RESPONSE BY SMITH, MR G TO OFCOM CALL FOR EVIDENCE ON AVMSD

24 September 2020

Smith, Mr G is a solicitor in private practice in London. He is the editor and main author of the legal textbook Internet Law and Regulation (Sweet & Maxwell, 5th edition 2020). He also writes the Cyberleagle blog.

This submission is made in Mr Smith's personal capacity. Views expressed are not attributable to the law firm at which he works or to any of its clients.

Introduction

1. Ofcom has issued a call for evidence which, although it mainly relates to its role as regulator in relation to the revised Audiovisual Media Services Directive, touches on its anticipated future role under the government's Online Harms proposals:

> "If confirmed, we will build on VSP [Video Sharing Platform] regulation to inform our approach to regulation of services under the online harms regime." and
>
> "Many of the provisions of the AVMSD pertaining to regulation of VSPs complement the Government's stated proposals for online harms regulation."

2. Ofcom goes on:

> "The evidence we gather will directly inform the drafting of our guidance for VSPs, but also inform our potential future role as the online harms regulator, should the Government confirm Ofcom taking on functions in this area."

3. In the light of the above it seems appropriate to explore some of the kinds of decisions that the government may anticipate intermediaries being obliged to make, and the role of Ofcom in relation to the supervision and enforcement of such decision-making.

4. At this stage the boundaries and contours of the proposed duty of care, and the precise role of Ofcom in supervising platforms' discharge of their duty, have yet to be delineated.

5. The most recent formal iteration of the government's proposals – the Initial Consultation Response - stated, introducing the proposed 'differentiated' duty of care:

> "We will not prevent adults from accessing or posting legal content, nor require companies to remove specific pieces of legal content. The new regulatory framework will instead require companies, where relevant, to explicitly state what content and behaviour is acceptable on their sites and then for platforms to enforce this consistently."

6. That emphasis on intermediaries' freedom to decide on their own substantive rules contrasts, on the face of it, with the government's May 2018 Response to its consultation on the Internet Safety Strategy Green Paper. In that document it set out its requirements for substantive contents of platforms' terms and conditions:

> "The government has made clear that we require all social media platforms to have [*inter alia*]: Terms and conditions that provide a minimum level of safety and protection for users".

7. It is unclear whether that remains the government's policy. If it is, then it appears inevitable that the ostensible freedom of intermediaries under the differentiated duty of care to set the contents of their own Terms and Conditions[1] (for lawful content for adults) would be circumscribed.

8. In any event, the government's Initial Consultation Response suggests that Ofcom's remit would include the effectiveness of intermediaries' actions in enforcing their T&Cs:

> "Recognising concerns about freedom of expression, the regulator will not investigate or adjudicate on individual complaints. Companies will be able to decide what type of legal content or behaviour is acceptable on their services, but must take reasonable steps to protect children from harm. They will need to set this out in clear and accessible terms and conditions and enforce these effectively, consistently and transparently."

9. The reference to effectiveness invites the question: 'effectiveness in achieving what?'. If the answer is 'effectiveness in reducing harm'[2], it is difficult to see how Ofcom could remain aloof from considering concrete examples of user content. Even if Ofcom were to have no investigative or adjudicative role in response to an individual complaint asserting that the intermediary should have taken action, it appears inevitable that assessing the incidence of harm would involve considering examples of content posted by users and forming a view on whether, in discharging its duty of care, the intermediary should (or need not) have inhibited them in some way.

10. The same would be true of the incidence of unlawful content. It seems inevitable that Ofcom would be drawn into considering, retrospectively, whether given items of content were or were not unlawful. That initial step is unavoidable if, as the government has suggested, the proposed duty of care should apply differently to lawful and unlawful content.

11. It might be suggested that Ofcom's interest could be limited to assessing systemic processes, disconnected from substantive evaluation of user content. One example of such a process might be limits on re-dissemination of posts generally (amplification). Putting aside questions of damage to the reach of legitimate speech[3] raised by that kind of blanket obligation, it is unclear how effectiveness of such a process would be assessed. A wholly abstract, systemic approach would seem to necessitate regarding amplification as inherently detrimental, regardless of the content amplified. In short, effectiveness is

---

[1] It is hard to see what relevance terms and conditions could have to search engines, which the government confirmed in evidence to the Commons Home Affairs Committee on 13 May 2020 remain in scope.

[2] This would be the answer according to the Carnegie UK Trust harm reduction model.

[3] Limitations on reach engage the right of freedom of expression: "There is no dispute that freedom of speech and expression includes the right to disseminate information to as wide a section of the population as is possible. The wider range of circulation of information or its greater impact cannot restrict the content of the right nor can it justify its denial." *Bhasin and Azad v India*, Supreme Court of India, 10 January 2020.

assumed at the outset. Evaluating the prevalence of actual harm would suggest the need for empirical assessment of actual content and of the harm said to flow from it.

12. Whether Ofcom would seek to discover whether harm (however defined[4]) actually resulted from reading and viewing specific kinds of content, and in what cases, or would simply deem certain categories of content to be harmful, is unclear. It is noteworthy, *en passant*, that Ofcom in its 2018 survey (conducted with the Information Commissioner's Office) suggested to participants that 'bad language' was harmful[5].

13. Even in the case of deemed harm, it is unclear how effectiveness of measures taken by intermediaries could be evaluated without attempting to measure the prevalence of the content deemed to be harmful. That requires an objective assessment of whether specific content present on platforms does or does not fall within those categories. If bad language were to be deemed harmful, what does and does not constitute 'bad language'?

Why concrete examples?

14. The preceding discussion points to the need to consider concrete examples of user content. Where speech is concerned, as George Carlin was famously aware[6], it always comes down to specifics: what are the words that you cannot say? While context will often be important, the starting point has to be the actual words (and images) used.

15. The question for an intermediary subject to a legal duty of care will be: "are we obliged to consider taking steps (and if so what steps) in respect of *these* words, or *this* image, in *this* context[7]?"

16. If we are to gain an understanding of where the lines would be drawn, we cannot shelter behind comfortable abstractions. We have to grasp the nettle of concrete examples, however uncomfortable that may be.

17. That is important from the perspective not only of the intermediary, but of the user. From a rule of law standpoint, it is imperative that the user should be able to predict, in advance, with reasonable certainty, whether what they wish to say is likely to be affected by the actions of an intermediary seeking to discharge its duty of care.

18. How, then, would Ofcom approach various kinds of speech that might be found on an online sharing platform? The primary interest in posing concrete examples at this stage is not so much in the actual conclusion that Ofcom might reach about how they should be addressed according to some aspect of a platform's duty of care (although that is obviously significant), but in whether Ofcom would consider the examples to be within

[4] The White Paper and Initial Response have not indicated that the legislation would define harm. If it were to do that, a list of matters included in open-ended "harm" would be a different matter from an exhaustive definition. The September 2020 Law Commission Consultation Paper on "Harmful Online Communications: The Criminal Offences" contains a detailed analysis of different kinds of speech harms.
[5] The survey question was "Which, if any, of the following harmful things have you experienced on the internet". One of the listed items was "Bad language".
[6] "Seven Words You Can Never Say on Television".
[7] However, an intermediary will frequently not be in a position to evaluate context.

its jurisdiction and the platforms' duty of care at all; and if so, how it would go about determining whether the item triggered the duty of care.

Triggering a duty of care

19. 'Triggering' a duty of care is convenient shorthand, but conceals some complexity. It is easy to think of a duty of care only in terms of an obligation to remove, or take other action, in respect of some kind of reprehensible content or conduct. But non-removal and removal are two sides of the same duty of care coin: the one risks liability for inaction, compelling the other risks illegitimate interference with fundamental rights[8] (which ought in principle to be just as much a concern for a regulatory regime as failure to remove or take other action).

20. The first step in considering what is triggered by a duty of care, in the context of a duty to prevent conduct by third parties to each other, is whether there is an obligation on the intermediary to put itself in a position to make a decision in relation to any given item of content. That aligns closely to the question whether an intermediary's obligation is proactive or reactive, and if so whether a proactive obligation applies (or is capable of applying[9]) only to certain identifiable kinds of content or behaviour.

21. Assuming that the item of content in question has, by whatever route, come to the attention of the intermediary, it is then faced with a two-stage question: (1) is this content that I have to take a substantive decision about? (2) if so, what should that decision be?

22. If the answer in any given case is that the speech in question would be out of scope of the duty, then the answer to the first question is 'No decision is necessary'. In order for the first question to be meaningful, it must be possible to come to that conclusion with certainty. Put another way, what clear and precise rule would be applied to enable a clear conclusion to be reached that the item of speech lies on one side or the other of the jurisdictional line?

23. If the answer to the first question is 'Yes, a decision is necessary', how should the second question be approached? There will (or should be) many cases in which the answer would be clear: the content does, or does not, merit some kind of action. But it is likely that in many (perhaps most) cases the answer will not be obvious.

24. That prediction stems from at least two factors: (1) the great variety of different ways in which people themselves react to content and (2) the vast range of subjective opinions that people hold about how other people will, or ought to, respond to content. In cases where the answer is not obvious, where should the default lie? Action, or no action?

---

[8] Voluntary action by a platform, not being state action, should in principle not be regarded as involving an interference with fundamental rights. The position is different where action is taken pursuant to a duty of care imposed by the state.

[9] The point is routinely made that an intermediary cannot proactively scan for specific kinds of content defined only by subject matter without scanning all the content in order to find them.

Concrete examples

25. Whatever abstract formulations may be devised to answer these questions, we are driven back to concrete examples. Would an intermediary be expected to make itself aware of the existence of a given example? Would that example require a decision or not? Why, in each case, is the answer yes or no? If an example requires a decision, how should that decision be reached? Is the decision clearly yes or no? If so, why is that so? If the decision is not clear is the default to be action, or no action - and why?

26. Some of the concrete examples that follow are taken from court cases. In one well-known case the court's view of the legality of the post changed as the case went through an appeal process. However, legality is not the only test. The government's proposals do not confine the notion of harm to illegality. They are largely concerned with content that carries a risk of harm. What constitutes harm? What constitutes risk of harm? (Is it a slight, small, medium, high or other risk of occurrence?) Is it that degree of risk to a single person, to a few people, to most people, to everyone? What characteristics of those notionally affected are to be assumed[10]?

27. These are legitimate questions when people's responses to the same material are highly variable and subjective, and the notion of what constitutes harm is malleable. Moreover, for the purposes of a duty of care not every harm (in the sense of detriment) is harmful (in the sense of meriting a legal response).

28. The mousetrap maker who suffers loss of business and consequential personal distress has no claim against the maker of the better mousetrap. As the Law Commission noted in its recent Consultation Paper on "Harmful Online Communications: The Criminal Offences":

> "Indeed, some harmful behaviour is permissible or even desirable. This is true of online communications as it is all other fields of human activity. Consider, for example, a highly effective online advertising campaign for a new brand of ecofriendly, ethically produced sports clothing. Such a campaign may harm the new brand's competitors, but is – all things being equal – at least permissible or even desirable." [5.11]

29. It is also noteworthy that, as matters stand, the government's proposals as regards 'lawful but harmful' content would place a duty of care on intermediaries that goes beyond lawful behaviour in the sense of a user merely not contravening a specific law. The duty of care would appear to apply to the intermediary even where the poster of the content would not itself owe any corresponding duty of care to avoid harm to its readers[11].

---

[10] For further discussion of this topic, see Section 5 of the author's Submission to the Online Harms White Paper Consultation: 'Speech is not a Tripping Hazard' (www.cyberleagle.com/2019/06/speech-is-not-tripping-hazard-response.html).
[11] See *Rhodes v OPO* [2015] UKSC 32.

30. A list of concrete examples follows. Some are real, some hypothetical. For convenience these are mostly framed as tweets, but any social media or shared post or blog would serve the purpose.

31. In the spirit of Ofcom having called for evidence, rather than for advocacy, the examples are presented without comment as to whether they should or should not trigger a duty of care.

32. To repeat, the purpose of these examples is less about what the answer is in any given case (although that is of course important in terms of whether the line is being drawn in the right place), but more about whether we are able to predict the answer in advance. If a legal framework does not enable us to predict clearly, in advance, what the answer is in each case, then there is no line and the framework falls at the first rule of law hurdle of "prescribed by law". It is not sufficient to make ad hoc pronouncements about what the answer is in each case, or to invoke high level principles. We have to know *why* the answer is what it is, expressed in terms that enable us to predict with confidence the answer in other concrete cases.

33. Some of the examples refer to S.127 Communications Act 2003. It may be pertinent that the Law Commission has recently described aspects of s.127 as "vague and ambiguous"[12].

[12] Consultation Paper on "Harmful Online Communications: The Criminal Offences", para 3.113.

A. An upsetting post

A social media user posts an explicit account of abuse that they suffered as a child. Reading the post would shock ordinary readers and could be traumatic for others who had suffered similarly. On the basis if the Supreme Court decision in *Rhodes* the post cannot, notwithstanding the distress that its contents might cause, be considered unlawful on the part of the person posting it.

B. Denial

(1) Someone uploads a social media post denying the existence of global warming.

(2) Someone tweets denying the efficacy of a recently launched Covid-19 vaccine.

(3) Someone posts a blog denying that the Holocaust took place.

(4) Someone posts a blog denying that dekulakisation in the 1930s Soviet Union involved large scale deaths.

(5) Someone posts a blog denying that large scale deaths occurred in the Chinese Cultural Revolution.

(6) Someone posts a blog denying that large scale deaths occurred under the Cambodian Khmer Rouge in the 1970s.

C. Threats

(1) Someone tweets "Crap! Robin Hood Airport is closed. You've got a week and a bit to get your shit together otherwise I am blowing the airport sky high!!".

This is the text of the tweet that was the subject of the Twitter Joke Trial. The author was prosecuted and convicted under S.127 Communications Act 2003, a decision that was subsequently overturned by the Divisional Court.

(2) Someone tweets: "Come, friendly bombs, and fall on Lunar House!"

(3) Someone tweets: "The first thing we do, let's kill all the activist lawyers! And that prat [X] is top of the list."

(4) Someone tweets: "This Tory Cabinet will be first against the wall when the revolution comes. And it can't come soon enough!"

(5) Someone tweets: "The only good Marxist is a dead Marxist."

(6) Someone tweets: "Burn, Baby, Burn!" in response to news of a protest march.

(7) Someone tweets: "Who will rid us of this turbulent priest?" in response to a news item about a radical Islamic cleric.

(8) Someone tweets: “I’d like to put a bomb under that eyesore of a shopping centre. And, with luck, take that mobile mast on the roof with it.”

(9) Someone tweets: “Judges are Enemies of the People and Stalin knew what to do with class enemies”.

D. Offensiveness

(1) An evangelical preacher uploads a video of a sermon containing the statement “Islam is heathen, Islam is satanic, Islam is a doctrine spawned in hell”.

(An uploaded video containing these words resulted in an unsuccessful prosecution of the preacher under S.127 Communications Act 2003.)

(2) Someone tweets a link to a clip of George Carlin performing “Seven Words You Can Never Say on Television”.

(3) Someone tweets a link to an uncut clip of The Major in the Fawlty Towers episode of “The Germans”.

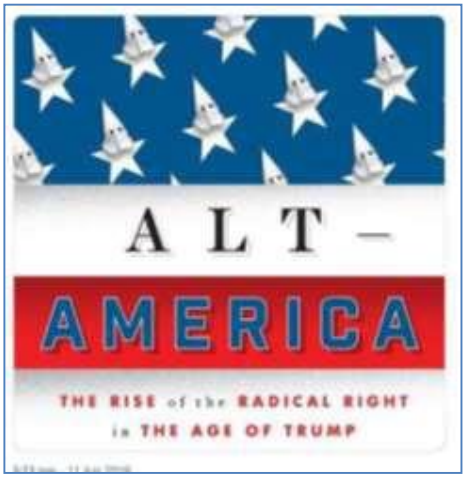
(4) Someone tweets a link to the High Court judgment in *R v Miller*, expressing support for the views contained in the Claimant’s tweets.

(5) Someone tweets a link to a newspaper report containing video footage of an effigy of Grenfell Tower being burned in a garden on Bonfire Night, 2018.

(6) Someone tweets: “Saudi Arabia has the right idea about women.”

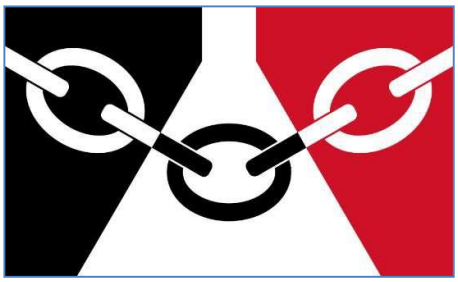
(7) Someone tweets a link to a picture of the ‘A brilliant idea hit her’ mug withdrawn from sale by Sainsbury’s in August 2020.

(8) This image[13] is included in a Twitter profile and a tweet.



[13] The illustration is from the cover of a book by David Neiwert.

(9) Someone posts a copy of the Flag of the Black Country:

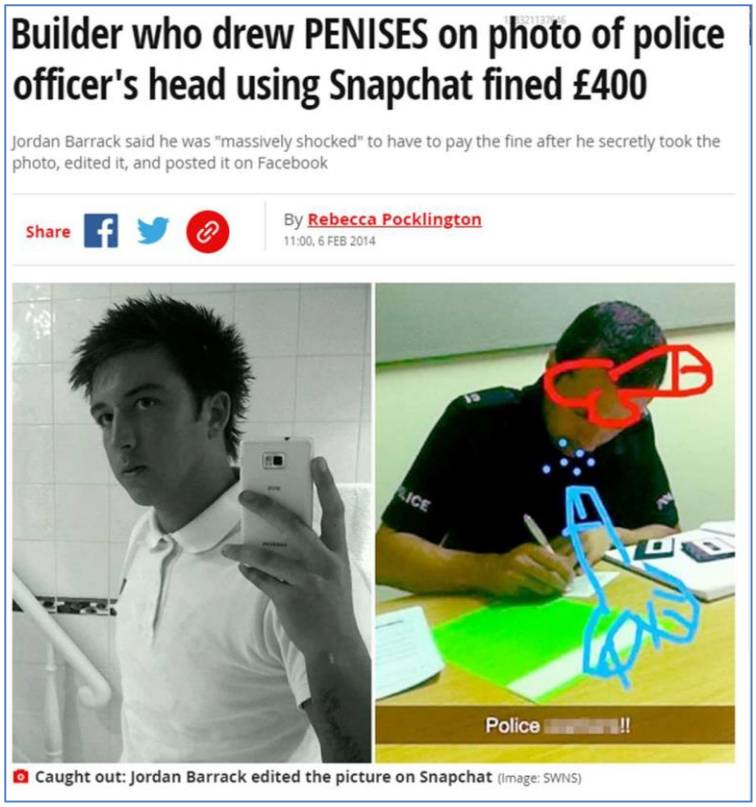


E. Humour and satire

(1) Someone tweets a link to a newspaper article dating from 2016, reporting the controversy about the 'Nazi pug' video that subsequently resulted in a conviction of Mark Meechan under S.127 Communications Act 2003. The article includes a playable link to the video.

(2) This 2014 Daily Mirror tweet linked to an article about a builder who was fined after pleading guilty to an offence under s.127 Communications Act 2003 resulting from a Facebook post.

The article included a copy of the photo, pixelated to remove certain text:



Builder who drew PENISES on photo of police officer's head using Snapchat fined £400

Jordan Barrack said he was "massively shocked" to have to pay the fine after he secretly took the photo, edited it, and posted it on Facebook

Share

By Rebecca Pocklington
11:00, 6 FEB 2014

Caught out: Jordan Barrack edited the picture on Snapchat (Image: SWNS)

An ordinary Twitter user tweets a link to the article.

(3) Someone tweets a link to an article about the 2005 Jyllands-Posten Muhammad cartoons controversy, which contains copies of the cartoons.

(4) Someone posts a link to the 2015 Muhammad cartoons published (and recently republished) by Charlie Hebdo.

(5) Someone posts a photograph of the cover of the book 'This is a Swedish Tiger'. (The satirical cover image is currently the subject of a Swedish prosecution of the book's author for copyright infringement.)