

Online Safety Bill

House of Lords

Committee Stage Amendments

April 2023

Illegal Content Duties and Safeguarding Freedom of Expression

1. Whilst JUSTICE welcomes the increased regulation of certain types of illegal content online, in particular child sexual abuse and exploitation content, we also recognise the need, identified in the Government's response to the Online Harms White Paper, to protect freedom of expression by ensuring that companies avoid taking an "*overly risk-averse approach to the identification and removal or material likely to be illegal.*"[1] It is JUSTICE's view that the Bill, in its current form, does not sufficiently safeguard against this risk of over-moderation by online companies.

2. In particular, the Bill does not account for the difficulty that in-scope providers are likely to face in moderating certain content, and moreover doesn't account for the fact that certain forms of illegal content are inherently difficult to moderate. As stated by Richard Wingfield, Head of Legal, at Global Partners Digital, in his evidence before the Select Committee on Communications and Digital:

> "*Making decisions about what is illegal speech is incredibly difficult. It takes time to gather evidence and to talk to witnesses, and it is most likely there will be a trial at the end of it, yet we are asking content moderators to understand our legal system and make decisions in minutes about whether somebody's speech is illegal or not*."[2]

3. JUSTICE is particularly concerned about the inclusion of content that amounts to an offence under section 5 of the Public Order Act ("**POA"**) within the category of priority illegal content to be regulated by the Bill.[3] Section 5 of the POA covers "*threatening or abusive words or behaviour, or disorderly behaviour*" that is likely to cause "*harassment, alarm or distress*." As recognised by the Joint Committee on Human Rights, "*it is hard to*

[1] HM Government, 'Online Harms White Paper: Full Government response to the consultation', 2020, p.31.
[2] House of Lords Select Committee on Communications and Digital, 'Corrected oral evidence: Freedom of expression online', 2020.
[3] Schedule 7.

see how providers, and particularly automated responses, will be able to determine whether content on their services fall on the legal or illegal side of this definition."[4]

4. Given the fine line between what constitutes abusive as opposed to merely offensive material, as well as the difficulties involved in determining whether or not content is likely to cause someone sufficient "*alarm or distress*," it is our view that placing a duty on online service providers to prevent users encountering content amounting to a section 5 POA offence will result in the disproportionate removal of legitimate content that does not meet the threshold for the offence. This risk is exacerbated by the fact that when determining whether content is illegal content on not, providers need only have "*reasonable grounds to infer*" that the elements necessary for the commission of the offense in question are satisfied.[5]

5. Moreover, the Bill designates as priority illegal content the inchoate versions of section 5 POA offences.[6] This means that under the Bill content that encourages, for instance, disorderly behaviour likely to cause alarm would also have to be moderated and removed by online companies. It is not difficult to imagine the broad range of content that could meet that definition given that companies only need to have reasonable grounds to think the threshold for the offence is met. For instance, content that portrays protest activity in a positive light, or which shows protest activity which could be viewed as disorderly, without directly condemning that activity.

6. Whilst the Bill does state that in making judgments as to whether content is illegal content or not, providers would have to consider that there are reasonable grounds to think that the mental element[7] of the offence is met,[8] given the low bar set by "*reasonable grounds*," the difficulties in determining the state of mind of any particular poster, particularly where

[4] Letter from Harriet Harman MP, Chair of the Joint Committee on Human Rights to the Secretary of State for Digital, Culture, Media and Sport, 19 May 2022.
[5] Clause 170.
[6] Schedule 7. Inchoate offences occur where an individual does not commit the substantive offence but instead commits an offence by encouraging, inciting, assisting, or attempting the commission of the substantive offence.
[7] The mental element of an offence is the state of mind an individual must have when carrying out the conduct prohibited by the offence. For instance, in the case of inchoate section 5 POA offences the individual encouraging the commission of the substantive section 5 offence must also intend or believe that the substantive offence would be committed.
[8] Clause 170(6).

automated technology is being used,[9] and the consequences that attach to failing to fulfil an illegal content duty,[10] online providers will have a strong incentive to infer this once the other elements of the offence are reasonably satisfied.

7. The inclusion of section 5 POA offences in the category of priority illegal content therefore has potentially significant implications for the Article 10 ECHR rights of users. The European Court of Human Rights ("**ECtHR**") has repeatedly warned against the collateral effect of measures designed to prevent the dissemination of illegal content online and has consistently reiterated that to be ECHR compliant domestic law must strictly target the illegal content in question.[11] Provisions which encourage an overly risk-averse approach to content removal, resulting in legitimate content being removed, may therefore fall short of the UK's obligations under the ECHR. **JUSTICE therefore considers that section 5 of the Public Order Act should be removed from the category of priority illegal content, set out in Schedule 7 of the Bill**.

Amendment 1

Page 202, leave out line 22.

Member's explanatory statement

This would remove offences under s.5 of the Public Order Act from the category of priority illegal content as set out in Schedule 7.

8. JUSTICE also considers that in order to protect users from overcautious content moderation by service providers, the Government should do more to strengthen the safeguards designed to protect freedom of expression within the Bill. The most obvious of these safeguards are the duties about freedom of expression contained in Clause 18 and 28 of the Bill, which would require all services to "*have particular regard to the importance*

[9] See i.e. Microsoft, 'Online Safety Bill – Parliamentary Briefing from Microsoft' 2022; Independent Reviewer of Terrorism Legislation, 'Missing Pieces: A Note on Terrorism Legislation in the Online Safety Bill' 2022.

[10] Under the Bill OFCOM would be able to issue those who fail to comply with a duty of care with a fine of up to £18 million or 10% of annual global turnover, whichever is higher. In cases of continued non-compliance OFCOM would be able to take measures to disrupt a company's business activities in the UK, including blocking access to services. See Chapter 6 (Enforcement Powers).

[11] See e.g., *Engels v Russia* (2020) App. No. 61919/16; *OOO Flavus and Others v Russia* (2020) App. Nos. 12468/15,23489/15 and 19074/16.; *Vladimir Kharitonov v Russia* (2020) App. No. 10795/14; *Cengiz and others v Turkey* (2016) App. Nos. 48226/10 and 14027/11.

of protecting users' right to freedom of expression within the law" when deciding on and implementing their safety measures and policies.

9. It is JUSTICE's view that the requirement to "*have regard to the importance of*" protecting freedom of expression, rather to simply protect it, significantly dilutes this duty. Moreover, JUSTICE recognises the concern highlighted by the Joint Committee on Human Rights, that the obligation to have regard to "*freedom of expression within the law*" lacks clarity given the number of interrelated legal duties that affect freedom of expression.[12] As John Howell MP, noted during October's debate on online harms, the ECHR is a "*key pillar*" for protecting the right to freedom of expression online.[13] With this in mind, **we recommend that the duties about freedom of expression in the Bill be amended to specify that when deciding on, and implementing, safety measures and policies, in-scope services must uphold the right to freedom of expression as protected by Article 10 ECHR.**

10. Under Article 10 ECHR any interferences with freedom of expression must be necessary and proportionate to achieving a legitimate aim.[14] Making explicit reference to Article 10 of the ECHR would impress on service providers the need to ensure that the measures they impose in complying with their duties under the Bill are proportionate and necessary restrictions on the freedom of expression. This would go some way to guarding against overzealous moderation, as it would make clear that the measures imposed by service providers should be no more restrictive than is necessary to, for instance, protect users from a particular kind of illegal content.

Amendment 2

Page 20, line 31, leave out "have particular regard to the importance of protecting users' right to freedom of expression within the law." and insert -

"uphold users' right to freedom of expression as protected by Article 10 of the European Convention of Human Rights."

---

[12] Letter from Harriet Harman MP, Chair of the Joint Committee on Human Rights, to Nadine Dorries MP, Secretary of State for Digital, Culture, Media and Sport, 19 May 2022.

[13] HC Deb 26 October 2022, vol 721, col 168WH.

[14] Such aims include protecting national security, preventing disorder or crime, and protecting health or morals. See European Court of Human Rights 'Guide on Article 10 of the European Convention on Human Rights,' (2022)

Member's explanatory statement

This would make explicit that all regulated user-to-user services when deciding on and implementing safety measure and policies, have a duty to uphold users' rights as protected by Article 10 of the European Convention of Human Rights.

Amendment 3

Page 21, line 4, leave out "within the law" and insert -

"as protected by Article 10 of the European Convention of Human Rights."

Member's explanatory statement

This would make explicit that when deciding on safety measures and policies Category 1 services must carry out an assessment of the impact of these measures and policies on users' right to freedom of expression as protected by the European Convention of Human Rights.

Amendment 4

Page 30, line 31, leave out "have particular regard to the importance of protecting the rights of users and interested persons to freedom of expression within the law." and insert -

""uphold users' and interested persons' right to freedom of expression as protected by Article 10 of the European Convention of Human Rights."

Member's explanatory statement

This would make explicit that all regulated search services, when deciding on, and implementing, safety measure and policies, have a duty to uphold users' and interested persons' rights as protected by Article 10 of the European Convention of Human Rights.

11. Another safeguard which JUSTICE considers could be strengthened is the requirement concerning complaints procedures. Under the Bill companies would have a duty to operate complaints procedures to allow, amongst other things, complaints from users whose content is taken down on the basis that it is illegal content or who are suspended from the service as a result of content which the provider considers to be illegal content.[15] However, whilst this requirement is designed to offer users some protection from overzealous content moderation, by enabling users to challenge the removal of their content, the Bill leaves it entirely up to online companies how they chose to operate their complaints

[15] Clause 17 and 27.

procedure and contains no provisions for the kinds of redress that might be available to those whose content is wrongly removed.

12. To ensure that online service providers implement robust complaints processes, the face of **the Bill should contain requirements as to what basic features these processes should have.** At a minimum, this should specify that where content is taken down wrongfully, the company must have a process for reinstating that content (or user) within a particular time-frame.

Amendment 5

Page 18, line 42, after subsection (2) insert -

(3) For complaints of the kind specified in subsections (4)(c), (d) and (e) and 5 (c) and (d) a duty to operate a complaints procedure that -

(a) Allows users to have actions taken against them reviewed and to have the outcome of that review decided within a period of time specified in the terms of service.

(b) Ensures that where a review of an action taken against a user finds in favour of that user, the action against that user is reversed within a period of time specified in the terms of service.

(c) is easy to access, easy to use, and transparent.

Member's explanatory statement

This would place a duty on user-to-user services to operate a complaints procedure that allows users to challenge the removal of their content, or suspension of their account, and to have their content or account reinstated within a specified time period if their challenge is successful.

Amendment 6

Page 29, line 22, after subsection (2) insert -

(3) For complaints of the kind specified in subsections 4(c) and (d) and 5(c) a duty to operate a complaints procedure that -

(a) allows interested persons to have actions taken against them reviewed and to have the outcome of that review decided within a period of time specified in the terms of service.

(b) ensures that where a review of an action taken against an interested person finds in favour of that interested person, the action against them is reversed within a period of time specified in the terms of service.

(c) is easy to access, easy to use, effective and timely, and transparent.

Member's explanatory statement

This would place a duty on search services to operate a complaints procedure that allows interested persons to challenge actions taken that affect content relating to them, and to have these actions reversed within a specified time period if their challenge is successful.

13. Moreover, the Bill should make explicit that when determining whether content is illegal content in the context of an appeal or complaints process, a service provider should have to have more than "*reasonable grounds to infer*" that the elements of the offence are met. The Bill already suggests that judgements about the status of content are likely to differ based on whether the judgement is made by a human moderator or an automated system, given the information likely to be available in each of those cases.[16] It is consistent with this position that a higher standard should apply in the context of appeals where more time is given over to considering that content and more information is likely to be available. **The Bill should therefore specify that when determining whether content is illegal content for the purpose of an appeal or complaints process, the provider must be satisfied that its more likely than not that the elements of the offence are met.**

Amendment 7

Page 145, line 26, after subsection (7) insert -

(8) In the case of judgements made in the context of a review process in compliance with section 17 or section 27, the provider must be satisfied that on the balance of probabilities the content in question is illegal content.

(9) For a provider to be satisfied that it is more likely than not that content is illegal content a provider must consider that -

(a) On the balance of probabilities all the elements necessary for the commission of the offence, including mental elements, are present or satisfied, and

(b) On the balance of probabilities no defence to the offence could be relied upon.

Member's explanatory statement

This would require providers when making a judgement about whether content is illegal content in the context of a review against action taken in relation to that content, to be satisfied that it is

[16] Clause 170.

more likely than not that the elements of the offence are met, and that no defence to the offence could be relied on.

JUSTICE

April 2023