guidance does not take express cognisance of the abuse received by women in public life (including journalists) and the gendered nature of that, its silencing effect and the impact of that in public life. This is not to suggest that such cases should automatically be prosecuted but that the guidance around the issues to be considered do not currently reflect this aspect.
8. We question whether there is, in the light of changing societal expectations, the need for a broader review to take into account non-criminal mechanisms. One possibility is to review the range of civil actions available. Not all hate speech that falls short of the criminal threshold falls neatly into a civil action (and this is a gap that the Law Commission review of hate crime will not address). We suggest that there be a consideration of other possible mechanisms. We are not convinced that Public Space Protection Orders, as currently formulated, would work for online spaces. We note the additional point that Scotland uses Antisocial Behaviour Orders and in the online context a uniform approach across the UK would be helpful to remove the risk of there being internal conflict of laws issues. An alternative mechanism is the Community Protection Notice (CPN). Two points should be noted: failing to comply with a CPN is a criminal offence and consideration should be given to whether this would always be proportionate. There could be a difference between someone continuing to abuse a

⁶ Note that Article 20(2) ICCPR specifies that “[a]ny advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence shall be prohibited by law”. The ECHR says nothing about hate speech expressly but the American Convention on Human Rights at Article 13(5) provides that any propaganda for war and advocacy of racial hatred “shall be considered as offenses punishable by law”.

⁷ See here for the police approach to non-crime hate incidents: <https://www.app.college.police.uk/app-content/major-investigation-and-public-protection/hate-crime/responding-to-non-crime-hate-incidents/>

victim with hateful speech and the situation online of failing to take down the offending post (a CPN could be a useful mechanism to ask platforms to take down content if they have not already done so in relation to their community standards).

9. A second point is the authorities that currently may issue a CPN. On the whole, these are bodies that do not have experience of the victims' perspective nor the importance of freedom of speech. Other bodies (eg Victims Commissioner, the EHRC or the Children's Commissioner) exist who could take on such a role or participate in the process (with appropriate resources). We note in this context the Law Commission's suggestion of a Hate Speech Commissioner.
10. In the online context, we are reminded of Ofcom's existing powers under s. 128 Communications Act with regards to persistent misuse of an electronic communications network or electronic communications services (as defined in the Act) where that misuse causes or is likely to cause "another person unnecessarily to suffer annoyance, inconvenience or anxiety", a test which is likely to be satisfied in ongoing abuse of the same person, but is less good a fit for one-off behaviour (which may be below the threshold for intervention unless it is serious one-off behaviour in which case it would be better dealt with by relevant criminal provisions), or a person who engages in ongoing low-level hate speech as against different people (but who may share the same protected characteristics). Again, the existence of a notification could be a tool to ensure content is taken down rather than relying on the platform's individual interpretation of its own terms and conditions. Of course, it would be preferable if there were some form of regulation of the platforms themselves. Even if there were such a system, some consideration should be given as to how rulings and understanding as to the acceptability of speech from the domestic legal system is understood and respected by the platform operators themselves.

Equal Speech

11. There is a difference between who holds the right to freedom of expression – which should be held equally by all – and to the scope of the speech protected. Some speech, by virtue of Article 17 ECHR, would fall outside Article 10. The freedom of speech is a limited right; interference with the right may be justified (under strict conditions). In this context, it should be noted that individuals should enjoy their rights equally, that there should not be discrimination in the protection awarded to rights. Hate speech has a silencing effect on those towards whom it is directed and the State is obliged to protect their rights.
12. This point is reaffirmed by the jurisprudence on Article 8 ECHR which protects personal autonomy, identity and integrity, as well as human dignity.⁸ Where a particularly important facet of an individual's existence or identity is at stake, the margin of appreciation accorded to a State will, in general, be restricted⁹ and the Court has recognised that the State is under an obligation to secure not just an individual's physical integrity but also that person's psychological integrity.¹⁰ In *Beizaras and*

⁸ *Pfeifer v Austria* App no 12556/03 [2007] ECHR 935 [33]

⁹ *Dickson v United Kingdom* [GC] App no 44362/04 ECHR 2007-V.

*Levickas v. Lithuania*¹¹, for example, the decision by the prosecutor's office not to initiate criminal proceedings against homophobic posts on Facebook which included calls for violence, constituted a violation of the State's obligations under Article 8 (in conjunction with Article 14).¹² State obligations have arisen in relation to the verbal harassment of a boy with disabilities,¹³ and with the negative stereotyping of the Roma in State-sponsored schoolbooks¹⁴.

13. The committee asks if hate speech law needs updating to reflect changing social values but we would urge it to consider the role of the mechanisms which do much to propagate or undermine these values and thus drive change. To paraphrase McLuhan – social media may be the message. While some debates may be characteristic of recent years, for instance, JK Rowling and the trans rights debate, new technologies can amplify older hatreds – as seen in the anti-Semitic and misogynistic abuse of Margaret Hodge MP¹⁵. These examples highlight the risk of the silencing effect of online abuse, more often experienced by those in minoritised groups. These modern media reveal that the criminal hate speech regime is not working – even with the revisions proposed by the Law Commission and even if the CPS and police had far greater resources it could still not cope. At Carnegie UK Trust we proposed a regulatory regime to address harms arising from social media based on a statutory duty of care to reduce harm enforced by a regulator. For the hate speech regime to work a regulatory regime is required as well as a criminal one. The government's online harms policy picks up many of our suggestions, though there are questions of scope around low-level abuse which will need to be examined. The government has appointed OFCOM to administer the regime, a regulator experienced in balancing rights and addressing hate speech.
14. At the time of writing, two days after the government has published its final proposals for online safety/harms. We are still working through them. In particular, the government's approach creates a curious relationship between the criminal hate speech regime and regulation that we are seeking to tease out, especially as it pertains to equal protection of the right to freedom of expression. We should be delighted to discuss this with the committee in the New Year.

18/12/2020

¹⁰ *Glass v United Kingdom* App no 61827/00 ECHR 2004-II [74]–[83]; *KU v Finland* (Application no. 2872/02), judgment 2 December 2008, para 43.

¹¹ *Beizaras and Levickas v. Lithuania* (Application 42188/15), judgment 14 January 2020.

¹² More generally, see *Király and Dömötör v. Hungary* (application 10851/13), judgment 17 January 2017.

¹³ *Dorđević v Croatia* App no 41526/10 ECHR 2012-V [87]–[88].

¹⁴ *Aksu v Turkey* [GC] App nos 4149/04 and 41029/04 ECHR 2012-I; on the facts Aksu lost but the principle was recognised.

¹⁵ Margaret Hodge calls for ban on social media anonymity; Guardian December 2020 – Hodge received 90,000 mentions over three days <https://www.theguardian.com/society/2020/dec/06/margaret-hodge-calls-for-ban-on-social-media-anonymity>